



***CAN'T FIND THE MOUSE MODEL YOU NEED?***

BUILD IT WITH  
**CRISPR!**

From simple indel KOs to more complex conditional KOs and reporter insertions, we've created a wide range of models using CRISPR/ Cas9 technology. If we aren't able to generate your CRISPR model, you don't pay.

**TO LEARN MORE,  
VISIT [JAX.ORG/CRISPR](http://JAX.ORG/CRISPR).**

1.800.422.6423

1.207.288.5845



**The Jackson  
Laboratory**

*Leading the search  
for tomorrow's cures*

# Cell

Best of 2016



Featuring a Collection of SnapShots

Seahorse Bioscience is  
now Agilent Technologies

# A MAJOR LEAP FORWARD IN PHENOTYPING

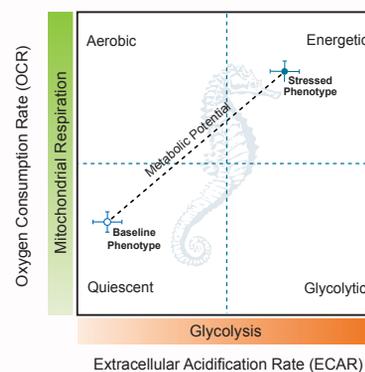
Determine the functional impact of genes, proteins  
or compounds on metabolism in an hour



## Agilent Seahorse XFp Cell Energy Phenotype Test Kit

Using as few as 15,000 live cells and an Agilent Seahorse XFp Analyzer, the Seahorse XFp Cell Energy Phenotype Test Kit identifies your cells' metabolic phenotype as well as their metabolic potential — the utilization of mitochondrial respiration and glycolysis.

**LEARN MORE** at [www.agilent.com](http://www.agilent.com)

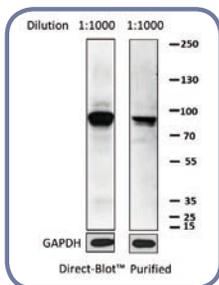


For Research Use Only. Not for use in diagnostic procedures.



## Make Cell Biology Research A Piece of Cake

BioLegend's Cell Biology products cover research areas including cellular metabolism, structure, communications, signaling, epigenetics, and the cell cycle, and applications such as western blotting, immunoprecipitation, and immunofluorescence microscopy.



Total protein extracts from Jurkat cells were probed with either Direct-Blot™ HRP anti-STAT1 (clone 10C4B40) or purified anti-STAT1 followed by secondary antibody.

Learn more at:  
[biolegend.com/cell\\_bio\\_cake](http://biolegend.com/cell_bio_cake)

### Featured Products

<b>Golgi:</b>	Giantin Antibody
<b>Centrosome:</b>	Pericentrin Antibody
<b>Cytoskeletal Filaments:</b>	Flash Phalloidin™ Localization Probes
<b>Nucleus:</b>	Nuclear Pore Complex Proteins and Lamin Antibodies
<b>Mitochondria:</b>	MitoSpy™ Localization Probes
<b>Functional Probes:</b>	TetraZ™ Cell Counting Kit
<b>Receptors and Signaling Proteins:</b>	gp130 Antibody STAT3 Antibody
<b>Cell Tracking:</b>	Tag-it™ Proliferation and Cell Tracking Probes



BioLegend is ISO 9001:2008 and ISO 13485:2003 Certified

Toll-Free Tel: (US & Canada): 1.877.BIOLEGEND (246.5343)

Tel: 858.768.5800

[biolegend.com](http://biolegend.com)

08-0061-06

World-Class Quality | Superior Customer Support | Outstanding Value

# Identifying and Isolating Mesenchymal Stem Cells from Bone Marrow Cultures

## Reliably Sort Stem Cell Populations with the Bio-Rad S3e™ Cell Sorter

Xiaoti Guo, PhD, Bio-Rad Laboratories, Inc., 2000 Alfred Nobel Drive, Hercules, CA 94547



Bone marrow (BM) mesenchymal stem cells (MSCs) are multipotent adult stem cells with the capacity to differentiate into cells that compose adipose, bone, cartilage, and muscle tissue, thus presenting a wide potential for cell-based therapies. According to the minimal criteria published by the International Society for Cellular Therapy and multiple studies (Boxall and Jones 2012, Dominici et al. 2006), CD73, CD90, and CD105 are the primary markers used to identify MSCs across species, and their expression levels tend to be stable during culture. In contrast, CD106 (VCAM-1) and STRO-1 have been shown to represent the most potent and undifferentiated MSCs, and they exhibit reduced expression levels after multiple passages (Boxall and Jones 2012, Halfon et al. 2011, Liu et al. 2008). For now, cell sorting by flow cytometry remains the most efficient approach to isolate MSCs from a highly heterogeneous environment like bone marrow. In this study, the S3e Cell Sorter was used to analyze and purify prospective MSCs from passage 1 (P1) and passage 3 (P3) rat bone marrow cultures. Sorting performance was evaluated and results were compared between P1 and P3. The data presented here show the capability of the S3e Cell Sorter to isolate rare MSCs from a mixed bone marrow culture with satisfying sort efficiency and postsort purity.

### Efficient Isolation of MSCs from a Heterogeneous Culture

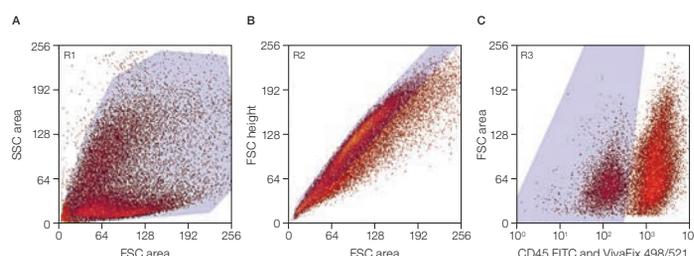
To identify potential MSCs and to test the capability of the S3e Cell Sorter to purify MSCs from heterogeneous BM culture, P1 rat BM cells were stained for CD90, CD73 or CD106, and CD45, and then with the viability dye VivaFix™ 498/521. Stained cells were analyzed on the S3e Cell Sorter equipped with a 100 µm nozzle tip and 30 psi sheath pressure. Viable MSCs and non-MSCs were then sorted on the S3e Cell Sorter and assessed for postsort purity. Representative graphs in Figure 1 illustrate the following gating strategy. Since P1 bone marrow displays a relatively high level of heterogeneity, a broad gate was applied to include both SSC<sup>low</sup> and SSC<sup>high</sup> populations (Figure 1A). Following doublet exclusion (Figure 1B) viable (VivaFix<sup>dim</sup>) and CD45<sup>neg</sup> populations were gated for stem cell surface marker expression (Figure 1C). As shown in Figure 2, about 15% of gated cells stain positive for both CD90 and CD73 (Figure 2B) and 3% coexpress CD90 and CD106 (Figure 2C), as determined by the phycoerythrin (PE) fluorescence minus one (FMO) control (Figure 2A). The dual-positive cells are prospective MSCs and the dual-negative ones are likely contaminating hematopoietic and stromal cells. Both prospective MSCs and non-MSCs were then sorted on the S3e Cell Sorter at up to 1,500 events per second (eps) using Purity mode. After sorting, a CD90+ CD73+ population was enriched from 15 to 95% (Figure 2D), and CD90+

CD106+ from 3 to 92% (Figure 2E). Non-MSCs were enriched from 40 to 99% (Figures 2F and 2G).

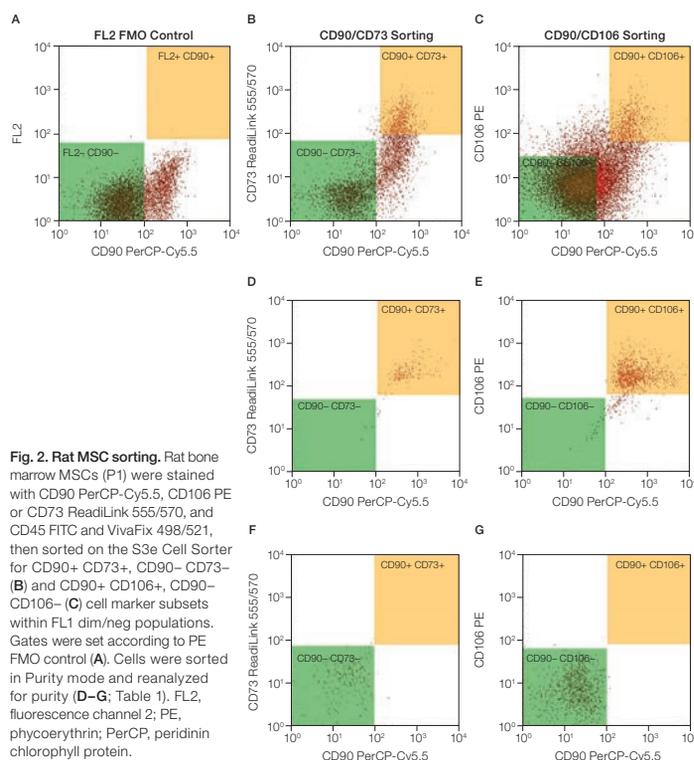
The sorting performance for each population is summarized in Table 1, which represents results from three independent experiments. Theoretical efficiencies are calculated for a 100 µm nozzle tip at 30 psi sheath pressure and various input rates and target cell frequencies with CytoCalc Software v3.2.

### P3 Cells Reveal Less Phenotypic Heterogeneity

To test whether rat bone marrow MSCs are able to maintain their phenotypic characteristics after in vitro expansion, P1 and P3 bone marrow MSCs were stained and analyzed side by side on the S3e Cell Sorter to compare their expression of MSC-related surface markers (Figure 3 and Table 2). Compared with P1 MSCs, P3 cells revealed less heterogeneity in morphology as determined by microscopy, and the SSC<sup>high</sup> population increased from 50% in P1 to 90% in P3 (data not shown). These more homogeneous P3 cells expressed a greater level of CD90 compared to P1 cells (86 vs. 56%), but the CD73 and CD106 expression within the CD90+ subset was drastically reduced (from 23 to 0.9% for CD73+; 6.4 to 0.02% for CD106+). These data suggest that although in vitro expansion may help eliminate contaminating cells from the culture, it may also compromise stem cell phenotype by downregulating stem cell marker expression.



**Fig. 1. Rat MSC gating strategy.** Rat bone marrow MSCs (P1) were stained with CD90 PerCP-Cy5.5, CD106 PE or CD73 ReadLink 555/570, and CD45 FITC and VivaFix 498/521, then acquired on the S3e Cell Sorter. After excluding debris and doublets (A–B), the fluorescence channel 1 (FL1) dim population was selected for further analysis (C). FITC, fluorescein isothiocyanate; FSC, forward scatter; PE, phycoerythrin; PerCP, peridinin chlorophyll protein; SSC, side scatter.



**Fig. 2. Rat MSC sorting.** Rat bone marrow MSCs (P1) were stained with CD90 PerCP-Cy5.5, CD106 PE or CD73 ReadLink 555/570, and CD45 FITC and VivaFix 498/521, then sorted on the S3e Cell Sorter for CD90+ CD73+, CD90– CD73– (B) and CD90+ CD106+, CD90– CD106– (C) cell marker subsets within FL1 dim/neg populations. Gates were set according to PE FMO control (A). Cells were sorted in Purity mode and reanalyzed for purity (D–G; Table 1). FL2, fluorescence channel 2; PE, phycoerythrin; PerCP, peridinin chlorophyll protein.

## Reliable Purification of BM-Derived MSCs

MSCs are rare in their resident tissues; the frequency of MSCs in BM is 0.001–0.01%, diluted in a heterogeneous mix of hematopoietic cells (Bernardo et al. 2009). When MSCs are isolated, a heterogeneous pool of cells with various proliferation and differentiation potentials can be obtained (Bianco et al. 2001). The rareness and heterogeneity together make it difficult to identify, isolate, and purify MSCs. The adherent culture or panning approach is able to eliminate some of the hematopoietic cells; however, flow cytometry-based cell sorting remains the most reliable approach to isolating a pure population of MSCs for downstream cellular and molecular characterization.

Using a 100  $\mu\text{m}$  nozzle tip, the S3e Cell Sorter was able to analyze and identify BM-derived MSCs from a mixed BM culture and sort target cells at satisfying purity and efficiency. Cells were sorted at up to 1,500 eps with 30 psi sheath pressure and 41 kHz drop drive frequency, and the sort efficiency ranged from 91 to 98%. Note that the *in vitro* expanded bone marrow MSCs are relatively large (20–35  $\mu\text{m}$ ) for the 100  $\mu\text{m}$  orifice. Therefore, to minimize disturbance of the break-off point and fanning of deflected streams, sort performance at a higher speed was not assessed at current settings. A nozzle tip with a larger diameter would be more suitable for MSC sorts in future studies.

## Dual Sorting Modes on the S3e Cell Sorter May Further Improve Purity and Yield

While non-MSCs were sorted at close to 100% purity, CD90+ CD73+ and CD90+ CD106+ dual-positive MSCs achieved only 95 and 92% purity, respectively. There are two possible reasons for this: a majority of the CD90+ cells are larger in size (SSC<sup>high</sup>) whereas CD90– cells make up the majority of the SSC<sup>low</sup> population. The large cells (relative to the orifice diameter of the nozzle tip) likely have compromised sort purity due to insufficient drop charge and excessive side stream fanning (Arnold and Lannigan 2010, Petersen et al. 2003). Another possible factor contributing to the suboptimal purity is that dual-positive cells are found with lower frequency (3% of target population) than dual-negative ones (40–50%), and sorting low abundance populations can result in lower purity and yield unless they are pre-enriched (Arnold and Lannigan 2010, Kuka and Ashwell 2013). The mixed mode sorting feature of the S3e Cell Sorter allows sorting of abundant populations in Purity mode and rare populations in Enrich mode. Hence, presorting rare populations in the Enrich followed by Purity modes will likely help improve MSC purity and yield.

## Better Cell Culture and Isolation Methods Needed to Minimize Loss of Stem Cell Phenotypes

The P1 cells used in this study were the first passage of Sprague-Dawley (SD) rat BM, in which a significant amount of contaminating hematopoietic cells remained in the culture, contributing mainly to the CD90– population. After two more passages in 5 weeks, most of the hematopoietic cells were removed from the culture, which greatly increased the frequency of the CD90+ population in P3. But the expression of two other stem cell markers, CD73 and CD106, drastically decreased in P3 cells. This finding is consistent with other reports that showed that CD106 expression in MSCs was downregulated by tenfold after 2 weeks of osteogenic differentiation (Liu et al. 2008). Even without induction, the expression of CD106 could also decrease at a later passage compared with an earlier culture (Halfon et al. 2011). As it is a surface marker considered to be linked with stem cell differentiation potential,

a reduction in CD106 expression may indicate compromised stem cell multipotency during *in vitro* culture (Boxall and Jones 2012, Liu et al. 2008). This could be caused by a serum component in the culture medium, which induces stem cell differentiation, or simply by the fact that cells with less potential outgrow the multipotent MSC subset. As an alternative approach, single cell sorting using Single mode on the S3e Cell Sorter can be used to separate the different stem cell subpopulations. The single colony-derived cells can then be cultured and evaluated separately.

## Conclusions

With the use of appropriate surface markers, the S3e Cell Sorter is capable of identifying bone marrow MSCs from a mixed BM culture and sorting target cells at satisfying purity and efficiency. Compared to P1 cells, P3 rat MSCs display a more homogeneous phenotype with an increased level of CD90 but decreased levels of CD90+ CD106+ and CD90+ CD73+ stem cell subsets, suggesting that prolonged culture may help remove contaminating cells but also could compromise stem cell phenotype and possibly affect their pluripotency. An improved long-term cell culture strategy should be investigated to minimize stem cell differentiation before therapeutic applications of MSCs can be realized.

## Acknowledgements

The author would like to acknowledge Drs. Haixiang Liang and Daniel A Grande (The Feinstein Institute for Medical Research) for kindly providing P1 and P3 rat MSCs, Drs. Qing Xiang and Ralph J Garippa (Memorial Sloan Kettering Cancer Center) for their generous support with equipment, Michael Kissner (Propel Labs) for his helpful technical discussions, and Brandon Williams and Melissa Ma (Bio-Rad Laboratories) for their help with manuscript preparation.

## References

- Arnold LW and Lannigan J (2010). *Curr Protoc Cytom* Chapter 1:Unit 1.24.1–1.24.30.
- Bernardo ME et al. (2009). *Ann N Y Acad Sci* 1176, 101–117.
- Bianco P et al. (2001). *Stem Cells* 19, 180–192.
- Boxall SA and Jones E (2012). *Stem Cells Int* 2012, 975871.
- Dominici M et al. (2006). *Cytotherapy* 8, 315–317.
- Halfon S et al. (2011). *Stem Cells Dev* 20, 53–66.
- Kuka M and Ashwell JD (2013). *J Immunol Methods* 400–401, 111–116.
- Liu F et al. (2008). *J Bone Miner Metab* 26, 312–320.
- Petersen TW et al. (2003). *Cytometry A* 56, 63–70.

Cy is a trademark of GE Healthcare. CytoCalc is a trademark of Applera Corporation. ReadLink is a trademark of AAT Bioquest, Inc.



Bio-Rad Laboratories, Inc.

Visit [bio-rad.com/info/S3eMSC-ASCB](http://bio-rad.com/info/S3eMSC-ASCB) for more information.

Table 1. Sort performance.\*

Cell Marker Subset	Presort Frequency (% of parent)	Postsort Purity (% of parent)	Sort Efficiency, %	Theoretical Efficiency, %**
CD90+	40	99	98	98.2
CD90+ CD73+	15	95	91	97
CD90– CD73–	40	97	94	97.8
CD90+ CD106+	3	92	96.2	96.6
CD90– CD106–	50	99	92	98.4

\* The size of rat MSCs range from 20 to 35  $\mu\text{m}$ . P1 MSCs were sorted in Purity mode at 1,000–1,500 eps with a 100  $\mu\text{m}$  nozzle tip, 30 psi sheath pressure, and 41 kHz drop drive frequency. Data represent three independent experiments.

\*\* Theoretical efficiencies are calculated with CytoCalc Software v3.2 for a 100  $\mu\text{m}$  nozzle tip at 30 psi sheath pressure at various input rates and target cell frequencies.

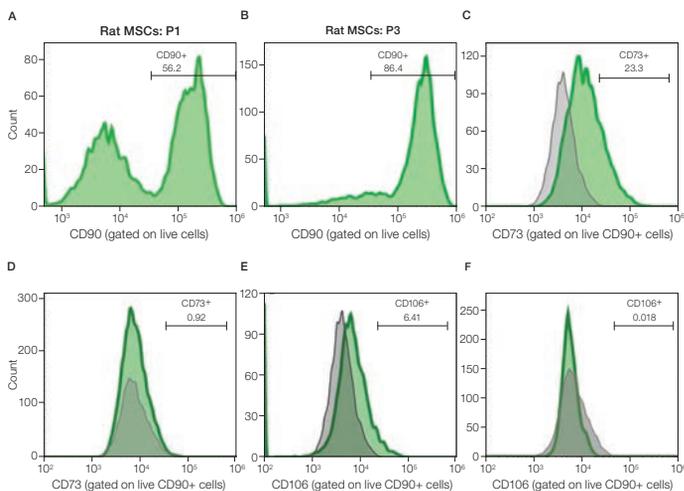


Fig. 3. Characterization of P1 vs. P3 rat bone marrow MSCs. Rat bone marrow MSCs were stained with anti-CD90 PE-Cy7, CD106 PE or CD73 ReadLink 555/570, and CD45 FITC and VivaFix 498/521, then acquired on the S3e Cell Sorter. Live CD90+ (A–B), CD90+ CD73+ (C–D), and CD90+ CD106+ (E–F) subsets were compared between P1 and P3 MSCs. Gray traces indicate FMO control for the PE channel. The frequency of each cell marker subset is shown in Table 2.

Table 2. Frequency of P1 vs. P3 rat bone marrow MSCs.

Cell Marker Subset	P1 Frequency, %	P3 Frequency, %
CD90+	56	86
CD90+ CD73+	23	0.9
CD90+ CD106+	6.4	0.02

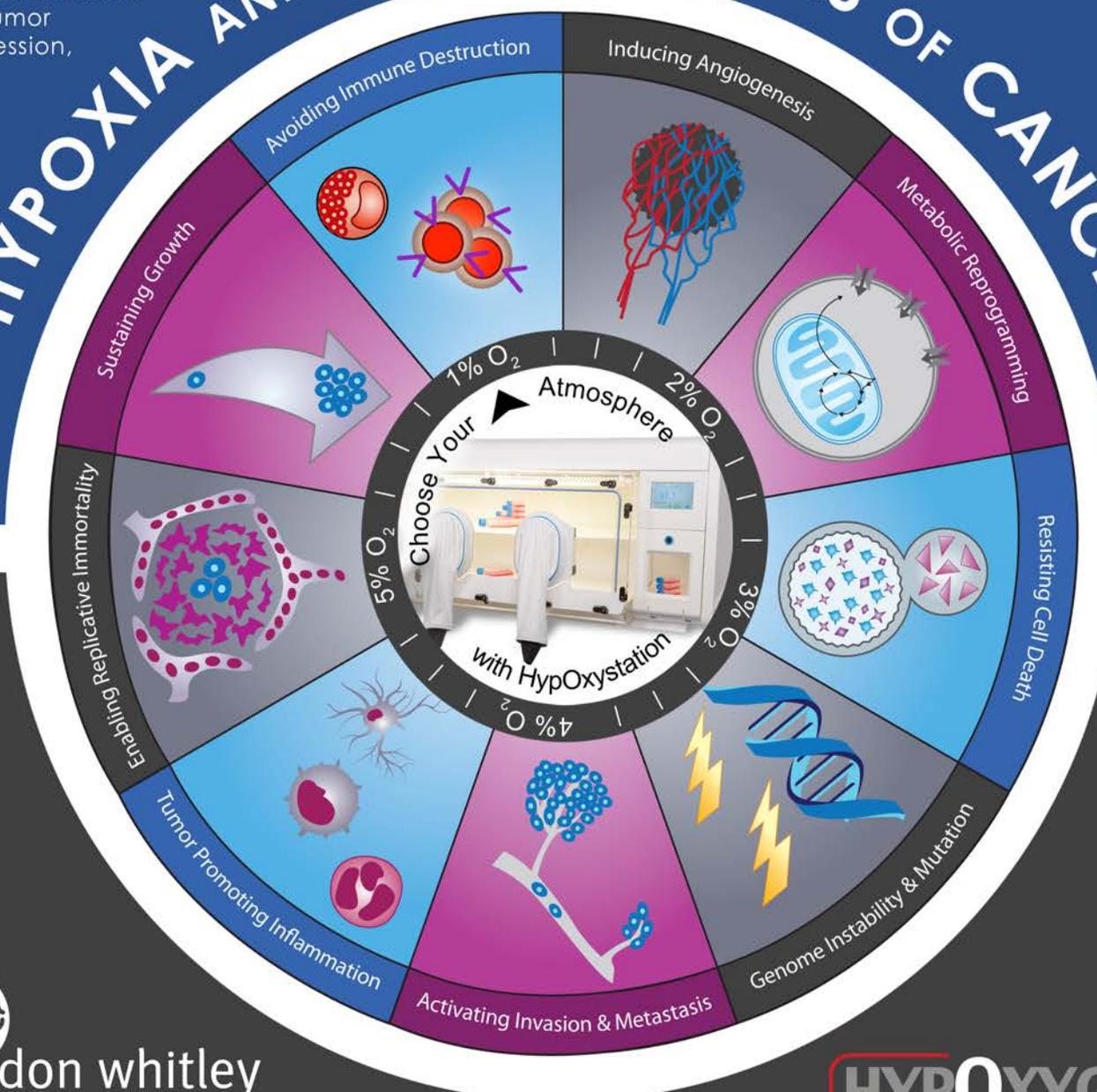
The iconic "Hallmarks of Cancer" first described by Douglas Hanahan and Robert Weinberg in 2000 continue to develop explosively as efforts to illuminate the "acquired capabilities of cancer" evolve. Hypoxia in the fast-growing, poorly perfused tumor setting is one of the main factors determining survival of cancer cells and driving mutagenesis and metastasis. Oxygen in the tumor microenvironment and within the cells is central to metabolism, immunology, epigenetics and therapy resistance of all the cancers.

Physiological low oxygen and hypoxia in cell culture exert effects on metabolism, maintenance, cell yield, and cell survival. Cell culture in the **HypOxystation** establishes authentic conditions for research into tumor progression,

differentiation and self-renewal of cancer stem cells, response to chemotherapeutic drugs and other events which determine malignancy and outcome of cancer diseases. The **HypOxystation** enables glove-less access to cultivate and manipulate cells under physiological and hypoxic conditions, in a HEPA-clean environment.

Wigerup and Pahlman, **HypOxystation** users in Sweden, recently published a brilliant review on "Therapeutic Targeting of Hypoxia and Hypoxia-Inducible Factors in Cancer" (Pharm & Ther. 2016, 164: 152-169). They describe how "tumor hypoxia affects most of the Cancer Hallmarks" and give an overview of strategies to target the hypoxic phenotype to eradicate cancer.

# HYPOXIA AND THE HALLMARKS OF CANCER



don whitley  
scientific

www.dwscientific.co.uk

**HYPOXYGEN**

Define your environment

www.HypOxygen.com



# Impress Yourself

## The new Eppendorf Cell Culture Consumables

The all new line of Eppendorf Cell Culture Consumables will truly delight your cells. The outstanding design, reliability and purity is based on more than 50 years of experience. Products created by experts, developed for perfectionists. Impress yourself!

- > Unsurpassed quality, clarity, purity and sterility, providing reliable cell culture conditions
- > Significantly improved design for more safety and consistency
- > Maximum safety and confidence during storage and transportation



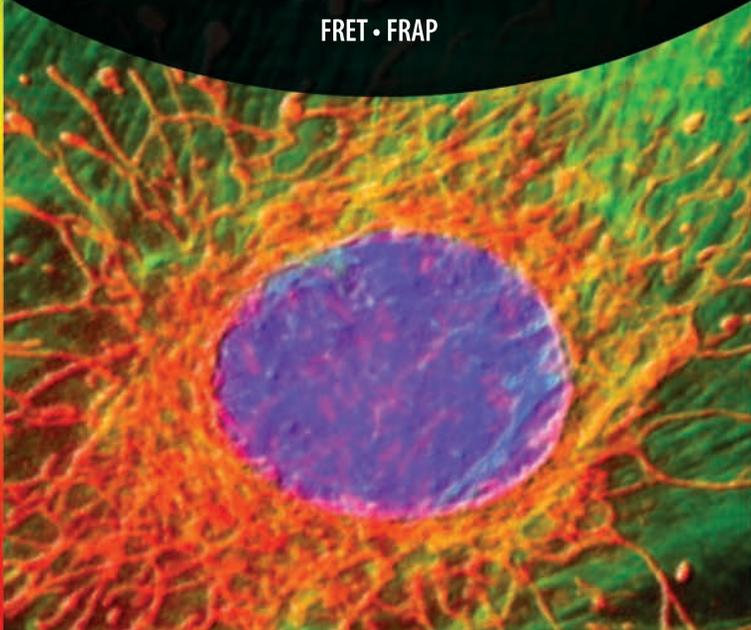
[ccc.eppendorf.com](http://ccc.eppendorf.com) • 800-645-3050

**ADVANCED FLUORESCENCE  
IMAGING SYSTEMS**

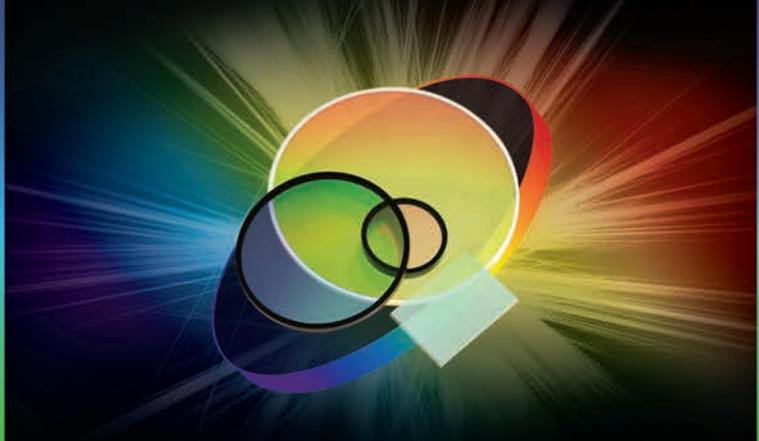


Imaging products from 89 North, Cairn Research  
and CrestOptics create light re-engineered.

Confocal Imaging • Super-Resolution Imaging  
High-Speed Ratiometric Imaging • Optogenetics • Photoactivation  
Photoconversion • Simultaneous Multichannel Imaging  
FRET • FRAP

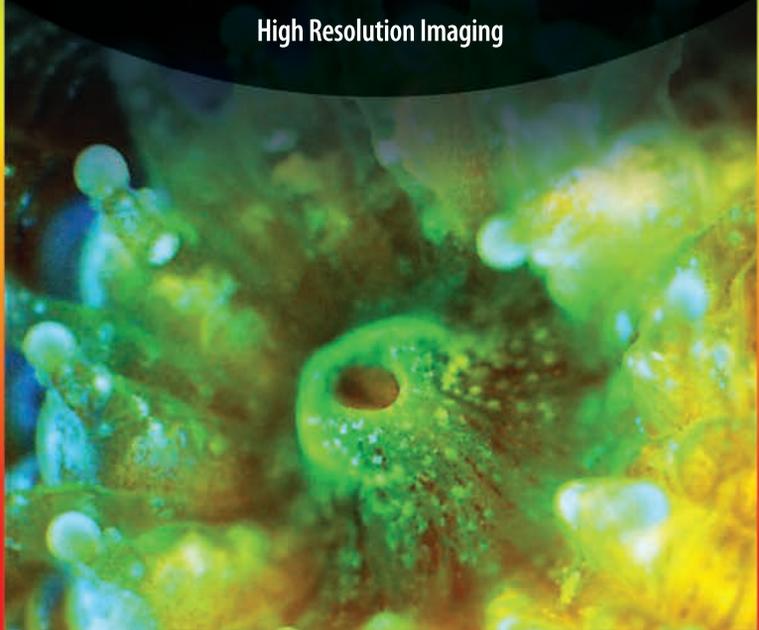


**DURABLE HIGH TRANSMISSION  
SPUTTERED FILTERS FROM UV TO NIR**



From super-resolution to Raman spectroscopy,  
Chroma filters give you more options.

Flow Cytometry • Machine Vision  
Point-of-Care • Laser Applications  
Raman • Fluorescence • Astronomy  
High Resolution Imaging



# Expedite Your Research

by partnering with our industry-renowned product development and biomarker testing scientists to deliver unique reagents and reliable data.

- Get custom products tailored to your exact specifications
- Outsource troubleshooting and QA oversight to the experts
- Deliver your project on time and on budget

## Custom Services

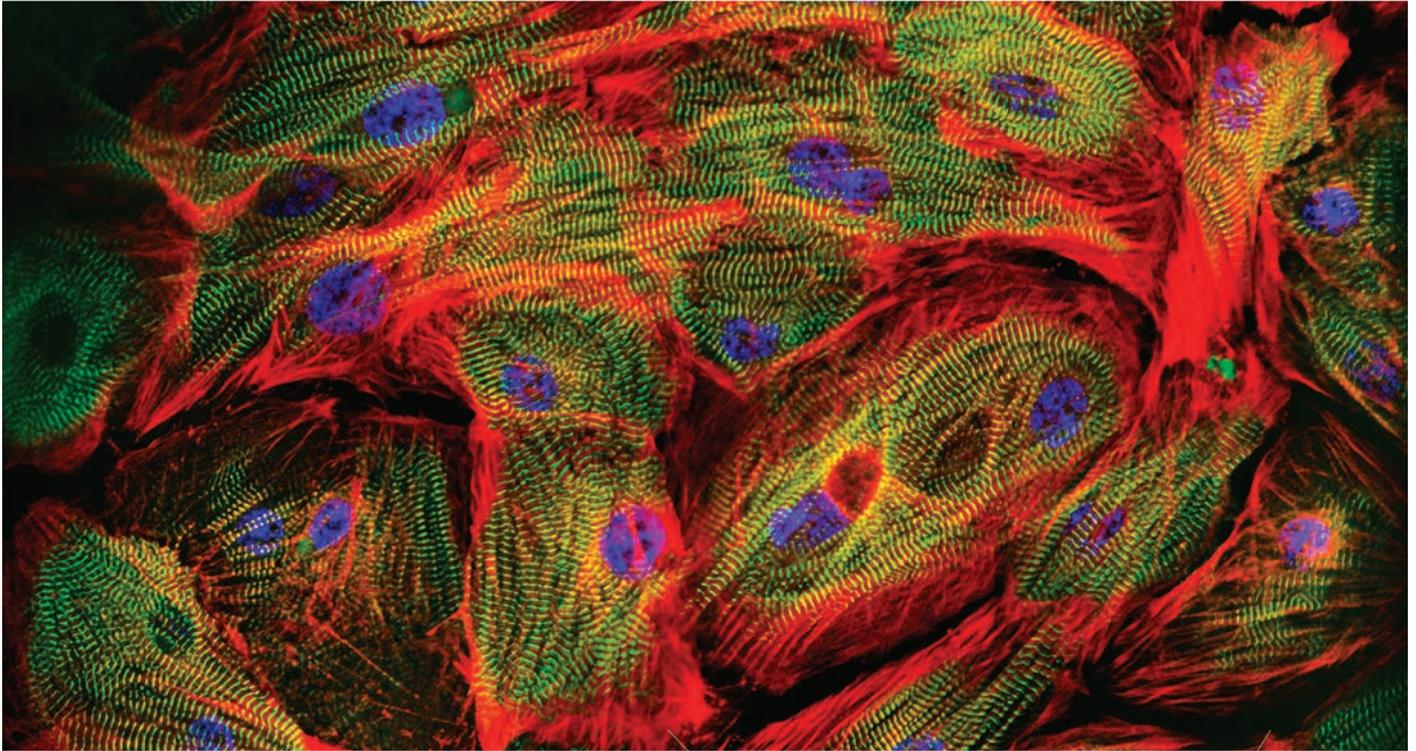


*Setting the standard in quality research reagents and support for over 30 years.*

Learn more | [rndsystems.com/customservices](http://rndsystems.com/customservices)



**CELLular**  
**Dynamics**  
international  
a FUJIFILM company



## Human Cells.

### Endless Possibilities from Discovery to Regenerative Medicine.

Cellular Dynamics International (CDI) has designed an induced pluripotent stem cell (iPSC)-based technology platform that enables the large scale production of high-quality, fully functioning human cells. Stem cells are early stage, undifferentiated cells that are found in all multicellular organisms. They are unique because they can survive and replicate indefinitely. CDI's technology offers the potential to create iPSCs from anyone, starting with a standard blood draw, and followed by the powerful capability to develop into virtually any cell type in the human body.

Our iCell® and donor-specific MyCell® products are highly pure, highly reproducible, and available in industrial quantity to enable drug discovery, toxicity testing, stem cell banking, and cell therapy development.

[cellulardynamics.com](http://cellulardynamics.com)

# Foreword



Four years ago, Cell Press launched the “Best of” reprint collections across a number of our journals, including *Cell*. We are pleased to now bring you a new *Best of Cell* that focuses on articles published over the course of 2016. For this edition, we made our selection by looking at review and research articles with the most full-text HTML and PDF downloads since publication, then choosing a representative group from each of the four published volumes. We’ve also included a selection of frontmatter articles and a collection of SnapShots that highlight some of the most impactful biology that we’ve had the privilege to publish this year. This collection covers a broad range of topics and includes papers that feature some of the exciting experimental approaches, such as genome engineering with CRISPR, disease modeling with organoids, an examination of the cellular phase of Alzheimer’s, and breaking cryo-EM resolution boundaries to facilitate drug discovery, that have helped drive science forward in recent years.

We recognize that no one measurement can be indicative of the “best” papers over a given period of time, especially when the articles are relatively new and their true significance may still need time to be established. We’ve selected several papers that were published later in the year and that may not have had the time to develop quite as significant usage statistics but show promise as groundbreaking in their own right. We hope that this combination of approaches to highlighting articles will give you an overview of different perspectives on the science that we published during 2016.

We hope that you will enjoy reading this special collection and that you will visit [www.cell.com/cell](http://www.cell.com/cell) to see the latest findings that we have had the privilege to publish. Also, be sure to visit [www.cell.com](http://www.cell.com) to find other high-quality articles published in the full collection of Cell Press journals. If you prefer to receive your updates via social media, you may want to follow us on Twitter ([@CellCellPress](https://twitter.com/CellCellPress)) or like our Facebook page ([www.facebook.com/CellCellPress/](https://www.facebook.com/CellCellPress/)).

Last but definitely not least, we are grateful for the generosity of our many sponsors, who helped to make this reprint collection possible.



**For information for the Best of Series, please contact:**  
Jonathan Christison  
Program Director, Best of Cell Press  
e: [jchristison@cell.com](mailto:jchristison@cell.com)  
p: 617-397-2893  
t: [@cellPressBiz](https://twitter.com/cellPressBiz)



# Discover what's beneath.

Explore the depths of amplicon sequencing with the NEXTflex™ Amplicon Panels. The NEXTflex Amplicon Panels are complete solutions for your targeted resequencing needs. These panels are designed to detect and screen relevant germline and somatic mutations using either Illumina or Ion Torrent-based sequencing platforms. Visit [BioScientific.com/panels](http://BioScientific.com/panels) to learn how the ready-to-use predesigned amplicon panels or validated NEXTflex™ custom solutions simplify targeted sequencing protocols.



#### AMPLICON PANELS

- Custom
- Mediterranean Fever
- BRCA 1/2
- Obesity
- Colon
- TP53
- CEBPA
- Phenylketonuria

*Your Discovery Awaits.*  
Visit [bioscientific.com/panels](http://bioscientific.com/panels) to learn more.

For research use only. Not for use in diagnostic procedures.

THE NGS EXPERTS™

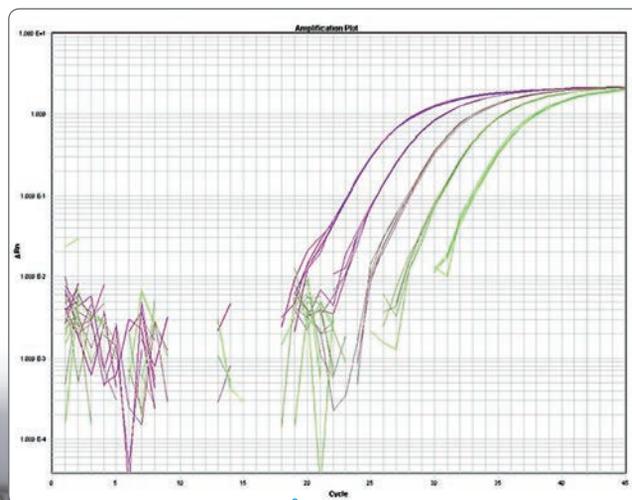
 **BIOO SCIENTIFIC®**  
a PerkinElmer company



# PipetmaX<sup>®</sup> + qPCR Assistant

Now everyone in your lab can automate qPCR protocols

- ▶ Easy to set up, no programming required
- ▶ Create flexible, single and multiplex qPCR methods
- ▶ Eliminate data entry errors with sample import and export
- ▶ Save time by combining multiple methods into one run



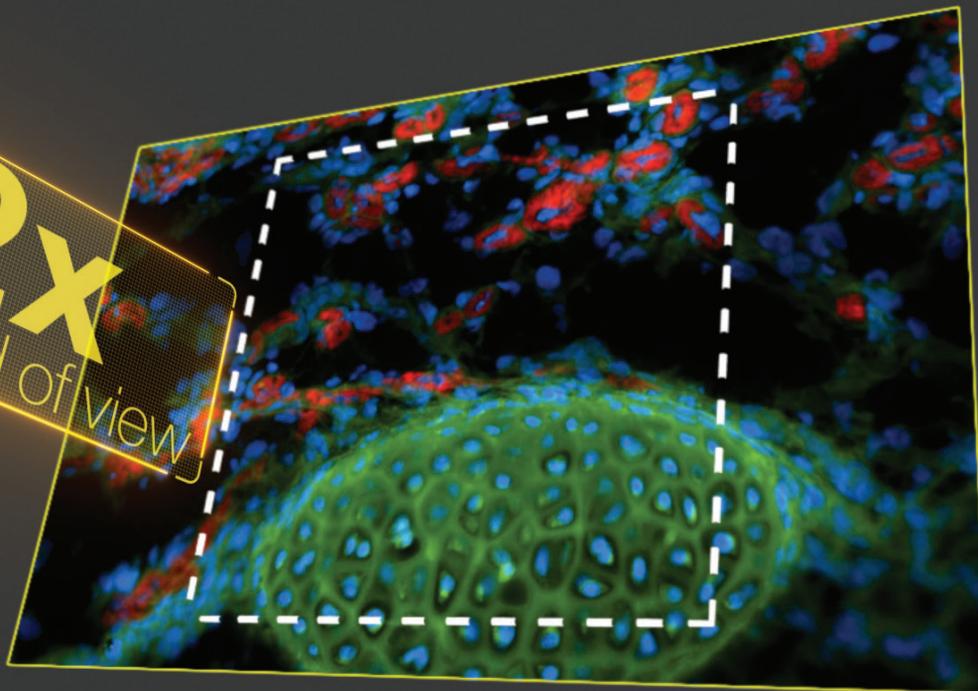
▶ Standard Amplification Curves for COX gene. Standard cDNA samples were serially diluted 5-fold and amplified in 5 replicates (dilutions left to right: 5x, 25x, 125x, 625x and 3125x).



instruments • accessories • software • solutions ▶ Meet PipetmaX and your new qPCR Assistant at [www.gilson.com](http://www.gilson.com)  
Visit [www.gilson.com](http://www.gilson.com) to download Speaking Volumes, our interactive publication for your tablet!

# NEW!

2x  
the field of view



## See More Than Before

Introducing the Eclipse Ti2, Nikon's all-new inverted microscope platform for advanced imaging.

Nikon Instruments has raised the bar for core imaging capability with the release of the new Eclipse Ti2 Inverted Research Microscope. With its unprecedented 25mm field of view, you can capture twice as much data with a single image. The fourth generation Perfect Focus System combined with a completely redesigned Z-drive provide a super-stable imaging platform with high-precision focusing for demanding applications. Built-in sensors and intelligent Assist Guide eliminate user error while the internal detector enables you to see the back-aperture for alignments.

There's so much more! For more information, contact your local Nikon sales representative, call 1-800-52-NIKON (1-800-526-4566) or visit [www.nikon-ti2.com](http://www.nikon-ti2.com). The new microscope will also be featured at the Nikon booth at the ASCB Annual Meeting.



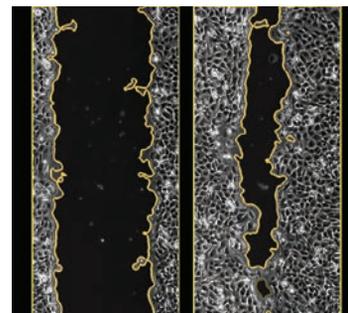
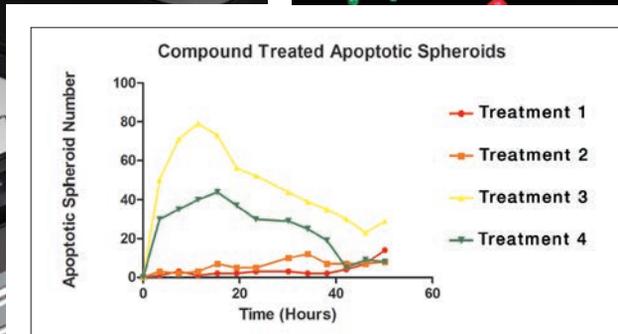
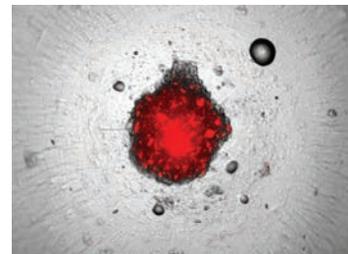
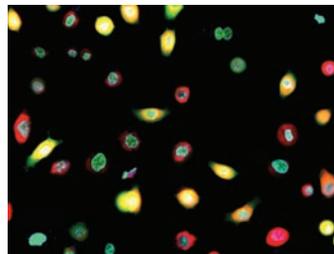
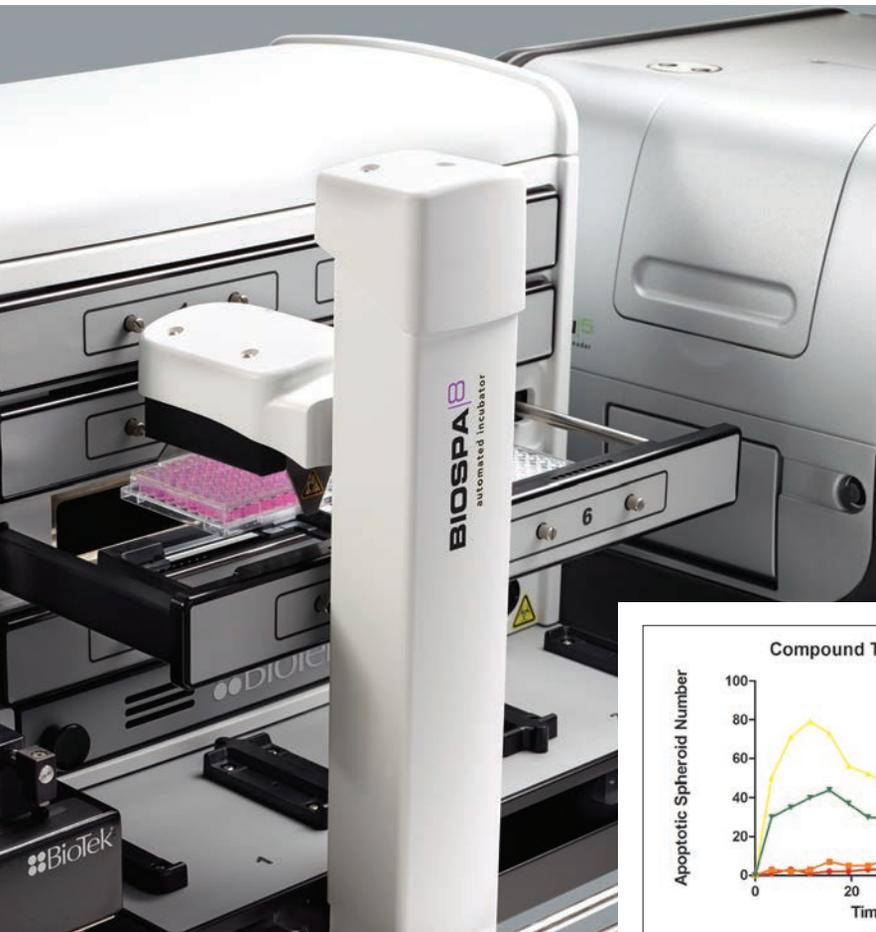
**ECLIPSE**  
**Ti2**

 nikon-instruments-inc.  nikoninstrumentsinc

Nikon Instruments Inc. ■ [www.nikoninstruments.com](http://www.nikoninstruments.com) ■ [nikoninstruments@nikon.net](mailto:nikoninstruments@nikon.net)



# BIOSPA SYSTEM FOR LIVE CELL IMAGING



## More Than an IncuCyte®

	BioSpa™ System	IncuCyte® Zoom
Live cell imaging	✓	✓
Temperature, humidity, gas control	Integrated, monitored	Incubator required
Labware capacity	8	6
Objective capacity	6	1
Magnification range	1.25x to 60x	4x to 20x
Fluorescence colors available	>15	2
Z-stacking for 3D biology	✓	
Laser autofocus	✓	
Advanced cell-level analysis	✓	
Washing, media replacement	✓	
Reagent dispensing	✓	

\*BioSpa system includes BioSpa 8 and Cytation 5

Think Possible



Learn more. Visit [www.biotek.com/biospasystem](http://www.biotek.com/biospasystem)

IncuCyte® ZOOM is a registered trademark of Essen BioScience

[www.biotek.com](http://www.biotek.com)

PHYSIOLOGICAL  
CELL CULTURE  
WORKSTATIONS

CULTURE AS  
NATURE INTENDED

**BAKER** RUSKINN

**Invivo<sub>2</sub>**

TOGETHER

WE CAN MAKE

MORE THAN

A DIFFERENCE

WE CAN  
MAKE HISTORY

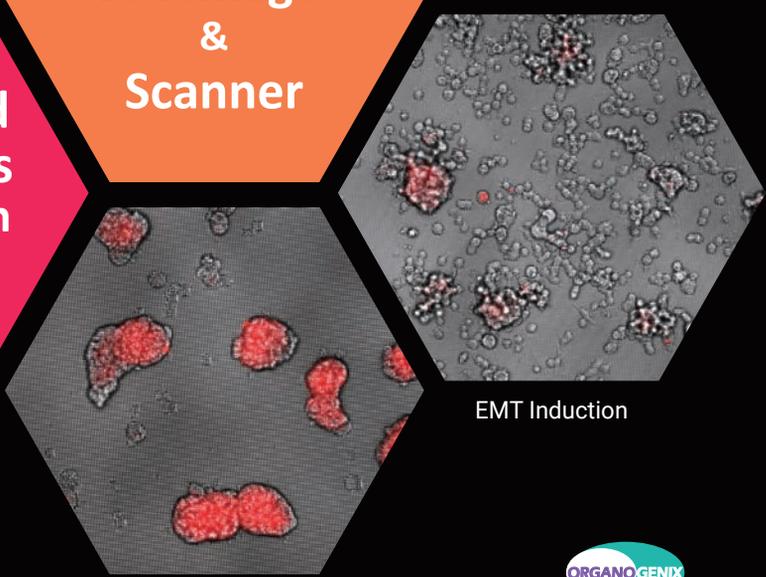
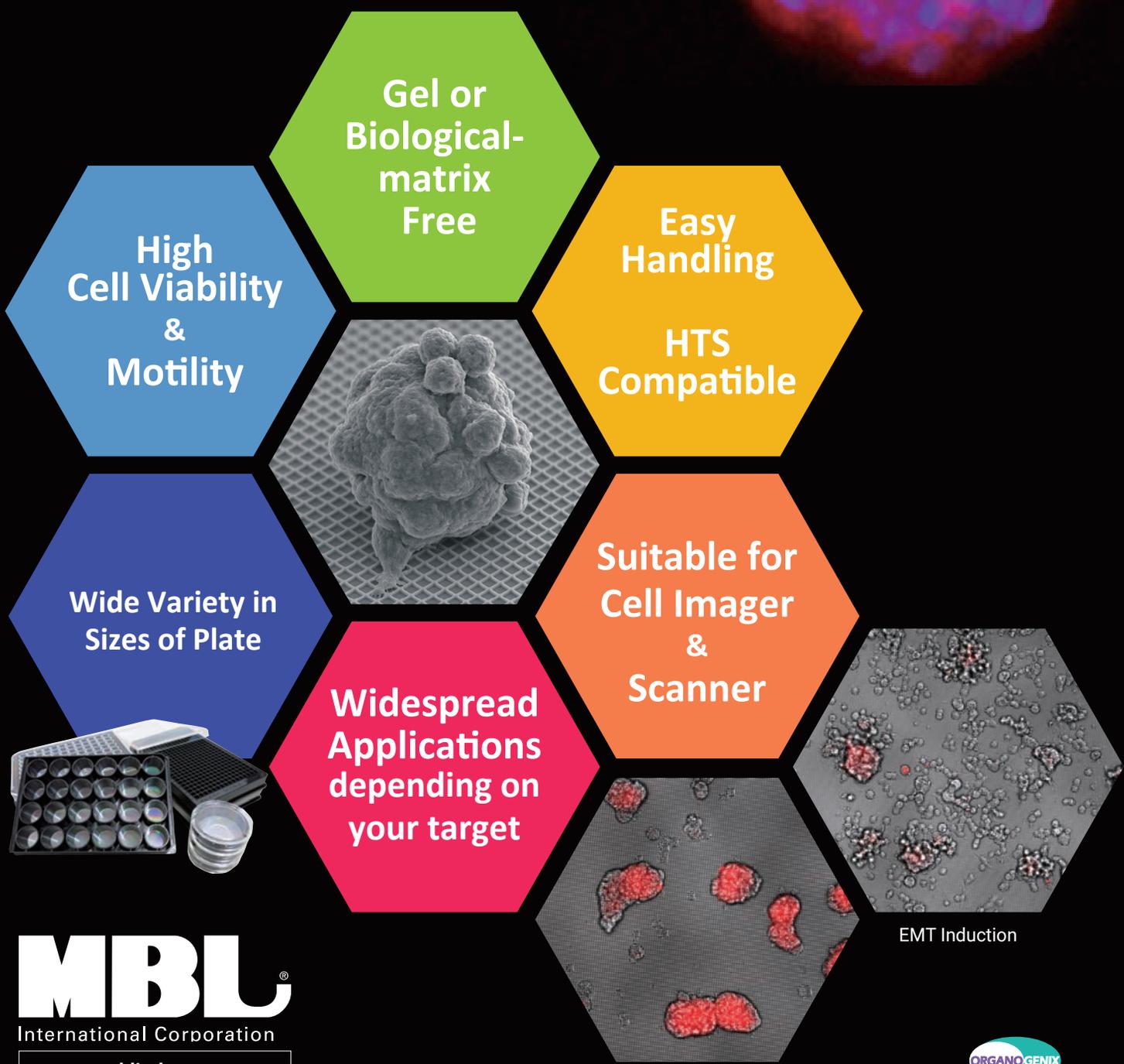
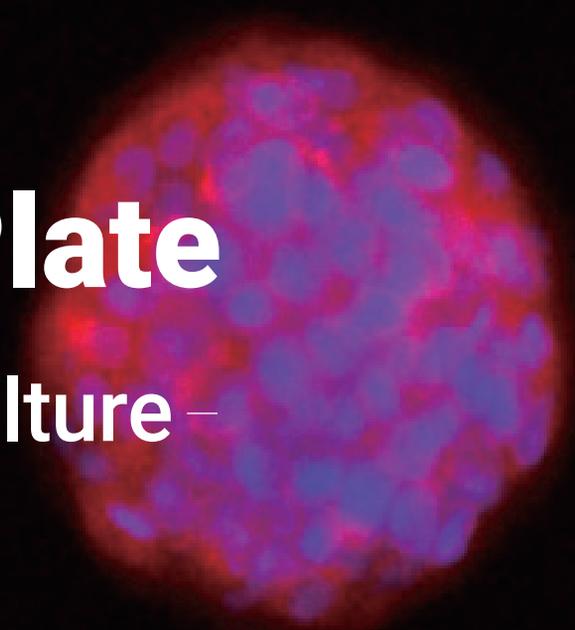
Since 1998 Baker Ruskin have been working with pioneering scientists around the world to make breakthroughs in cell research.



# 3D cell culture system

# Nano Culture Plate

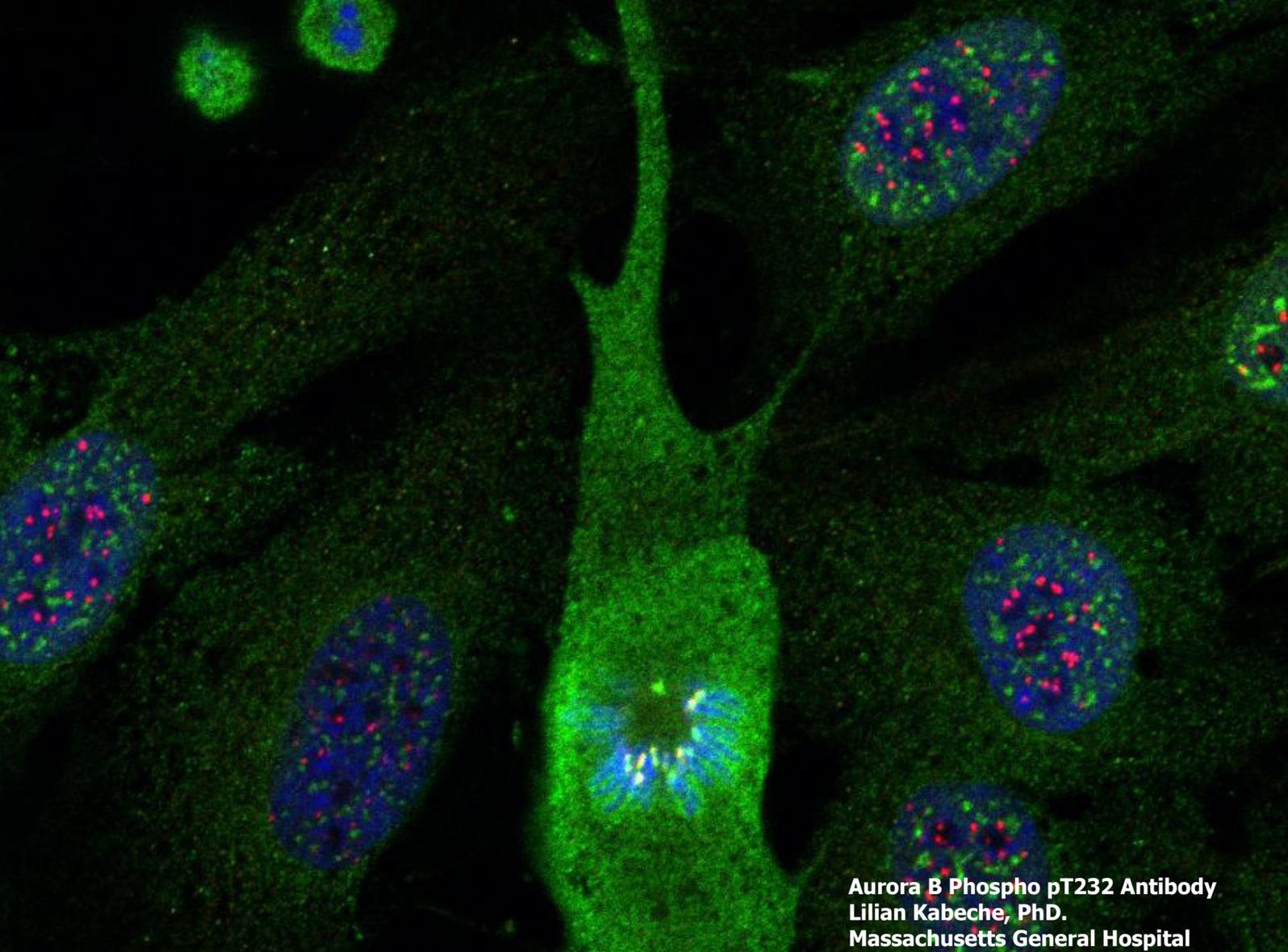
– Magnify Your Nanoculture –



**MBL**<sup>®</sup>  
International Corporation  
[mblintl.com](http://mblintl.com)



EMT Inhibition



Aurora B Phospho pT232 Antibody  
Lilian Kabeche, PhD.  
Massachusetts General Hospital

## Do You Know Where Your Antibodies Are Coming From?

Rockland is a frequently referenced antibody manufacturer and collaborates with leading organizations such as the National Cancer Institute, St. Jude Children's Research Hospital, and Lankenau Institute for Medical Research to develop high-quality antibodies and antibody based tools. **Rockland is a name synonymous with reliability, specificity, and sensitivity.**

- Extensive line of antibodies and reagents
- Non-normalized antibodies for higher concentrations
- Multi-assay validated products
- Custom antibody production and assay development
- Published staff scientists
- USDA accredited animal facilities
- Trial size antibodies available

 **ROCKLAND**<sup>™</sup>  
antibodies & assays  
[www.rockland-inc.com](http://www.rockland-inc.com)

**Protect your experiments with  
Rockland antibodies.**

1-800-656-7625

# Cell

Best of 2016

## SnapShots

Transcription Elongation

*Christine E. Cucinotta and Karen M. Arndt*

Expanding the Ubiquitin Code

*Krishnaraj Rajalingam and Ivan Dikic*

Epigenomic Assays

*Martin Krzywinski and Martin Hirst*

Signaling in Symbiosis

*Sharon R. Long*

The Bacterial Cytoskeleton

*Gero Fink, Andrzej Szewczak-Harris, and Jan Löwe*

Neuronal Regulation of Aging

*Heather J. Weir and William B. Mair*

Microglia in Disease

*Simon Beggs and Michael W. Salter*

Types of Ubiquitin Ligases

*Francesca Ester Morreale and Helen Walden*

## Reviews

**Biology and Applications of CRISPR Systems:  
Harnessing Nature's Toolbox for Genome Engineering**

*Addison V. Wright, James K. Nuñez, and Jennifer A. Doudna*

The Basis of Oncoimmunology

*A. Karolina Palucka and Lisa M. Coussens*

Modeling Development and Disease with Organoids

*Hans Clevers*

The Cellular Phase of Alzheimer's Disease

*Bart De Strooper and Eric Karran*

## Articles and Resources

**Mitochondrial Dynamics Controls T Cell Fate through  
Metabolic Programming**

*Michael D. Buck, David O'Sullivan, Ramon I. Klein Geltink, Jonathan D. Curtis, Chih-Hao Chang, David E. Sanin, Jing Qiu, Oliver Kretz, Daniel Braas, Gerritje J.W. van der Windt, Qiongyu Chen, Stanley Ching-Cheng Huang, Christina M. O'Neill, Brian T. Edelson, Edward J. Pearce, Hiromi Sesaki, Tobias B. Huber, Angelika S. Rambold, and Erika L. Pearce*

**Genomic and Transcriptomic Features of Response to  
Anti-PD-1 Therapy in Metastatic Melanoma**

*Willy Hugo, Jesse M. Zaretsky, Lu Sun, Chunying Song, Blanca Homet Moreno, Siwen Hu-Lieskovan, Beata Berent-Maoz, Jia Pang, Bartosz Chmielowski, Grace Cherry, Elizabeth Seja, Shirley Lomeli, Xiangju Kong, Mark C. Kelley, Jeffrey A. Sosman, Douglas B. Johnson, Antoni Ribas, and Roger S. Lo*

(continued)

**Progressive Loss of Function in a Limb Enhancer during Snake Evolution**

*Evgeny Z. Kvon, Olga K. Kamneva, Uirá S. Melo, Iros Barozzi, Marco Osterwalder, Brandon J. Mannion, Virginie Tissières, Catherine S. Pickle, Ingrid Plajzer-Frick, Elizabeth A. Lee, Momoe Kato, Tyler H. Garvin, Jennifer A. Akiyama, Veena Afzal, Javier Lopez-Rios, Edward M. Rubin, Diane E. Dickel, Len A. Pennacchio, and Axel Visel*

**Domestication and Divergence of *Saccharomyces cerevisiae* Beer Yeasts**

*Brigida Gallone, Jan Steensels, Troels Prael, Leah Soriaga, Veerle Saels, Beatriz Herrera-Malaver, Adriaan Merlevede, Miguel Roncoroni, Karin Voordeckers, Loren Miraglia, Clotilde Teiling, Brian Steffy, Maryann Taylor, Ariel Schwartz, Toby Richardson, Christopher White, Guy Baele, Steven Maere, and Kevin J. Verstrepen*

**Zika Virus Infection during Pregnancy in Mice Causes Placental Damage and Fetal Demise**

*Jonathan J. Miner, Bin Cao, Jennifer Govero, Amber M. Smith, Estefania Fernandez, Omar H. Cabrera, Charise Garber, Michelle Noll, Robyn S. Klein, Kevin K. Noguchi, Indira U. Mysorekar, and Michael S. Diamond*

**Microbial Reconstitution Reverses Maternal Diet-Induced Social and Synaptic Deficits in Offspring**

*Shelly A. Buffington, Gonzalo Viana Di Prisco, Thomas A. Auchtung, Nadim J. Ajami, Joseph F. Petrosino, and Mauro Costa-Mattioli*

**Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery**

*Alan Merk, Alberto Bartesaghi, Soojay Banerjee, Veronica Falconeri, Prashant Rao, Mindy I. Davis, Rajan Pragani, Matthew B. Boxer, Lesley A. Earl, Jacqueline L.S. Milne, and Sriram Subramaniam*

**Programmable RNA Tracking in Live Cells with CRISPR/Cas9**

*David A. Nelles, Mark Y. Fang, Mitchell R. O'Connell, Jia L. Xu, Sebastian J. Markmiller, Jennifer A. Doudna, and Gene W. Yeo*

**High-Throughput, High-Resolution Mapping of Protein Localization in Mammalian Brain by In Vivo Genome Editing**

*Takayasu Mikuni, Jun Nishiyama, Ye Sun, Naomi Kamasawa, and Ryohei Yasuda*

**Ultra-High Resolution 3D Imaging of Whole Cells**

*Fang Huang, George Sirinakis, Edward S. Allgeyer, Lena K. Schroeder, Whitney C. Duim, Emil B. Kromann, Thomy Phan, Felix E. Rivera-Molina, Jordan R. Myers, Irnov Irnov, Mark Lessard, Yongdeng Zhang, Mary Ann Handel, Christine Jacobs-Wagner, C. Patrick Lusk, James E. Rothman, Derek Toomre, Martin J. Booth, and Joerg Bewersdorf*

# THE BENCHMARK IN ANTIBODIES

Since the day it was founded, Proteintech has been making all of its products to the highest standards possible while taking complete responsibility for the quality of each product.

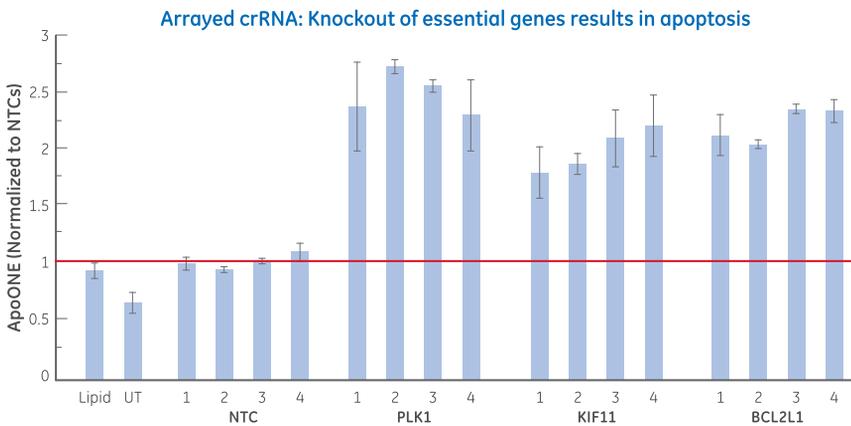
- **Proteintech makes every single antibody in its 12,000-strong catalog.**
- **Each Proteintech product is unique and cannot be bought under a different label.**
- **Antibodies tested in siRNA knockdown experiment to demonstrate specificity.**
- **It works in any species and any application or get a full money-back refund.**

# CRISPR

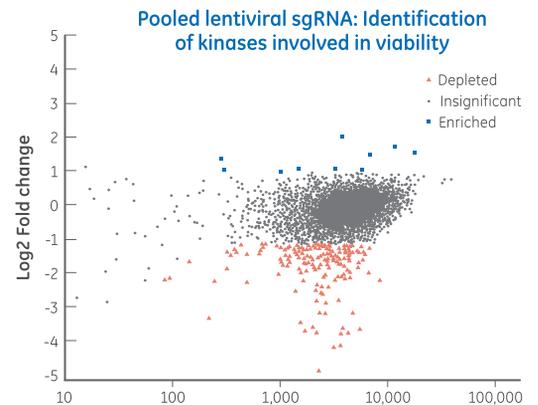
## Fast, efficient, loss-of-function screening with Dharmacon CRISPR Libraries

Rapidly assess multiple target sites per gene across hundreds or thousands of genes with Dharmacon™ Edit-R™ CRISPR-Cas9 screening libraries. Our arrayed synthetic crRNA and pooled lentiviral sgRNA collections leverage our powerful CRISPR RNA design algorithm to deliver high functionality and superior specificity – for high-confidence, loss-of-function screens.

**Edit-R crRNA Libraries | Edit-R Pooled Lentiviral sgRNA Libraries**



U2OS-UbiiEGFP-Cas9 stable cell seeded at 10,000 cells/well in 96-well format were transfected the following day with four different crRNA:tracrRNA complexes at 25 nM concentration targeting PLK1, KIF11 or BCL2L1 or four non-targeting crRNA controls (NTC). Wells transfected with lipid alone (Lipid) or left untreated (UT) were also included as controls. The effects on apoptosis were assayed using Casp3/9 homogeneous assay (ApoONE, Promega) at 48 hours post-transfection. Data normalized to average of NTC (non-targeting) crRNA controls.



Hits are identified by how strongly and significantly the experimental constructs are depleted at T1 compared to the constructs in the reference sample (T0).  $\geq 1.5$ -fold depletion |  $\leq 0.05$  padj value

# Immunotherapy research reagents

Proteins, Antibodies, Genes, ELISA kits

155 molecules  
12 species  
8000 products

Immune checkpoints  
Cytokines  
Tumor antigens

Superior quality  
Lowest price  
Quick shipping

## ▶ Immune checkpoint

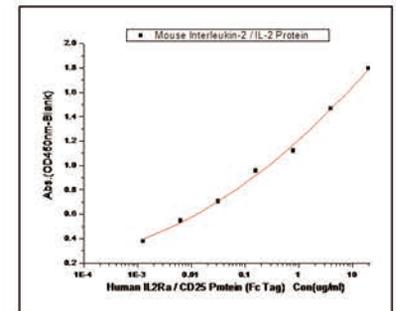
PD1	PD-L1	CTLA4	B7-H3	BTLA
VISTA	CD40	ICOSL	4-1BBL	PD-L2
CEACAM1	CD80	CD86	CD200	CD27

## ▶ Cytokines

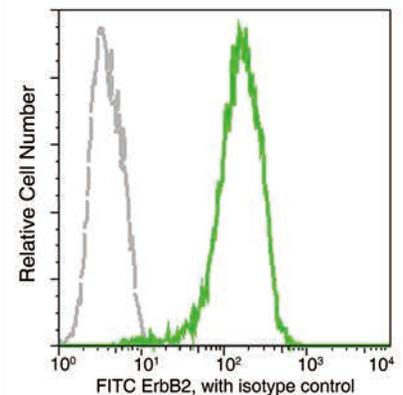
IL-2	IL-4	GM-CSF	TNF- $\alpha$	(IFN)- $\beta$
IL-6	IL-6R	IL-1	CXCL10	XCL1
IL12	IL-13	IL-15	IL-18	IL-21

## ▶ Tumor antigens

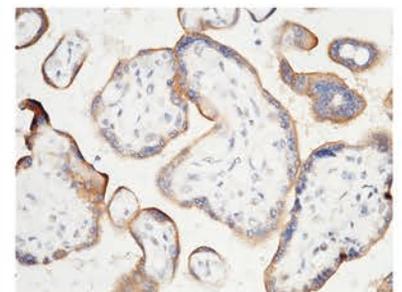
CEA	p53	TYR	gp100	CD20
LAGE-1	PSA	PSMA	EphA3	CDK4
NRAS	muc1	AFP	CD19	gp38



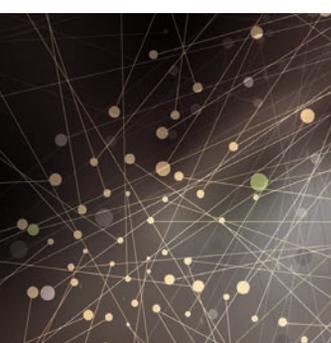
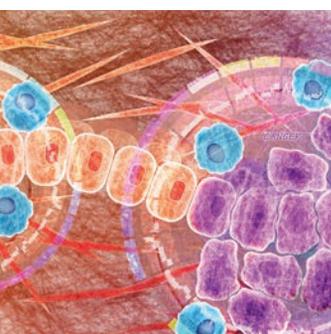
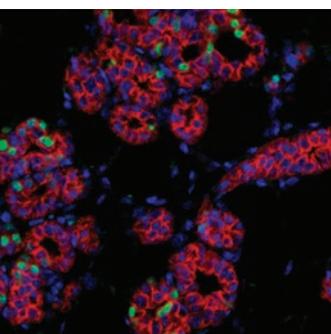
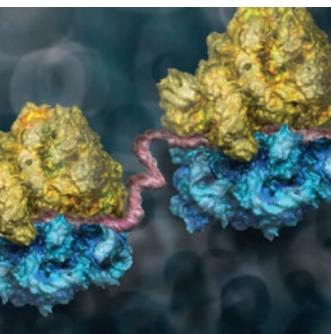
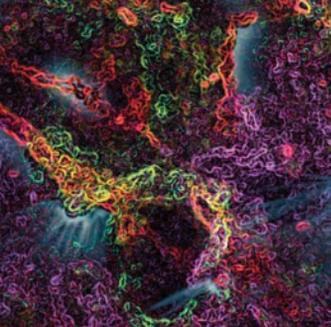
Activity validation for IL2 protein :  
(51061-MNAE) binding to human IL2Ra



FCM-ERBB2 antibody  
(10004-R511)-SKBR3 cells



IHC-PD-L1 antibody  
(10084-R015)-human placenta



# 2017 SCIENTIFIC CONFERENCES

Presenting the most significant research on cancer etiology, prevention, diagnosis, and treatment

**Precision Medicine Series:  
Opportunities and Challenges of  
Exploiting Synthetic Lethality in Cancer**

*Cochairpersons: René Bernards,  
William C. Hahn, and Louis M. Staudt*  
January 4-7, 2017 • San Diego, CA

**AACR International Conference  
on New Frontiers in Cancer Research**

*Cochairpersons: Peter A. Jones and Frank McCormick*  
January 18-22, 2017 • Cape Town, South Africa

**AACR Annual Meeting 2017**

*Program Committee Chairperson: Kornelia Polyak*  
April 1-5, 2017 • Washington, DC

**Head and Neck Cancer Conference: Optimizing  
Survival and Quality of Life Through Basic, Clinical,  
and Translational Research**

*Chairperson: Jeffrey N. Myers*  
*Cochairpersons: Jennifer Rubin Grandis,  
J. Silvio Gutkind, and Quynh-Thu Le*  
April 23-25, 2017 • San Diego, CA

**AACR International Conference  
on Translational Cancer Medicine**

*Cochairpersons: Carlos L. Arteaga  
and Carlos Gil M. Ferreira*  
May 4-6, 2017 • São Paulo, Brazil

**Hematologic Malignancies:  
Translating Discoveries to Novel Therapies**

*Conference Chairperson: Jonathan D. Licht*  
*Conference Cochairpersons: Lucy A. Godley,  
Louis M. Staudt, and Catherine J. Wu*  
May 6-9, 2017 • Boston, MA

**International Conference  
on Malignant Lymphoma (ICML)**

*ICML President: Franco Cavalli*  
*Chairperson, Local Organizing Committee:  
Michele Ghielmini*  
June 14-17, 2017 • Lugano, Switzerland

**EACR-AACR-SIC Special Conference 2017:  
The Challenges of Optimizing Immuno-and Targeted  
Therapies: From Cancer Biology to the Clinic**

*Cochairpersons: Anton J.M. Berns,  
Nancy E. Davidson, and Silvia Giordano*  
June 24-27, 2017 • Florence, Italy

**Third CRI-CIMT-EATI-AACR International  
Cancer Immunotherapy Conference**

September 6-9, 2017 • Mainz/Frankfurt, Germany

**Advances in Modeling Cancer in Mice:  
Technology, Biology, and Beyond**

*Cochairpersons: Cory Abate-Shen, Kevin M. Haigis,  
Katerina A. Politi, and Julien Sage*  
September 24-27, 2017 • Orlando, FL

**Tenth AACR Conference on The Science of  
Cancer Health Disparities in Racial/Ethnic  
Minorities and the Medically Underserved**

*Cochairpersons to be announced*  
September 25-28, 2017 • Atlanta, GA

**AACR-NCI-EORTC Molecular  
Targets and Cancer Therapeutics**

*Cochairpersons to be announced*  
October 26-30, 2017 • Philadelphia, PA

**Prostate Cancer**

*Conference Cochairpersons:  
Johann S. de Bono, Karen E. Knudsen,  
Peter S. Nelson, and Mark A. Rubin*  
December 2-5, 2017 • Orlando, FL

**San Antonio Breast Cancer Symposium  
Presented by CTRC-AACR-BCM**

*Codirectors: Carlos L. Arteaga,  
Virginia G. Kaklamani, and C. Kent Osborne*  
December 5-9, 2017 • San Antonio, TX

Learn more and register at  
[AACR.org/Calendar](http://AACR.org/Calendar)

**AACR** American Association  
for Cancer Research

FINDING CURES TOGETHER™

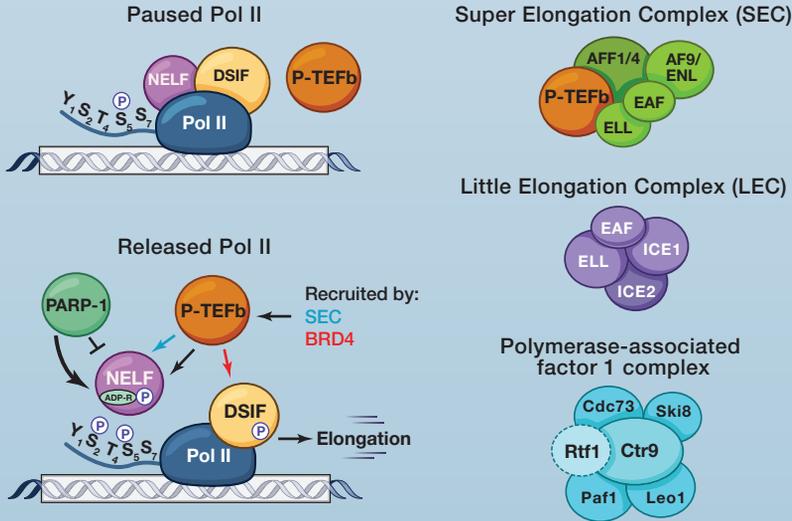
# SnapShot: Transcription Elongation

# Cell

Christine E. Cucinotta and Karen M. Arndt

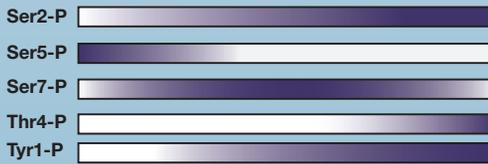
Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA

## Pol II pause, release, elongation rate, and processivity factors



Protein	Function
P-TEF-b/Cdk9/CycT1/ (S.c.:Bur1/Bur2)	Cyclin-dependent kinase, releases Pol II from pausing and phosphorylates Pol II, NELF, and DSIF
DSIF (S.c.:Spt4-Spt5)	Stabilizes Pol II pausing, recruits elongation factors, stimulates elongation
NELF	Negative elongation factor
PARP-1	Poly[adenosine diphosphate (ADP-ribose)] polymerase, transfers ADP-ribose to NELF to inhibit NELF
BRD4	Recruits and activates P-TEFb
TFIIS	Resumes Pol II elongation from backtracked-arrest
TFIIF	Prevents transient pausing of Pol II
Elongin	Increases transcription elongation rate
Super elongation complex	Recruits P-TEFb, increases Pol II elongation rate
Little elongation complex	Required for snRNA expression in metazoans
Paf1C	Associated with Pol II and Spt4-Spt5, recruits chromatin remodelers, histone chaperones, and modifiers
GDOWN1	Tightly associated with Pol II, stabilizes pausing
RNA processing factors	CCR4-NOT, THO/TREX, Xrn2: mRNA processing and export factors that also regulate elongation

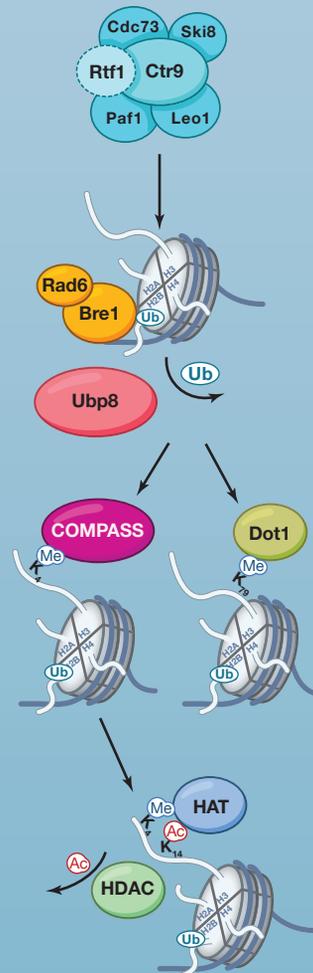
## Pol II CTD phosphorylation



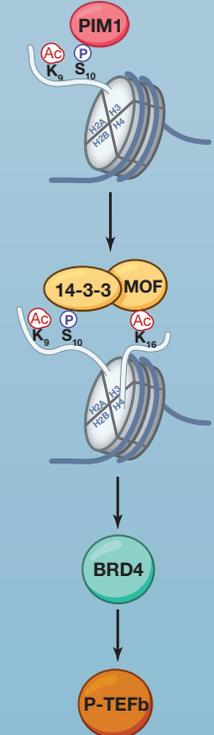
Kinase	Phosphatase
Cdk9 (S.c.: Bur1) Cdk12 (S.c.: Ctk1)	Fcp1
Cdk7 (S.c.: Kin28) Cdk9 (S.c.: Kin1)	Rtr1, Ssu72
Cdk7 (S.c.: Kin28) Cdk9 (S.c.: Kin1)	Ssu72
Plk3, Cdk9	?
?	Rtr1 (S.c.)

## Histone modifications involved in elongation

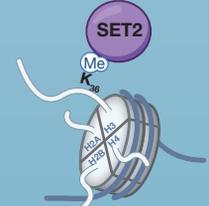
### H2B K123 monoubiquitylation cascade



### H3 S10 phosphorylation cascade

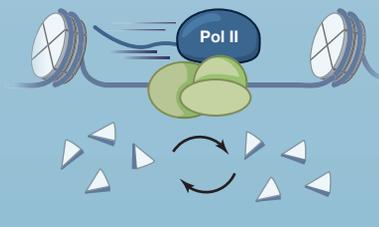


### H3 K36 methylation

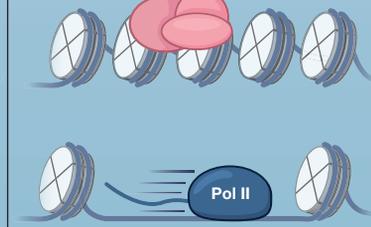


## Histone chaperones and chromatin remodelers

### Nucleosome disassembly and reassembly



### Nucleosome spacing



Protein	Function	Role in elongation
FACT	Histone chaperone complex of H.s.: hSPT16, Ssrp1, S.c.: Spt16, Nhp6, Pob3	Displaces H2A-H2B dimer in the wake of transcribing Pol II, evicts nucleosomes; reassembles nucleosomes, regulates H2BK123ub, prevents cryptic transcription
Spt6	Histone chaperone	Prevents cryptic transcription, required for proper histone occupancy during elongation
Asf1	Histone chaperone	Controls H3 exchange during transcription
Nap1	Histone chaperone	Binds H2A-H2B and forms hexasome structures through RSC, promotes nucleosome assembly
Rtt106	Histone chaperone	Binds H3-H4 and promotes transcription-coupled H3 deposition, prevents cryptic transcription
Chd1	Chromatin remodeler	Controls nucleosome spacing and histone exchange, promotes Pol II promoter escape (mammals)
ISW1	Chromatin remodeler	Controls nucleosome spacing and histone exchange
RSC	Chromatin remodeler	Helps Pol II passage through nucleosomes and maintains proper nucleosome occupancy

QUALIFIED ANTIBODIES | MADE IN THE USA | VALIDATED IN THE USA

OH MY  
#¥\$&!

**BETHYL**  
LABORATORIES, INC  
Really Good Antibodies

### Oh My Gosh!—Great Reactions Start with Really Good Antibodies

That's because every antibody is manufactured and validated on-site using a unique process that verifies target specificity with paired antibodies raised against distinct protein epitopes. We've heard our customers' great reactions for over 40 years and now we'd like to hear what you'll say.

Request a free trial size antibody by visiting [bethyl.com/1613](http://bethyl.com/1613)

Terms & Conditions Apply. Please see webpage for complete details.

©2016 Bethyl Laboratories, Inc. All rights reserved.



# MP Premium Dextran Sulfate Sodium Salt (DSS)

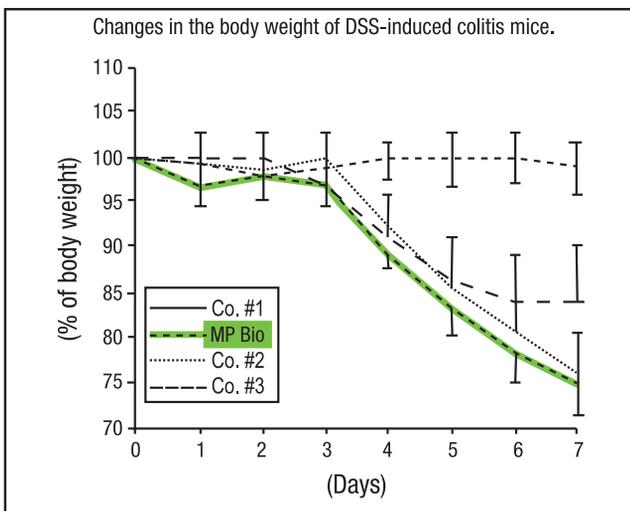
## The Gold Standard in Colitis Research



At least 3,000 Scientists Prove  
MP DSS is the Most Effective  
in the Industry!

Let us **PROVE** it to you!

### High Purity, Reproducible Quality! Best Results!



- Validated by 3,000+ scientific publications
- Highest purity (99%) and stability
- Offers Highest Sulfur content (18-20%)  
< 0.2% free sulfate
- Highest efficiency
- Best value in the industry.

*The mice were fed normal chow and 2.0% wt/wt DSS in water for 1 week and then sacrificed. The weight of individual mice was monitored daily. The data represent means=SD (data obtained from Bamba et.al., 2012)*

### FREE SAMPLE (\$180 value) Just Ask!

Get a 10g sample of our MP Gold Standard DSS for colitis research absolutely FREE.  
Use promo code: DSSCP when placing your order.

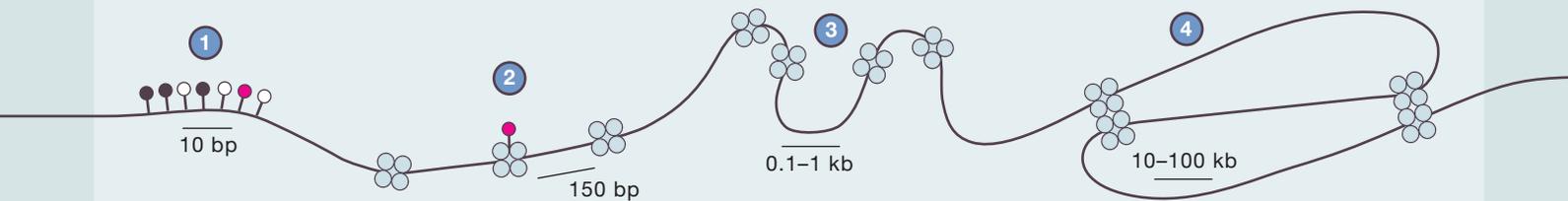
[www.mpbio.com/DSS](http://www.mpbio.com/DSS) or call 800.854.0530



# SnapShot: Epigenomic Assays

Martin Krzywinski<sup>1</sup> and Martin Hirst<sup>1,2</sup>

<sup>1</sup>Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency Research Centre, BC Cancer Agency, Vancouver, BC V5Z 1L3, Canada; <sup>2</sup>Department of Microbiology and Immunology, Michael Smith Laboratories Centre for High-Throughput Biology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

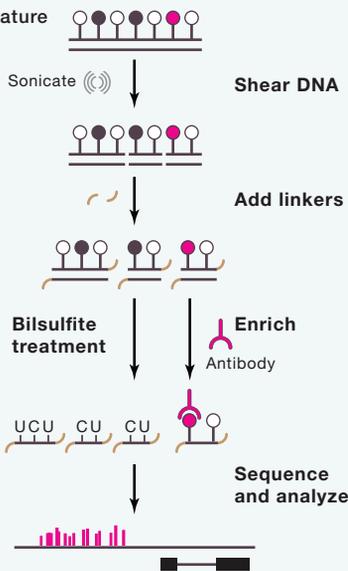


## 1 DNA methylation analysis

Cytosine modification is a feature of the mammalian genome.

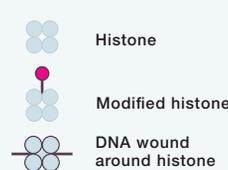


5-methyl-cytosine (5mC) and its oxidative derivatives (e.g., 5hmC) are measured genome-wide using enrichment- and conversion-based methodologies followed by massively parallel sequencing. Bisulfite conversion provides quantitative measurements of 5mC but is unable to distinguish 5mC. Antibody enrichment provides qualitative measurement of 5mC and 5hmC. Bisulfite-converted or -enriched DNA is purified, subjected to library construction and clonally sequenced. Specialized algorithms are required to align bisulfite-converted reads to a reference genome.

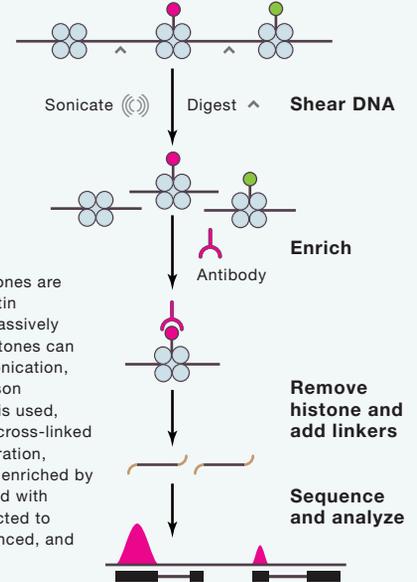


## 2 Chromatin immunoprecipitation sequencing

Histones are decorated with chemical modifications.



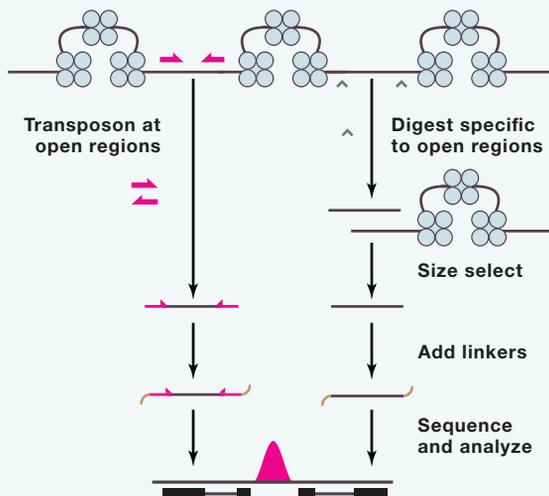
Genomic positions of modified histones are measured genome-wide by chromatin immunoprecipitation followed by massively parallel sequencing (ChIP-seq). Histones can be liberated from the genome by sonication, enzymatic digestion, or by transposon insertion (not shown). If sonication is used, chromatin must first be chemically cross-linked (see Step 4). Following histone liberation, specific chemical modifications are enriched by immuno-absorption. DNA associated with enriched histones is purified, subjected to library construction, clonally sequenced, and aligned to a reference genome.



## 3 Open chromatin analysis

Nucleosome-depleted regions are enriched in regulatory elements.

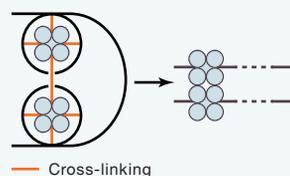
Genomic positions of open chromatin are measured genome-wide by massively parallel sequencing of a collection of DNA fragments liberated from intact chromatin by either transposon insertion, enzymatic digestion, or sonication. The resulting DNA fragments are subjected to size selection or phenol-chloroform extraction to deplete nucleosome-associated DNA. The resulting DNA is purified, subjected to library construction, clonally sequenced, and aligned to a reference genome.



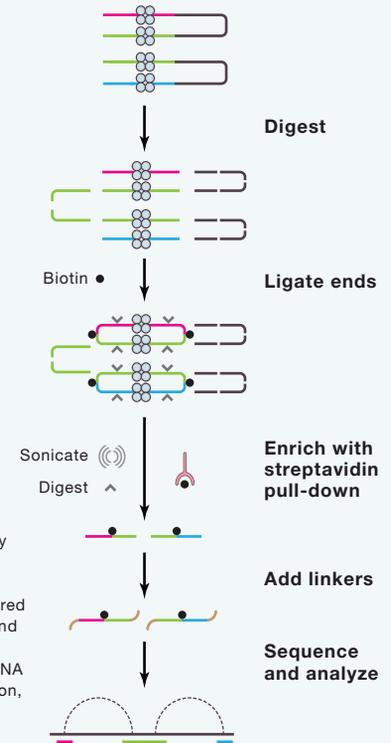
## 4 3D chromatin capture

Chromatin loop contacts reveal distal regulatory elements and structural domains.

Genomic positions of long-range chromatin contacts are measured genome-wide by massively parallel sequencing a collection of DNA fragments generated by proximity ligation. Intact chromatin is cross-linked to physically link genomically distal nucleosomes that are adjacent in 3-dimensional space.



Cross-linked chromatin is enzymatically digested and the resulting DNA end labeled with biotin and subjected to proximity ligation. Ligated DNA is sheared by sonication or enzymatic digestion and the ligated junctions enriched by streptavidin pull-down. The resulting DNA is purified, subjected library construction, clonally sequenced, and aligned to a reference genome.

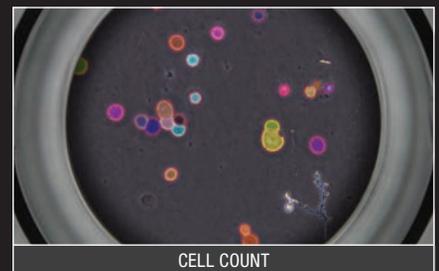
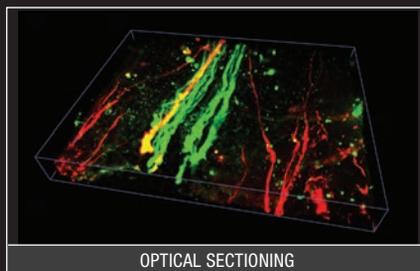
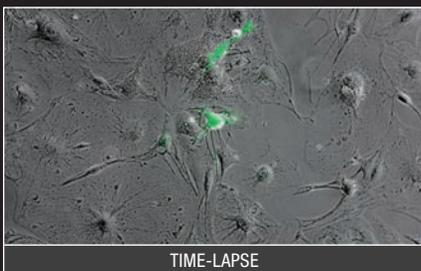
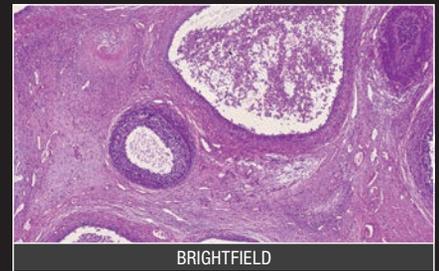
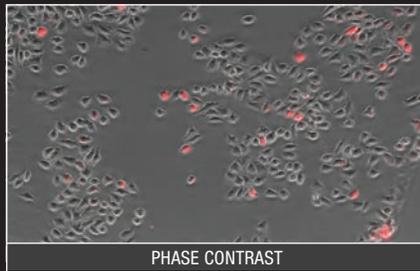




# ALL-IN-ONE FLUORESCENCE MICROSCOPE



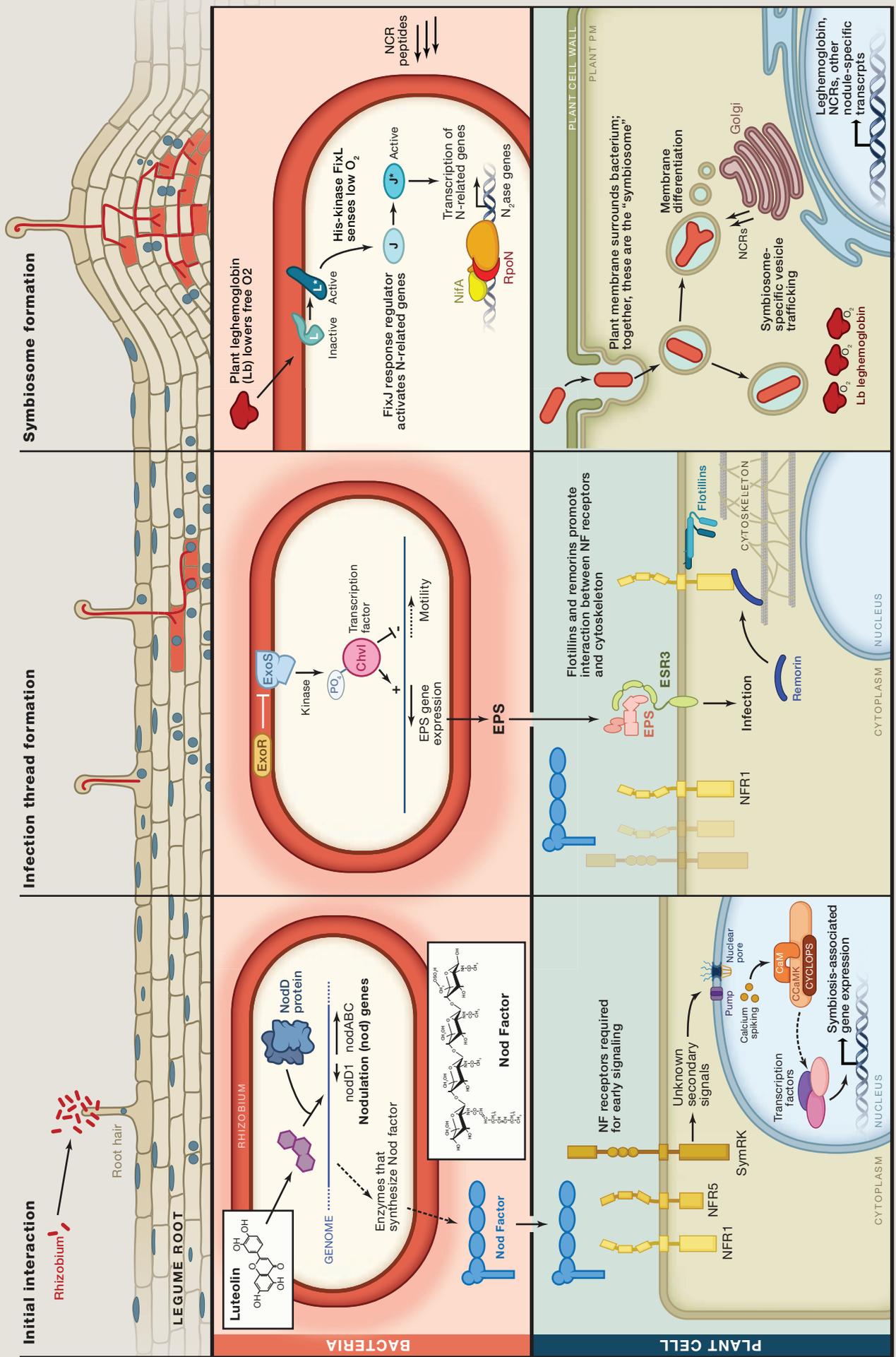
KEYENCE optimizes the imaging process with a fully automated, easy-to-use system that captures high-resolution images and allows researchers to achieve their desired results faster.



For more information, visit: [www.keyence.com/bz-x700](http://www.keyence.com/bz-x700)

# SnapShot: Signaling in Symbiosis

Sharon R. Long  
Department of Biology, Stanford University, Stanford, CA 94305, USA





Make scientific  
breakthroughs faster with  
**SciFinder<sup>®</sup>**, the most  
complete source of chemistry  
and related information.

CAS scientists thoroughly analyze the literature indexed in SciFinder so you can quickly search the authoritative CAS databases for substances, reactions, and patent and journal references, giving you confidence you aren't missing important results.

- Save time with powerful, intuitive search features that allow you to quickly find the most relevant answers to your research questions
- Build your plan, design a synthesis and easily manage results with insightful organization and planning tools
- Stay current in your field by easily creating alerts for timely updates on the latest research
- Affordable pricing programs are available for organizations of all types and sizes

SciFinder is designed by and for scientists like you to enhance your discovery efforts every step of the way.

Learn more about SciFinder at [www.cas.org/SciFinder](http://www.cas.org/SciFinder).



Your neurons are  
talking.  
Are you listening?



Maestro MEA platform with AxIS™ software showing real-time neural activity across a 96-well plate.

## Maestro™ microelectrode array (MEA) *in vitro* neural activity analysis

Record network-level neural activity directly on the benchtop with Maestro. Over minutes or months, gain unprecedented access to electrical network function from cultured neural populations.

Straightforward and easy to use, any researcher can measure activity from neurons in 12-, 48-, or 96-well plates.

**Connect to your network.**

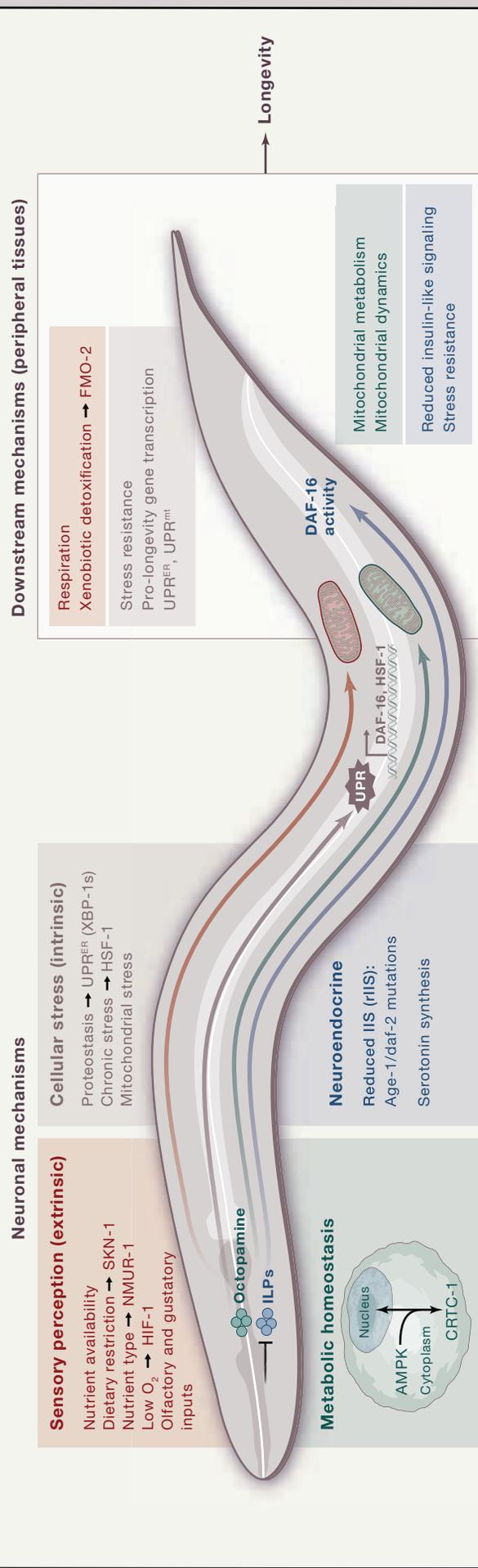


# SnapShot: Neuronal Regulation of Aging

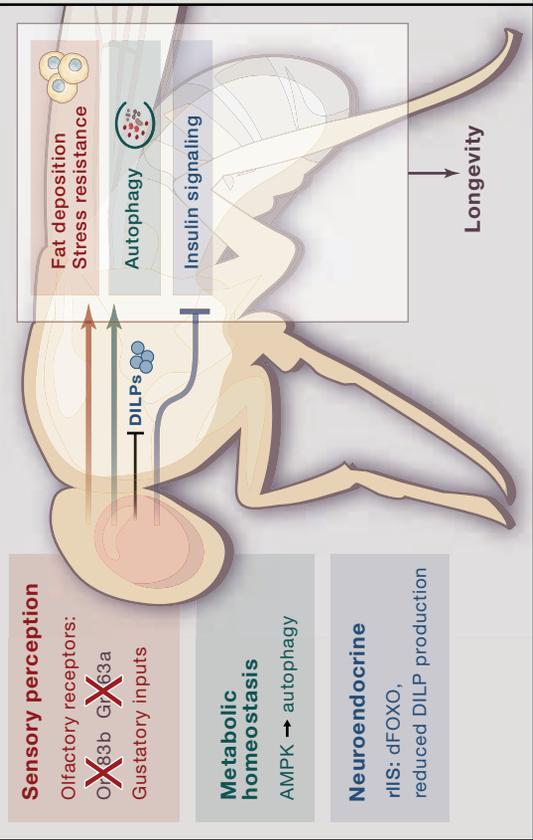
Heather J. Weir and William B. Mair  
Department of Genetics and Complex Diseases, Harvard T.H. Chan School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA

# Cell

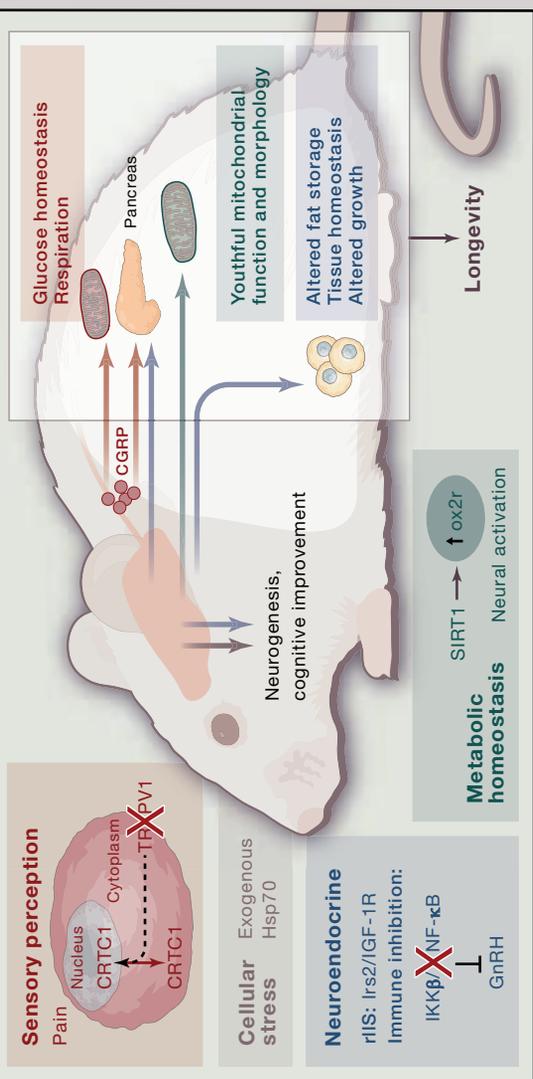
## C. elegans



## Drosophila



## Mouse



# Knock Out Any Gene!



## CRISPR/CAS 9 Genome Editing Kits

OriGene's system for genome disruption and gene replacement delivers pre-designed plasmids and all the vectors needed to knock out any human or mouse gene. Knockout is as simple as 1-2-3:

- 1 Search the gene symbol on [origene.com](http://origene.com) and order
- 2 Follow the simple protocol for transfection and Puro selection
- 3 Validate the knockout

### Kit Components Include:

- 2 guide RNA vectors to ensure efficient cleavage
- Donor vector with predesigned homologous arms
- Knockin GFP-Puro for selection
- Scramble gRNA as negative control included



Scan to view  
CRISPR/Cas-9 Video

Come to OriGene,  
the trusted molecular biology  
expert, for your CRISPR needs.



[www.origene.com/CRISPR-CAS9](http://www.origene.com/CRISPR-CAS9)

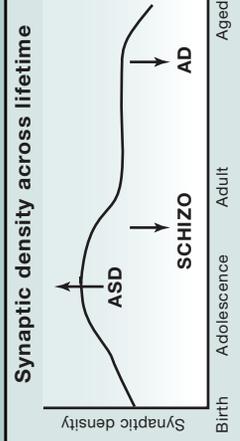
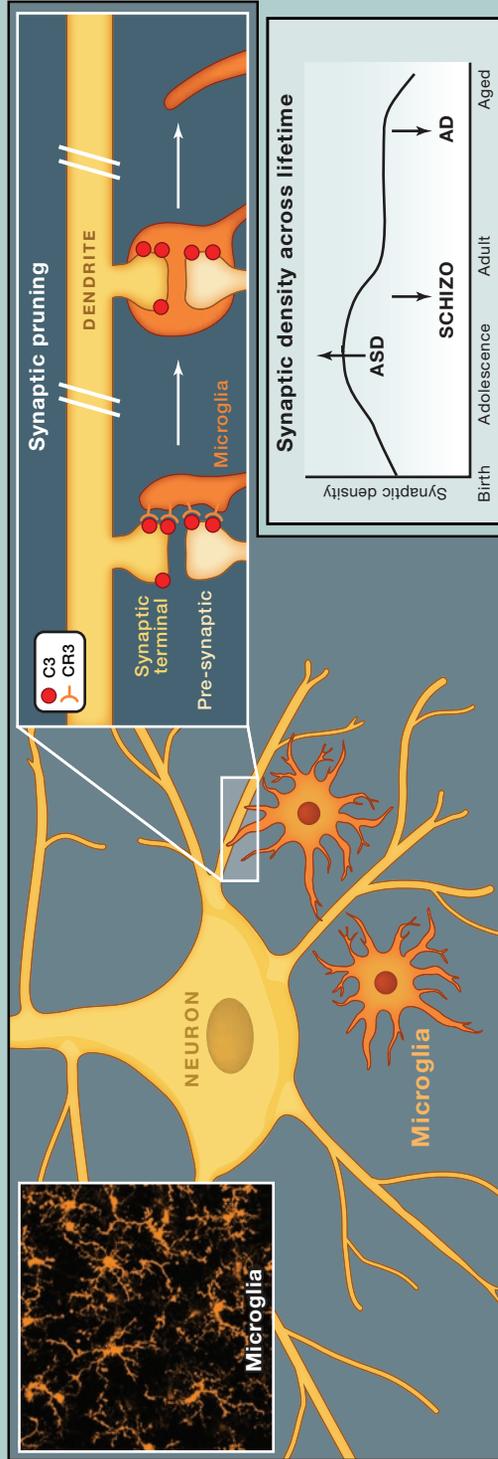
# SnapShot: Microglia in Disease

Simon Beggs<sup>1,2</sup> and Michael W. Salter<sup>1</sup>

<sup>1</sup>Program in Neurosciences & Mental Health, Hospital for Sick Children, Toronto ON M5G 1X8, Canada

<sup>2</sup>Developmental Neurosciences Programme, UCL Institute of Child Health, London WC1N 1EH, UK

# Cell



### AUTISM SPECTRUM DISORDERS

Decreased C3 expression  
Retention of immature synapses  
Abnormal connectivity

### SCHIZOPHRENIA

Healthy High risk Schizophrenia

Increased microglial activity and pruning

### ALZHEIMER'S DISEASE

Amyloid plaque

Increased C3 with age  
Microglia drawn to plaque deposits

### PAIN

CSF1 P2X4 BDNF

Pain signaling to brain  
Spinal inhibition

P2X4R+ microglia in spinal dorsal horn

# Reliable Quantification

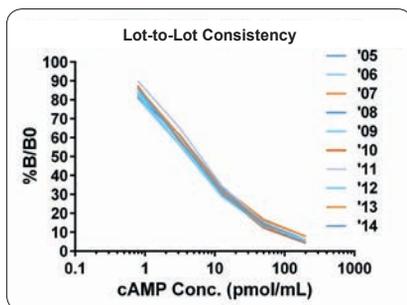
## High-Sensitivity Screening Assays

### Rapidly and Reliably Quantify Second Messengers with High-Sensitivity Assays

As our understanding of neurodegenerative disease origins increases, so does the need for innovative, high-quality screening assays for identifying potential therapeutics. Enzo's expertise in assay development and labeling and detection provide a unique collection of innovative research tools to monitor neural signaling networks.

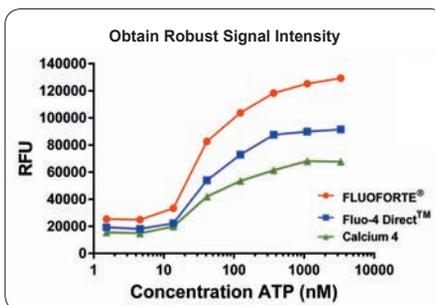
#### cAMP/cGMP ELISA Kits

Most sensitive and reliable ELISA kits for detection of cAMP/cGMP



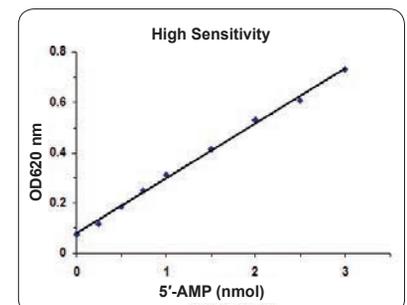
#### FLUOFORTE<sup>®</sup> Calcium Assays

Brighter, more robust fluorescent calcium mobilization assays



#### Cyclic Nucleotide Phosphodiesterase Assay

Unique colorimetric assay used to screen inhibitors and modulators of cyclic nucleotide PDE activity

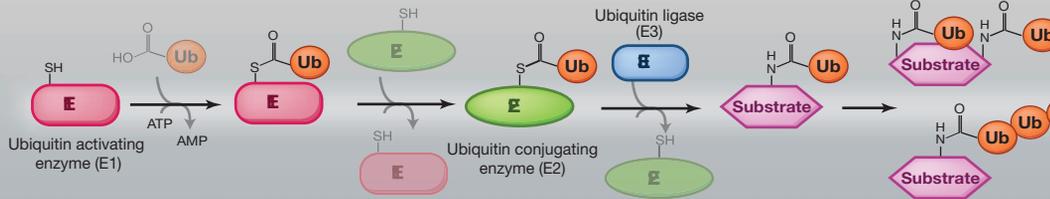


# SnapShot: Types of Ubiquitin Ligases

Francesca Ester Morreale and Helen Walden

MRC Protein Phosphorylation and Ubiquitylation Unit at University of Dundee, Scotland DD1 1H, UK

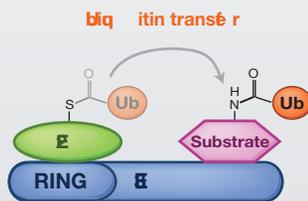
## UBIQUITINATION



Protein degradation by the proteasome, cell-cycle progression, transcriptional regulation, DNA repair, signal transduction

## RING E3s

~600 in humans



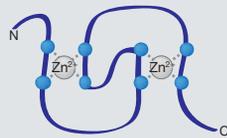
### E2 ubiquitination mechanisms

Neddylation (e.g., CRLs)  
Phosphorylation (e.g., c-CBL, APC, Mdm2)  
Small molecule: E3 interaction (e.g., TRAF-2: sphingosine-1-phosphate)

### Characteristic domains

**RING** (Really Interesting New Gene):  
Zn binding domain, recruits the Ub-charged E2

**U-box**: same fold as RING without Zn coordination



### Types of RING E3

#### Monomeric RING

Example: c-CBL



#### Homodimeric RING

Examples: cIAP, RNF4, BIRC7, IDOL, TRIM5α



#### Heterodimeric RING

Examples: BRCA1-BARD1, Mdm2-MdmX, RING1B-Bmi1



#### Monomeric U-box

Example: E4B

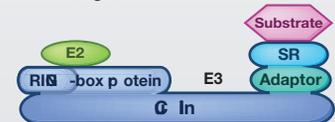


#### Homodimeric U-box

Examples: CHIP, Prp19



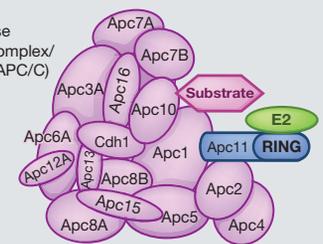
#### Cullin-RING ligases



Composed of:  
**Substrate receptor (SR)** - example: F-box protein Skp2  
**Adaptor protein** - example: Skp1  
**Cullin** - example: Cul1  
**Ring-box protein (Rbx)** - example: Rbx1

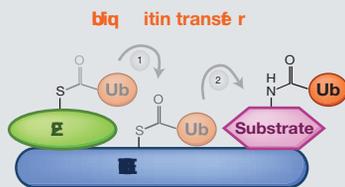
#### Other multisubunit E3s

The anaphase promoting complex/cyclosome (APC/C)



## HECT E3s

~30 in humans

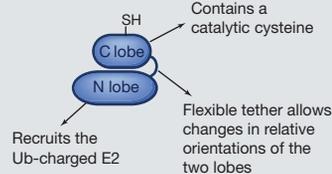


### E2 ubiquitination mechanisms

Autoinhibition through intramolecular interactions released by:  
phosphorylation (e.g., Itch), neddylation (e.g., Smurf1), protein interaction (e.g., Smurf1, Smurf2)

### Characteristic domain

**HECT** (homologous to E6-AP carboxyl terminus)



### ECT Family



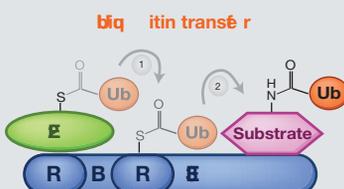
Examples: Smurf1 (2 WW motifs), Smurf2 (3 WW motifs), Itch (4 WW motifs)



Example: E6AP

## RBR E3s

~12 in humans

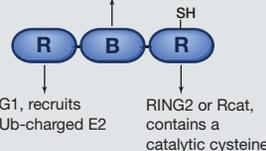


### E2 ubiquitination mechanisms

Autoinhibition through intramolecular interactions released by:  
phosphorylation/phosphoubiquitin interaction (Parkin), binding to neddylated CRLs (e.g., HHARI), protein interactions (e.g., HOIP)

### Characteristic domain

IBR (InBetweenRING) or BRcat



RING1, recruits the Ub-charged E2

RING2 or Rcat, contains a catalytic cysteine

### Examples of RBR E3s





# Quantify with Confidence

*Ever wonder about the quantitative nature of image data?  
We do. Introducing the new **ORCA-Flash4.0 V3** and the  
**W-VIEW GEMINI-2C** dual channel, dual camera optics.*

**ENGINEERED FOR SUPER RESOLUTION—READY FOR ANYTHING.**

Society for Neuroscience [Booth 1613](#)  
American Society for Cell Biology [Booth 309](#)

**HAMAMATSU**  
PHOTON IS OUR BUSINESS

908-231-0960 [hamamatsucameras.com](http://hamamatsucameras.com)

# Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering

Addison V. Wright,<sup>1</sup> James K. Nuñez,<sup>1</sup> and Jennifer A. Doudna<sup>1,2,3,4,5,6,\*</sup>

<sup>1</sup>Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA

<sup>2</sup>Howard Hughes Medical Institute HHMI, University of California, Berkeley, Berkeley, CA 94720, USA

<sup>3</sup>Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720, USA

<sup>4</sup>Center for RNA Systems Biology, University of California, Berkeley, Berkeley, CA 94720, USA

<sup>5</sup>Innovative Genomics Initiative, University of California, Berkeley, Berkeley, CA 94720, USA

<sup>6</sup>Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, Berkeley, CA 94720, USA

\*Correspondence: [doudna@berkeley.edu](mailto:doudna@berkeley.edu)

<http://dx.doi.org/10.1016/j.cell.2015.12.035>

**Bacteria and archaea possess a range of defense mechanisms to combat plasmids and viral infections. Unique among these are the CRISPR-Cas (clustered regularly interspaced short palindromic repeats-CRISPR associated) systems, which provide adaptive immunity against foreign nucleic acids. CRISPR systems function by acquiring genetic records of invaders to facilitate robust interference upon reinfection. In this Review, we discuss recent advances in understanding the diverse mechanisms by which Cas proteins respond to foreign nucleic acids and how these systems have been harnessed for precision genome manipulation in a wide array of organisms.**

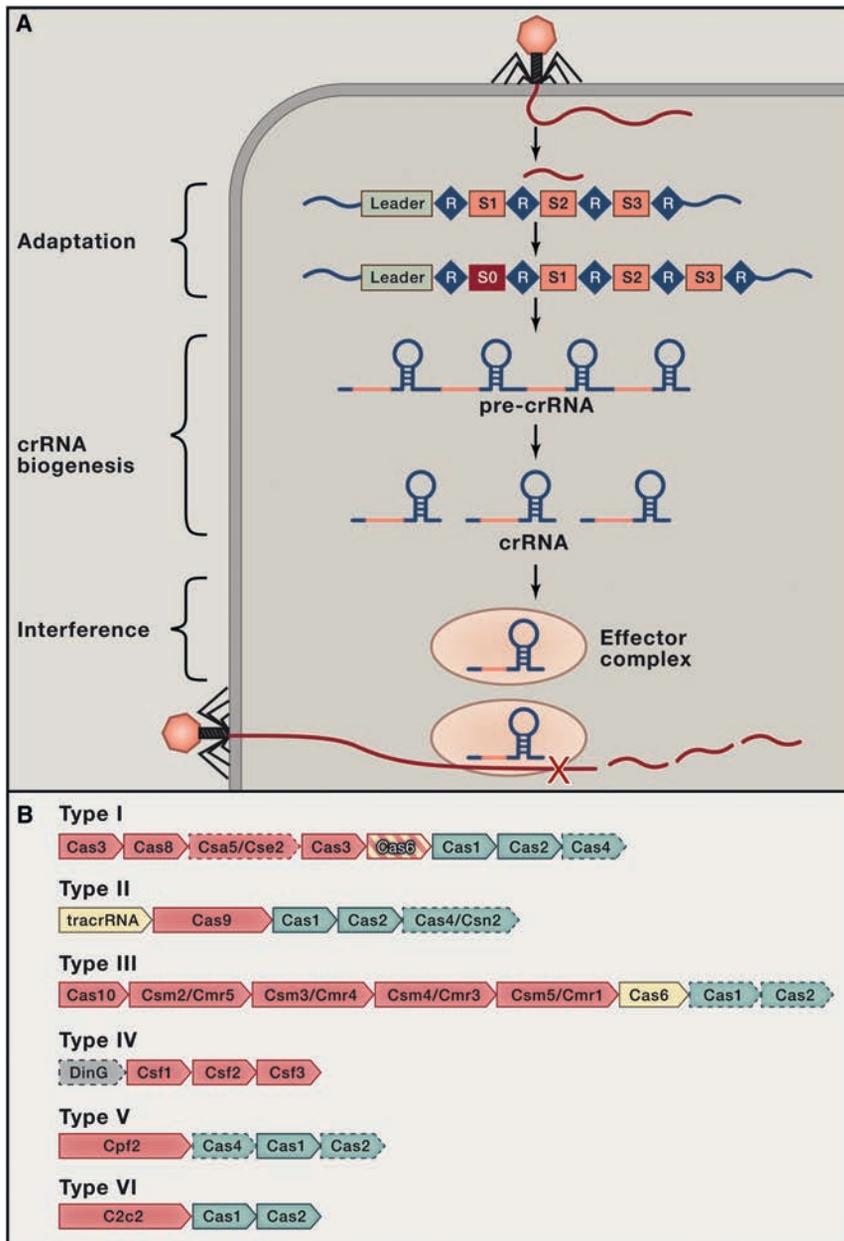
CRISPR-Cas (clustered regularly interspaced short palindromic repeats-CRISPR associated) adaptive immune systems are found in roughly 50% of bacteria and 90% of archaea (Makarova et al., 2015). These systems function alongside restriction-modification systems, abortive infections, and adsorption blocks to defend prokaryotic populations against phage infection (Labrie et al., 2010). Unlike other mechanisms of cellular defense, which provide generalized protection against any invaders not possessing countermeasures, CRISPR immunity functions analogously to vertebrate adaptive immunity by generating records of previous infections to elicit a rapid and robust response upon reinfection.

CRISPR-Cas systems are generally defined by a genomic locus called the CRISPR array, a series of ~20–50 base-pair (bp) direct repeats separated by unique “spacers” of similar length and preceded by an AT-rich “leader” sequence (Jansen et al., 2002; Kunin et al., 2007). Nearly two decades after CRISPR loci were first identified in *Escherichia coli*, spacers were found to derive from viral genomes and conjugative plasmids, serving as records of previous infection (Bolotin et al., 2005; Ishino et al., 1987; Mojica et al., 2005; Pourcel et al., 2005). Sequences in foreign DNA matching spacers are referred to as “proto-spacers.” In 2007, it was shown that a spacer matching a phage genome immunizes the host microbe against the corresponding phage and that infection by a novel phage leads to the expansion of the CRISPR array by addition of new spacers originating from the phage genome (Barrangou et al., 2007).

CRISPR immunity is divided into three stages: spacer acquisition, CRISPR RNA (crRNA) biogenesis, and interference (Figure 1A) (Makarova et al., 2011b; van der Oost et al., 2009). During spacer acquisition, also known as adaptation, foreign DNA is identified, processed, and integrated into the CRISPR locus as a new spacer. The crRNA biogenesis or expression

stage involves CRISPR locus transcription, often as a single pre-crRNA, and its subsequent processing into mature crRNAs that each contain a single spacer. In the interference stage, an effector complex uses the crRNA to identify and destroy any phage or plasmid bearing sequence complementarity to the spacer sequence of the crRNA.

These steps are carried out primarily by Cas proteins, which are encoded by *cas* genes flanking the CRISPR arrays. The specific complement of *cas* genes varies widely. CRISPR-Cas systems can be classified based on the presence of “signature genes” into six types, which are additionally grouped into two classes (Figure 1B) (Makarova et al., 2011b; Makarova et al., 2015; Shmakov et al., 2015). Types I–III are the best studied, while Types IV–VI have only recently been identified (Makarova and Koonin, 2015; Makarova et al., 2015; Shmakov et al., 2015). The signature protein of Type I systems is Cas3, a protein with nuclease and helicase domains that functions in foreign DNA degradation to cleave DNA that is recognized by the multi-protein-crRNA complex Cascade (CRISPR-associated complex for antiviral defense). In Type II systems, the signature *cas9* gene encodes the sole protein necessary for interference. Type III systems are signified by Cas10, which assembles into a Cascade-like interference complex for target search and destruction. Type IV systems have Csf1, an uncharacterized protein proposed to form part of a Cascade-like complex, though these systems are often found as isolated *cas* genes without an associated CRISPR array (Makarova and Koonin, 2015). Type V systems also contain a Cas9-like single nuclease, either Cpf1, C2c1, or C2c3, depending on the subtype (Shmakov et al., 2015; Zetsche et al., 2015a). Type VI systems have C2c2, a large protein with two predicted HEPN (higher eukaryotes and prokaryotes nucleotide-binding) RNase domains (Shmakov et al., 2015). Type I, III, and IV systems are considered



**Figure 1. Function and Organization of CRISPR Systems**

(A) CRISPR immunity occurs in three stages. Upon introduction of foreign DNA, the adaptation machinery selects protospacers and inserts them into the leader end of the CRISPR locus. During crRNA biogenesis, the CRISPR locus is transcribed and sequence elements in the repeats direct processing of the pre-crRNA into crRNAs each with a single spacer. The crRNA then assembles with Cas proteins to form the effector complex, which acts in the interference stage to recognize foreign nucleic acid upon subsequent infection and degrade it.

(B) CRISPR systems are extremely diverse but can largely be classified into six major types. Representative operons for each type are shown here. Genes only present in some subtypes are shown with dashed outlines. Genes involved in interference are colored red, those involved in crRNA biogenesis are colored yellow, and those involved in adaptation are colored blue. Type IV systems are notable for their frequent occurrence in the absence of CRISPR loci.

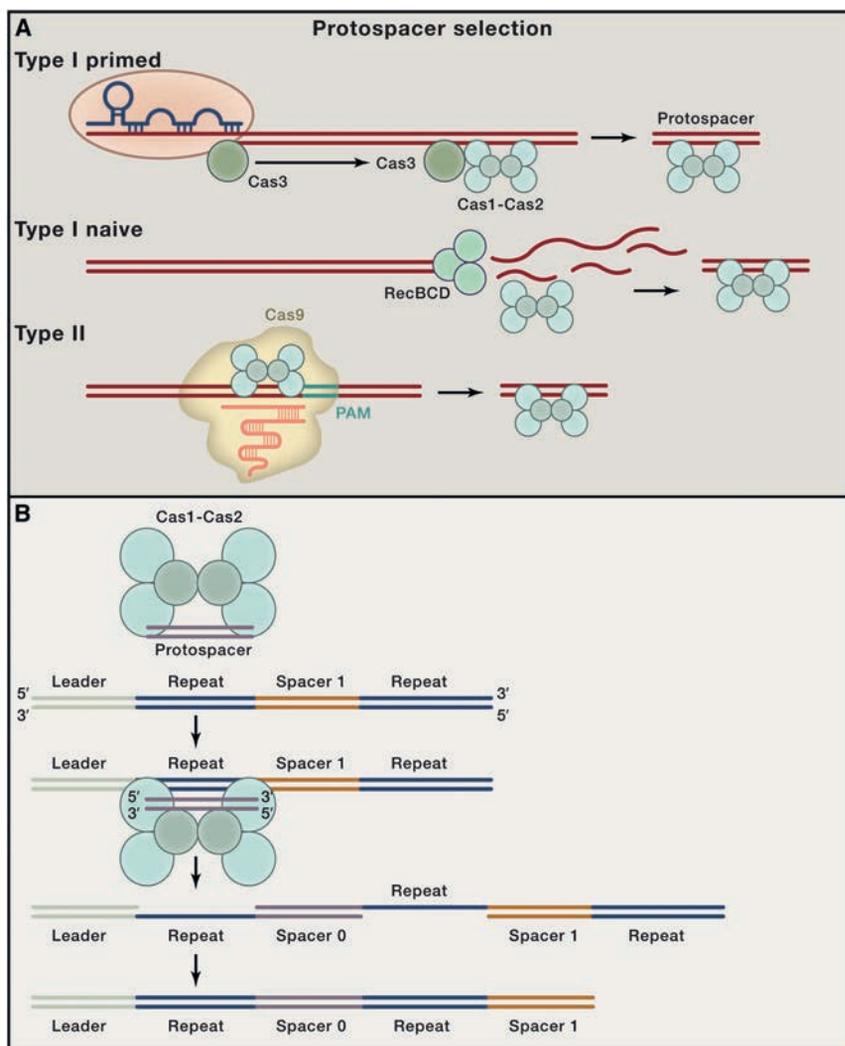
**Acquisition: Creating Genetic Records of Past Infections**

CRISPR immunity begins with the detection and integration of foreign DNA into the host cell's chromosome. In the *Streptococcus thermophilus* Type II-A system, where acquisition was first detected experimentally, new spacers from bacteriophage DNA are inserted into the leader end of the CRISPR locus, causing duplication of the first repeat to maintain the repeat-spacer architecture (Figure 1A) (Barrangou et al., 2007). Subsequent studies using the *E. coli* Type I-E system verified that Cas1 and Cas2 mediate spacer acquisition (Datsenko et al., 2012; Swarts et al., 2012; Yosef et al., 2012). The selection of new protospacer sequences is nonrandom and, in most systems, depends on the presence of a 2–5 nucleotide protospacer adjacent motif (PAM) found next to the protospacer

sequence (Deveau et al., 2008; Mojica et al., 2009). PAM-specific selection of protospacers is critical for immunity, as crRNA-guided interference in most systems depends on the PAM sequence for foreign DNA detection and destruction, which avoids self-targeting at the PAM-free CRISPR locus. Interestingly, spacers originating from the host genome are present in almost 20% of CRISPR-containing organisms, suggesting alternative roles of the CRISPR-Cas machinery in directing other processes such as endogenous gene regulation and genome evolution (Westra et al., 2014). Spacer acquisition has been observed experimentally in various systems across Types I–III. Here, we focus on recent mechanistic studies of acquisition in Type I-E and Type II-A systems, in which the most comprehensive studies have been done.

Class 1 systems based on their multi-subunit effector complexes, while the single-subunit effector Type II, V, and VI systems are grouped into Class 2 (Makarova et al., 2015; Shmakov et al., 2015).

The study of CRISPR biology has revealed enzyme mechanisms that can be harnessed for precision genome engineering and other applications, leading to an explosion of interest in both native CRISPR pathways and the use of these systems for applications in animals, plants, microbes, and humans. In this Review, we discuss recent advancements in the field that reveal unexpected divergence, as well as unifying themes underlying the three stages of CRISPR immunity. In each case, we highlight the ways in which these systems are being harnessed for applications across many areas of biology.



**Figure 2. Protospacer Selection and Integration in Adaptation**

(A) The selection of protospacers for acquisition is poorly understood, but studies suggest at least three distinct mechanisms for the selection of substrates for integration. In Type I systems, primed adaptation occurs when Cascade binds a partially mismatched target. The nuclease/helicase Cas3 is recruited to the target site and then likely translocates along the target DNA to a new site. The new location is then selected as a protospacer to be used by Cas1-Cas2 in the integration reaction. In *E. coli*, naive adaptation involves the nuclease/helicase RecBCD. The degradation products appear to serve as substrates for Cas1-Cas2, but how the variable-length single-stranded products of RecBCD activity are converted into double-stranded protospacers of appropriate size is unknown. In Type II systems, Cas9 recognizes PAM sites and likely recruits Cas1-Cas2 to acquire the flanking sequence.

(B) Cas1-Cas2 act as an integrase to insert protospacers into the CRISPR locus as new spacers. The complex with protospacer bound recognizes the leader-adjacent repeat and catalyzes a pair of transesterification reactions. The 3' OH of each protospacer strand makes a nucleophilic attack on the repeat backbone, one at the leader-side and one at the spacer side. The resulting gapped product is then repaired, causing duplication of the first repeat.

mentary strand, suggesting that Cas1 recognizes PAM sites on potential protospacers before they are processed for integration (Wang et al., 2015).

After a spacer is acquired from a new invader, the resulting crRNA assembles with Cas proteins to form Cascade, the interference complex capable of targeting PAM-adjacent DNA sequences matching the spacer sequence of the crRNA (Brouns et al., 2008; Jore et al., 2011; Lintner et al., 2011).

Upon target binding, the helicase/nuclease Cas3 is recruited to the site and processively degrades the foreign DNA (Hochstrasser et al., 2014; Mulepati and Bailey, 2011; Sinkunas et al., 2011; Sinkunas et al., 2013; Westra et al., 2012). Strikingly, when Cascade encounters a mutant PAM or protospacer that prevents Cas3 degradation, hyperactive spacer acquisition from the targeted plasmid or genome is triggered in a process called “priming” (Figure 2A) (Datsenko et al., 2012; Li et al., 2014; Richter et al., 2014; Savitskaya et al., 2013; Swarts et al., 2012). Priming increases the host’s repertoire of functional spacers, allowing the host to adapt to invaders that evade the CRISPR-Cas system by mutation. Cascade is capable of binding escape mutant target sites, and recent single-molecule studies showed that the presence of Cas1 and Cas2 allows for the recruitment of Cas3 to these sites (Blosser et al., 2015; Redding et al., 2015; Richter et al., 2014). The recruited Cas3 can then translocate in either direction, in contrast to the unidirectional movement observed at perfect targets, without degrading the target DNA (Redding et al., 2015). Cas1 and Cas2 may accompany the translocating Cas3 and

### Type I Acquisition

Acquisition in *E. coli* occurs via two mechanisms—naive and primed (Figure 2A). Naive acquisition initiates upon infection by previously unencountered DNA and relies on the Cas1-Cas2 integrase complex to recognize and acquire new spacers from foreign DNA. Overexpression of Cas1 and Cas2 in the absence of other Cas proteins leads to the acquisition of 33 bp spacers at the leader-proximal end of the CRISPR array (Datsenko et al., 2012; Yosef et al., 2012). The PAM of the *E. coli* CRISPR-Cas system was identified as 5'-AWG-3', with the G becoming the first nucleotide of the integrated spacer (Datsenko et al., 2012; Díez-Villaseñor et al., 2013; Levy et al., 2015; Nuñez et al., 2014; Savitskaya et al., 2013; Shmakov et al., 2014; Swarts et al., 2012; Yosef et al., 2012; Yosef et al., 2013). In addition to the PAM, a dinucleotide motif, AA, found at the 3' end of the protospacer was also shown to be present in a disproportionately large number of spacers (Yosef et al., 2013). A recent crystal structure of the Cas1-Cas2 complex bound to an unprocessed protospacer revealed sequence-specific contacts with the 5'-CTT-3' sequence on the PAM-comple-

be activated for protospacer selection, allowing for robust acquisition on either side of the target site.

Primed acquisition has also been shown experimentally in the *P. atrosepticum* Type I-F system, in which Cas2 and Cas3 are naturally fused as a single polypeptide that associates with Cas1, as well as in the *Haloarcula hispanica* Type I-B system, where naive acquisition was not experimentally observed (Li et al., 2014; Richter et al., 2014; Richter et al., 2012). Acquisition in *H. hispanica* also requires Cas4, a 5' → 3' exonuclease found in most Type I subtypes as well as Type II-B and Type V systems, and which might be involved in generating 3' overhangs on protospacers prior to integration (Lemak et al., 2013; Li et al., 2014; Makarova et al., 2015). Although Cas1 and Cas2 may be the minimal proteins required for spacer acquisition in some systems, the association of Cas1, Cas2, and the interference machinery allows the host to coordinate robust adaptive immunity in Type I systems.

### Self- versus Non-Self-Recognition

The mechanism underlying the preference for foreign over self DNA during protospacer selection remained poorly understood until a recent study on spacer acquisition during naive acquisition. Spacer acquisition in *E. coli* was shown to be highly dependent on DNA replication, and foreign-derived spacers were preferred over self-derived spacers by about 100- to 1,000-fold (Levy et al., 2015). Analysis of the source of self-derived spacers demonstrated that protospacers were acquired largely from genomic loci predicted to frequently generate stalled replication forks and double-stranded DNA breaks (Levy et al., 2015). Such harmful dsDNA breaks are repaired by the helicase/nuclease RecBCD complex, which degrades the broken ends until reaching a Chi-site, after which only the 5' end is degraded (Dillingham and Kowalczykowski, 2008). Due to the lower frequency of Chi sites in foreign DNA, RecBCD is predicted to preferentially degrade plasmids and viral DNA, resulting in the generation of candidate protospacer substrates for Cas1 and Cas2 (Levy et al., 2015) (Figure 2A). RecBCD degrades DNA asymmetrically, yielding single-stranded fragments ranging from tens to hundreds of nucleotides long from one strand and kilobases long from the other (Dillingham and Kowalczykowski, 2008). It is unclear how Cas1-Cas2 substrates, which are 33 bp long and partially double stranded with 3' overhangs, are generated from RecBCD products (Nuñez et al., 2015a; Nuñez et al., 2015b; Wang et al., 2015). It is possible that ssDNA products re-anneal to produce partial duplexes, followed by processing to 33 bp by an unknown mechanism prior to integration into the CRISPR locus. Recent crystal structures of Cas1-Cas2 with bound protospacer reveal that the complex defines the length of the duplex region of the protospacer via a ruler mechanism and may cleave the 3' overhangs to their final length (Nuñez et al., 2015b; Wang et al., 2015). The involvement of a helicase/nuclease in both Type I-E primed and naive acquisition (Cas3 and RecBCD, respectively), as well as in Cas4-containing subtypes, hints at a conserved mechanism for protospacer generation. It is also worth noting that RecBCD is conserved primarily in Gram-negative bacteria, while Gram-positive bacteria and archaea rely on AddAB and HerA-NurA, respectively, to fill a similar role (Blackwood et al., 2013; Dillingham and Kowalczykowski, 2008). Whether CRISPR-Cas systems in these organ-

isms have evolved to cooperate with these evolutionarily distinct machineries remains to be tested.

### Mechanism of Protospacer Integration

Cas1 and Cas2 play central roles in the acquisition of new spacers, where they function as a complex (Nuñez et al., 2014). Crystal structures of Cas1 and Cas2, with or without bound protospacer, revealed two copies of a Cas1 dimer bridged by a central Cas2 dimer (Nuñez et al., 2014; Nuñez et al., 2015b; Wang et al., 2015). Cas1 functions catalytically, while Cas2 appears to serve a primarily structural role (Arslan et al., 2014; Datsenko et al., 2012; Nuñez et al., 2014; Yosef et al., 2012).

The first insight into the mechanism of protospacer integration was gained by Southern blot analysis of the genomic CRISPR locus of *E. coli* cells overexpressing Cas1 and Cas2 (Arslan et al., 2014). This revealed integration intermediates consistent with two transesterification reactions, where each strand of the protospacer is integrated into opposite sides of the leader-proximal repeat (Figure 2B). This integrase-like model was further bolstered by the *in vitro* reconstitution of protospacer integration into a plasmid-encoded CRISPR locus using purified Cas1-Cas2 complex (Nuñez et al., 2015a). The integration reaction required double-stranded DNA protospacers with 3'-OH ends that are integrated into plasmid DNA via a direct nucleophilic transesterification reaction, reminiscent of retroviral integrases and DNA transposases (Engelman et al., 1991; Mizuuchi and Adzuma, 1991).

Although deep sequencing of *in vitro* integration products revealed preferential protospacer integration adjacent to the first repeat, confirming that Cas1-Cas2 directly recognize the CRISPR locus, integration also occurred at the borders of every repeat at varying levels (Nuñez et al., 2015a). This contrasts with spacer acquisition only occurring at the first repeat in *E. coli* *in vivo* (Datsenko et al., 2012; Swarts et al., 2012; Yosef et al., 2012). To determine if the Cas1-Cas2 complex has sequence specificity for the leader-repeat sequence, a recent study took advantage of the Cas1-catalyzed disintegration reaction, a reversal of the integration reaction also observed with retroviral integrases and transposases (Chow et al., 1992; Rollie et al., 2015). Disintegration activity was stimulated when using the correct leader-repeat border sequences, highlighting intrinsic sequence-specific recognition by Cas1. Furthermore, disintegration was faster at the leader-repeat junction compared to the repeat distal end (Rollie et al., 2015). Taken together, protospacer integration likely begins at the leader-repeat junction via sequence-specific recognition by Cas1, followed by a second nucleophilic attack at the repeat distal end. This ensures precise duplication of the first repeat, as observed *in vivo*, after DNA repair by host proteins. The integration mechanism is hypothesized to be highly specific, as almost all acquired spacers with a corresponding AAG PAM are oriented with the 5'-G at the leader-proximal end, leading to functional crRNA-dependent targeting by Cascade and Cas3 (Shmakov et al., 2014). A preference for integration in the proper orientation was observed *in vitro* when protospacers with a 5'-G were used (Nuñez et al., 2015a); however, inclusion of part of the PAM in spacers has only been observed in *E. coli*, raising the question of how Cas1-Cas2 in other systems properly orient the integration reaction.

### Type II Acquisition

While most mechanistic work on acquisition has been performed in Type I systems, recent studies in Type II systems have also shed light on key aspects of spacer acquisition. One generalizable finding in Type II systems is the dependence of acquisition on infection by defective phage (Hynes et al., 2014). A significant problem with CRISPR immunity is the time required for foreign DNA to be identified, integrated into the CRISPR locus, transcribed, processed, and assembled into an interference complex that must then begin the search for appropriate targets. Since lytic phage can kill cells within 20 min, providing insufficient time for this multi-step process, Hynes and colleagues tested the hypothesis that initial immunization takes place from infection by a defective phage. Supplementation of active phage with UV-irradiated phage or phage susceptible to a restriction-modification system stimulated spacer acquisition compared to that observed with active phage alone (Hynes et al., 2014). The authors speculate that acquisition from compromised phage might also represent the dominant mode of acquisition in wild populations, allowing for a small subset of the population to acquire resistance and escape without needing to outpace a rapidly reproducing phage.

### Type II Acquisition Machinery

Type II systems are subdivided into II-A, II-B, and II-C based on the presence or absence of an additional *cas* gene alongside the minimal complement of *cas1*, *cas2*, and *cas9*. Type II-A systems contain *csn2*, while Type II-B systems, which are least commonly found, contain *cas4* (Chylinski et al., 2014; Makarova et al., 2011b). Type II-C systems comprise only the minimal three genes. *Csn2* has been shown to be essential for acquisition in several Type II-A systems (Barrangou et al., 2007; Heler et al., 2015; Wei et al., 2015b). It forms a tetramer with a toroidal architecture that binds and slides along free DNA ends, though its function in CRISPR systems is unclear (Arslan et al., 2013; Ellinger et al., 2012; Koo et al., 2012; Lee et al., 2012). *Cas4*, discussed above, is likely involved in acquisition in Type II-B systems. Type II-C systems, which constitute the majority of identified Type II systems (Chylinski et al., 2014; Makarova et al., 2015), are possibly functional for acquisition in the absence of auxiliary acquisition factors, though in the case of the *Campylobacter jejuni* system, acquisition was only observed following infection by phage encoding a *Cas4* homolog (Hooton and Conerton, 2014).

Recently, two simultaneous studies demonstrated that, in addition to *Cas1*, *Cas2*, and *Csn2*, *Cas9* plays a necessary role in the acquisition of new spacers in Type II systems (Heler et al., 2015; Wei et al., 2015b). Both groups, one working with the CRISPR1 Type II-A system of *S. thermophilus*, the other with the Type II-A system of *Streptococcus pyogenes* and the CRISPR3 system of *S. thermophilus*, also Type II-A, showed that wild-type or catalytically inactive *Cas9* (d*Cas9*) supported robust spacer acquisition, whereas deletion of *Cas9* abolished spacer acquisition. It is proposed that *Cas9* serves to recognize PAM sites in potential protospacers and mark them for recognition by *Cas1* and *Cas2* (Figure 2A). This hypothesis was confirmed by mutating the PAM-interacting residues of *Cas9*, resulting in complete loss in PAM-specificity in the newly acquired spacers (Heler et al., 2015). This presents a striking contrast to

the *E. coli* Type I-E system, where *Cas1*-*Cas2* recognize PAM sequences independently.

Intriguingly, expression of d*Cas9* results in the acquisition of primarily self-targeting spacers, suggesting that many acquisition events lead to self-targeting and suicide (Wei et al., 2015b). Microbial populations may rely on a few individuals to acquire phage resistance while the rest succumb to infection or CRISPR-mediated suicide. Some systems, such as that found in *E. coli*, may evolve to use host processes to bias acquisition away from self-targeting. Alternatively, *S. thermophilus* might have mechanisms of self-non-self-discrimination that were masked in the strain overexpressing CRISPR proteins. Phage challenge experiments with wild-type *S. thermophilus* revealed that some sequences were acquired as spacers disproportionately often across multiple experiments, suggesting that the Type II acquisition machinery has preferences in addition to *Cas9*-dependent PAM selection, though no clear pattern emerged with respect to the genomic location or sequence of protospacers that indicated a basis for the preferences (Paez-Espino et al., 2013).

Additionally, it was demonstrated that the four proteins of the *S. pyogenes* CRISPR system (*Cas1*, *Cas2*, *Csn2*, and *Cas9*) form a complex, suggesting that *Cas9* directly recruits the acquisition proteins to potential targets (Heler et al., 2015). While drawing comparisons between the involvement of *Cas9* in acquisition and primed acquisition in Type I systems is tempting, neither group saw evidence that acquisition was affected by the presence of existing spacers matching or closely matching the infecting phage or plasmid (Heler et al., 2015; Wei et al., 2015b). In addition, while the *trans*-activating crRNA (*tracrRNA*) that forms a complex with *Cas9* and the crRNA is necessary for acquisition, it is unclear whether a corresponding crRNA is also required (Heler et al., 2015; Wei et al., 2015b). Future mechanistic work will be required to shed light on the similarities between *Cas9*-mediated spacer acquisition and the primed acquisition in Type I systems.

### Type II Protospacer Integration

The sequence requirements for protospacer integration in Type II-A systems were recently demonstrated in *S. thermophilus* (Wei et al., 2015a). Similar to *E. coli*, the leader and a single repeat were sufficient to direct integration. Furthermore, only the ten nucleotides of the leader proximal to the first repeat are required to license the integration of new spacers, in contrast to the 60 nt minimal requirement in *E. coli* (Wei et al., 2015a; Yosef et al., 2012). A limited mutational study of the repeat showed that the first two nucleotides are necessary for acquisition, while the final two nucleotides can be mutated without consequence (Wei et al., 2015a). Thus, *Cas1*-*Cas2*-catalyzed integration at the leader-repeat junction is sequence specific, while the attack at the repeat-spacer junction is determined by a ruler mechanism, in agreement with observations from experiments in the *E. coli* system (Diez-Villaseñor et al., 2013). Together, these findings support the functional conservation of the *Cas1*-*Cas2* integrase complex despite divergent mechanisms of protospacer selection between Types I and II CRISPR-Cas systems.

### CRISPR Integrases as Genome-Modifying Tools

As with many other *Cas* proteins, the *Cas1*-*Cas2* integrase complex shows promise for use in modifying genomes. While *Cas1*-*Cas2* catalyze a reaction similar to that of many integrases

and transposases, they exhibit several fundamental differences that make them uniquely suited to certain applications. Cas1-Cas2 complexes lack sequence specificity for the DNA substrate to be integrated, a property that could make the system ideal for barcoding genomes. Genome barcoding allows for tracking lineages originating from individual cells, facilitating studies of population evolution, cancer, development, and infection (Blundell and Levy, 2014). Cas1-Cas2 complexes integrate short DNA sequences, in contrast with current techniques based on recombinases that integrate entire plasmids, resulting in potential fitness costs and unwanted negative selection (Blundell and Levy, 2014). Interestingly, *in vitro* integration of DNA substrates into plasmid targets revealed integration into non-CRISPR sites (Nuñez et al., 2015a), suggesting that Cas1-Cas2 can be harnessed to integrate into a wide array of target sequences. A greater understanding of the minimal functional recognition motif for various Cas1-Cas2 integrases will facilitate the development of this technology.

### crRNP Biogenesis: Generating Molecular Sentinels for the Cell

CRISPR immune systems use RNA-programmed proteins to patrol the cell in search of DNA molecules bearing sequences complementary to the crRNA. Assembly of these molecular sentinels begins with transcription of the CRISPR locus to generate long, precursor CRISPR RNAs (pre-crRNAs), followed by processing into short crRNA guides (Brouns et al., 2008; Carte et al., 2008). The promoter is embedded within the AT-rich leader sequence upstream of the repeat-spacer array, or sometimes within the repeat sequences (Zhang et al., 2013). Here, we briefly review the processing of pre-crRNAs catalyzed by the Cas6 endoribonuclease family in Type I and III systems and a distinct processing pathway in Type II systems that involves endogenous RNase III, Cas9, and a tracrRNA. The crRNA biogenesis pathway has been extensively reviewed elsewhere (Charpentier et al., 2015; Hochstrasser and Doudna, 2015).

#### Processing by Cas6 Endoribonucleases

Type I and Type III systems employ Cas6 endoribonucleases to cleave pre-crRNAs sequence specifically within each repeat (Brouns et al., 2008; Carte et al., 2008; Haurwitz et al., 2010). Although Cas6 homologs are variable in sequence, they share a conserved cleavage mechanism that results in crRNA guides comprising an entire spacer sequence flanked by portions of the repeat sequence on the 5' and 3' ends. Mature crRNA guides consist of an 8 nt 5' handle derived from the repeat sequence and variable lengths of the repeat at the 3' handle, which is further trimmed by as-yet-unidentified cellular nuclease(s) in Type III systems (Hale et al., 2008). A notable exception is in Type I-C systems, which utilize a Cas5 variant for crRNA processing, leaving an 11 nt 5' handle and 21–26 nt at the 3' end (Garside et al., 2012; Nam et al., 2012b). In other Type I systems, Cas5 subunits serve a non-catalytic role capping the 5' end of the crRNA in Cascade complexes.

In Type I-C, I-D, I-E, and I-F systems, the repeats form stable hairpin structures that allow for structure- and sequence-specific cleavage by Cas6 at the base of the hairpin (Gesner et al., 2011; Haurwitz et al., 2010; Sashital et al., 2011). After cleavage, the hairpin constitutes the 3' handle of the crRNA. The Cas6 pro-

teins in *Haloflex volcanii* (Cas6b), *E. coli* and *T. thermophilus* (Cas6e), and *Pseudomonas aeruginosa* (Cas6f) remain stably bound to the 3' handle and eventually become part of the Cascade complex (Brendel et al., 2014; Brouns et al., 2008; Gesner et al., 2011; Haurwitz et al., 2010; Sashital et al., 2011).

Type I-A, I-B, III-A, and III-B repeat sequences are non-palindromic and predicted to be unstructured in solution (Kunin et al., 2007). Thus, the respective Cas6 is thought to rely on sequence for specificity rather than structure. Interestingly, a crystal structure of the Type I-A Cas6 bound to its cognate RNA structure reveals Cas6 inducing a 3 bp hairpin in the RNA that positions the scissile phosphate in the enzyme active site (Shao and Li, 2013). It remains unknown whether other Cas6s that recognize non-palindromic repeats have a similar mechanism of RNA stabilization. Following or concurrent with the maturation of the crRNAs, the Cas proteins involved in interference assemble into the final effector complex that functions to recognize and destroy targets bearing sequence complementarity to the crRNA. In systems where Cas6 remains bound to the crRNA, it may serve to nucleate the assembly of the subunits that constitute the effector complex backbone along the crRNA. In type III systems, the number of backbone subunits defining the complex length is variable, and any unprotected crRNA remaining is degraded (Hale et al., 2008; Staals et al., 2014).

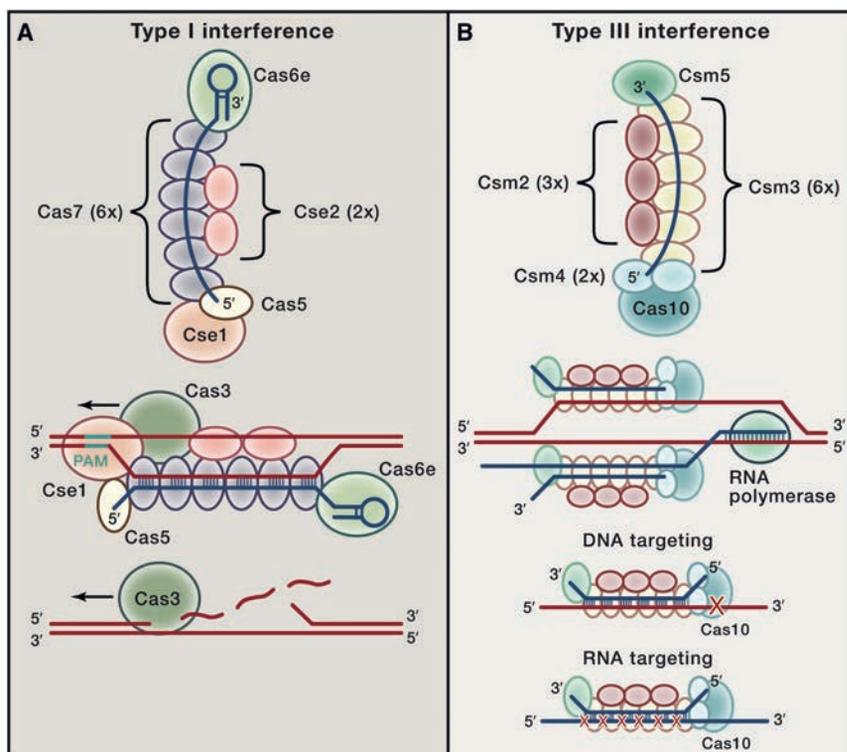
#### Processing in Type II Systems

Type II systems rely on a different mechanism to process pre-crRNAs. In Types II-A and II-B, pre-crRNA cleavage specificity is aided by a tracrRNA that has sequence complementarity to the CRISPR repeat sequence (Deltcheva et al., 2011). The gene encoding the tracrRNA is typically located either proximal to or within the CRISPR-cas locus (Chyliński et al., 2014). Upon crRNA:tracrRNA base pairing, which is stabilized by Cas9, endogenous RNase III cleaves the pre-crRNA at the repeat. The reliance on RNase III, which is not found in archaea, may explain why Type II systems are limited to bacteria (Garrett et al., 2015). An unknown nuclease trims the 5' end of the crRNA to remove the flanking repeat sequence and portions of the spacer. In *S. pyogenes*, the 30 nt spacer sequence is trimmed to the 20 nt that base-pairs with complementary foreign sequences during interference (Deltcheva et al., 2011; Jinek et al., 2012).

In the *Neisseria meningitidis* and *C. jejuni* Type II-C systems, each repeat sequence encodes a promoter, resulting in varying lengths of pre-crRNAs depending on the transcription start site (Dugar et al., 2013; Zhang et al., 2013). Although RNase III-mediated pre-crRNA processing can still occur, RNase III is dispensable for interference in these systems (Zhang et al., 2013). Thus, Cas9 is able to complex with the pre-crRNA and unprocessed tracrRNA for functional target interference without further processing of the pre-crRNAs.

#### Cas6 as a Biotechnology Tool

The Cas6 homolog from Type I-F systems, Cas6f (also known as Csy4), was the first Cas protein to be repurposed as a tool. Following demonstration of the sequence specificity of Cas6f binding and cleavage, the protein has been used for the purification of tagged RNA transcripts from cells (Haurwitz et al., 2010; Lee et al., 2013; Salvail-Lacoste et al., 2013; Sternberg et al., 2012). Subsequent studies showed that Cas6f could be used to alter the translation and stability of tagged mRNAs, allowing



**Figure 3. Interference by Class 1 Systems**

(A) Interference in Type I systems is carried out by Cascade and Cas3. Cascade is a large complex composed of the crRNA, bound at either end by Cas5 and Cas6, multiple Cas7 subunits along the crRNA, a large subunit (Cse1, Csy1, Cas8, or Cas10), and sometimes small subunits (Cse2 and Csa5). The Type I-E complex is schematized here. The large subunit recognizes the PAM in foreign DNA and initiates unwinding of the target DNA and annealing to the crRNA. Cas3 is recruited to the resulting R-loop and makes a nick. It then translocates along the displaced strand and processively degrades it.

(B) Type III systems contain either Csm or Cmr complexes, which share a similar architecture. The Csm complex from Type III-A systems is shown here. The crRNA is bound at either end by Csm5/Cmr1 and Csm4/Cmr3, which have homology to Cas6 and Cas5, respectively. Csm3/Cmr4 form the backbone of the complex, Cas10 serves as the large subunit, and Csm2/Cmr5 are the small subunit. These complexes can target both RNA and actively transcribed DNA. Cas10 catalyzes cleavage of target DNA, while the backbone subunit catalyzes cleavage of the target RNA at every sixth base, which is unpaired with the crRNA. Rather than recognizing a PAM sequence, these complexes only cleave if the 5' and 3' handles of the crRNA do not anneal to the target.

for post-transcriptional regulation of protein expression (Borchardt et al., 2015; Du et al., 2015; Nissim et al., 2014). Cas6f has also been used alongside Cas9 to process multiple guide RNAs from a single transcript, greatly facilitating multiplexed editing (Tsai et al., 2014).

### Interference: Precise, Programmable DNA Binding and Cleavage

Implementation of CRISPR systems to provide immunity involves RNA-guided recognition and precision cutting of DNA molecules, a property that makes them useful for genome engineering and control of gene expression. The extreme diversity of the crRNP targeting complexes is largely responsible for the variability observed in different CRISPR types. Whereas Types I and III use multi-protein complexes, Types II and V rely on a single protein for interference. Extensive studies have elucidated the mechanisms and structures of several complexes from each of the three major types, revealing the commonality of target binding through crRNA base-pairing and high divergence in the machineries and modes of target cleavage. For more in-depth recent reviews focused exclusively on CRISPR interference, refer to Tsui and Li (2015) and Plagens et al. (2015).

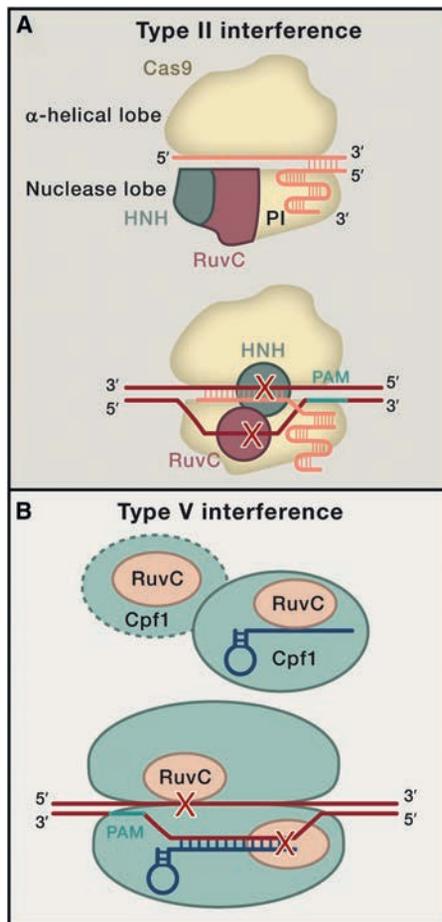
#### Type I Interference

In Type I systems, the roles of target DNA recognition and degradation are segregated into two distinct components. The crRNA-guided Cascade complex binds and unwinds the DNA target sequence (Brouns et al., 2008) and then recruits Cas3 to degrade the target in a processive manner through the combined action of its HD nuclease and helicase domains (Figure 3A) (Makarova et al., 2011b; Mulepati and Bailey, 2013; Sinkunas et al., 2013; Westra et al., 2012). Each Type I subtype (I-A through I-F) has

a distinct complement of Cascade components and, in some cases, significant variation of the *cas3* gene (Makarova et al., 2011b).

The *E. coli* Cascade complex has served as the model system for understanding the mechanism of Type I interference. In addition to the central 61 nt crRNA bearing the 32 nt spacer, the complex comprises five proteins in different stoichiometries: (Cse1)<sub>1</sub>, (Cse2)<sub>2</sub>, (Cas5)<sub>1</sub>, (Cas7)<sub>6</sub>, and (Cas6)<sub>1</sub>. The Cas7 subunits form the “backbone” that polymerizes along the crRNA and determines the crescent-shaped, semi-helical architecture seen in all structurally characterized Cascade complexes (Hochstrasser et al., 2014; Jackson et al., 2014; Jore et al., 2011; Mulepati et al., 2014; Wiedenheft et al., 2011a; Zhao et al., 2014). Cas6 (Cas6e in Type I-E systems) remains bound to the 3' hairpin following CRISPR maturation, while Cas5 binds the 5' handle (Brouns et al., 2008; Jore et al., 2011). A “small subunit” (Cse2 in Type I-E) is often found in two copies forming the “belly” of the structure and helps stabilize the crRNA and target DNA (Jackson et al., 2014; Mulepati et al., 2014; Zhao et al., 2014). A “large subunit” (Cse1 in Type I-E, Cas8 in most other subtypes) binds at the 5' end of the crRNA and recognizes the PAM sequences and recruits Cas3 to an authenticated target (Figure 3A) (Hochstrasser et al., 2014; Sashital et al., 2012). While Cas6 does not always remain with the complex and the small subunit is often found as a fusion with the large subunit, the overall architecture of Cascade complexes is generally conserved (Makarova et al., 2011b; Plagens et al., 2012; Sokolowski et al., 2014).

Cascade pre-arranges the spacer segment of the crRNA in six five-base segments of pseudo A-form conformation, with the sixth base flipped out and bound by a Cas7 subunit (Jackson et al., 2014; Mulepati et al., 2014; Zhao et al., 2014). To initiate



**Figure 4. Interference by Class 2 Systems**

(A) In Type II systems, Cas9 forms the effector complex with a crRNA and a tracrRNA. Cas9 is composed of the nuclease lobe and the  $\alpha$ -helical lobe. The nuclease lobe contains both the HNH and RuvC-like nuclease domains as well as the PAM-interacting domain. The 3' hairpins of the tracrRNA bind the nuclease lobe, while the stemloop and spacer line the channel between the two lobes. Binding to a matching, PAM-adjacent target causes the HNH domain to move into position to cleave the annealed strand, while the displaced strand is fed into the RuvC active site for cleavage.

(B) Cpf1 is the effector protein in Type V-A systems, the best characterized Type V subtype. It binds the crRNA alone. The structure of Cpf1 is unknown, but it contains an active RuvC-like nuclease domain for target cleavage. Cpf1 recognizes a PAM and makes two staggered cuts in a matching sequence. It has been proposed that Cpf1 acts as a dimer, with each monomer providing a RuvC active site, though there may be another unidentified nuclease domain.

interference, Cascade first recognizes trinucleotide PAM sites in the target strand of foreign DNA through specific interactions with Cse1 (Sashital et al., 2012). Upon PAM binding, the DNA target is unwound starting at the PAM-proximal end of the proto-spacer to form an R loop structure (Hochstrasser et al., 2014; Rollins et al., 2015; Rutkauskas et al., 2015; Sashital et al., 2012; Szczelkun et al., 2014; van Erp et al., 2015). Each stretch of five exposed bases in the crRNA is free to bind the target DNA, leading to a stable but highly distorted and discontinuous crRNA:target strand duplex (Mulepati et al., 2014; Szczelkun et al., 2014). Cascade undergoes a conformational change upon target binding that enables recruitment of Cas3 to the

Cse1 subunit (Hochstrasser et al., 2014; Mulepati et al., 2014). Cas3 binds and nicks the displaced strand using its catalytic center of the HD nuclease domain (Gong et al., 2014; Huo et al., 2014; Mulepati and Bailey, 2013; Sinkunas et al., 2013; Westra et al., 2012). The ATP-dependent helicase activity of Cas3 is then activated, causing metal- and ATP-dependent 3'  $\rightarrow$  5' translocation and processive degradation of the non-target strand (Gong et al., 2014; Huo et al., 2014; Westra et al., 2012). Cas3 initially degrades only 200–300 nt of the nontarget strand, though it continues translocating for many kilobases (Redding et al., 2015). Exposed ssDNA on the target strand may then become a substrate for other ssDNA nucleases or an additional Cas3 molecule to complete the degradation of foreign DNA (Mulepati and Bailey, 2013; Redding et al., 2015; Sinkunas et al., 2013). In addition to the PAM, target interference also relies on a seed region at the 3' end of the spacer segment of the crRNA (Semenova et al., 2011; Wiedenheft et al., 2011b). Single point mutations of the seed region of the *E. coli* Cascade complex, at the 1 to 5 and 7 to 8 position of the spacer, is enough to decrease target DNA binding and subsequent interference (Semenova et al., 2011).

Differences in the *cas3* gene among Type I subtypes suggest some variability in interference mechanism. In some Type I-E species, Cas3 is fused to Cse1 by a linker that allows it to stably associate with the Cascade complex (Westra et al., 2012). In Type I-A systems, the Cas3 helicase and nuclease domains exist as separate polypeptides that both associate with the Cascade complex (Plagens et al., 2014). In Type I-F systems, Cas3 is fused to Cas2, lending further genetic support for the interaction between the interference and acquisition machinery during primed acquisition (Makarova et al., 2015; Richter and Fineran, 2013; Richter et al., 2012). How these fusions and domain separations affect the processive degradation observed in Type I-E systems requires further study.

### Type II Interference

In contrast to the multi-subunit effector complexes seen in Type I and Type III systems (but similar to Cpf1 of Type V systems), the Type II signature protein Cas9 functions as an individual protein, along with a crRNA and tracrRNA, to interrogate DNA targets and destroy matching sequences by cleaving both strands of the target (Figure 4A) (Gasiunas et al., 2012; Jinek et al., 2012). Extensive studies on Cas9 have yielded a range of structures of *S. pyogenes* Cas9 in different substrate-bound states, as well as structures of several orthologs (Anders et al., 2014; Jiang et al., 2015; Jinek et al., 2014; Nishimasu et al., 2015; Nishimasu et al., 2014). Many of these structures, as well as the mechanism of Cas9 target search and recognition, are reviewed elsewhere (van der Oost et al., 2014); here, we focus on the most recent advances.

Structures of Cas9 have revealed two distinct lobes, the nuclease lobe and the  $\alpha$ -helical or REC lobe (Anders et al., 2014; Jinek et al., 2014; Nishimasu et al., 2015; Nishimasu et al., 2014). The nuclease lobe is composed of the HNH nuclease domain, which cleaves the target strand, a RuvC-like nuclease domain, which cleaves the non-target strand and is separated into three distinct regions in the primary sequence by the intervening  $\alpha$ -helical lobe and the HNH domain, and a C-terminal PAM-interacting domain (Anders et al., 2014; Jinek

et al., 2014; Nishimasu et al., 2015; Nishimasu et al., 2014). The  $\alpha$ -helical lobe contains an arginine-rich “bridge helix,” which connects the two lobes and interacts with the guide RNA, and is the most variable region of Cas9, with insertions or deletions accounting for much of the wide variation in size seen in Cas9 orthologs (Chylinski et al., 2014; Jinek et al., 2014; Nishimasu et al., 2014).

Cas9 initiates its target search by probing duplexed DNA for an appropriate PAM before initiating target unwinding (Sternberg et al., 2014). The target unwinds from the seed region, the first 10–12 nucleotides following the PAM, toward the PAM-distal end (Szczelkun et al., 2014). A perfect or near-perfect match leads to cleavage of both DNA strands, with mismatches being more tolerated outside of the seed region (Cong et al., 2013; Jiang et al., 2013; Jinek et al., 2012; Sternberg et al., 2014). The mechanism by which mismatches distant from the cleavage site prevent cleavage appears to rely on the structural flexibility of the HNH domain, which has yet to be crystallized in proximity to the scissile phosphate (Anders et al., 2014; Nishimasu et al., 2015; Nishimasu et al., 2014). FRET assays show that the HNH domain swings into a catalytically competent position only upon binding to a cognate double-stranded DNA substrate, underscoring the multiple steps of conformational control of Cas9-catalyzed DNA cleavage (Sternberg et al., 2015). The RuvC domain is in turn allosterically regulated by the HNH domain. Cleavage of the non-target strand requires movement of the HNH domain into an active position, even when the mismatched substrates allow full unwinding of the non-target strand (Sternberg et al., 2015).

Recent crystal structures of *S. pyogenes* Cas9-sgRNA surveillance complex and of the smaller *Staphylococcus aureus* Cas9 in a target-bound state provided new insights into Cas9 function (Jiang et al., 2015; Nishimasu et al., 2015). The sgRNA-bound structure revealed how binding of sgRNA shifts Cas9 from the auto-inhibited state observed in the apo form to a conformation competent for target search (Jiang et al., 2015; Jinek et al., 2014). As previously observed in low-resolution electron microscopy structures, a nucleic acid binding cleft is formed between the two lobes upon sgRNA binding (Jinek et al., 2014). Furthermore, two PAM-interacting arginine residues are pre-positioned to allow for scanning of potential target DNA, a finding that may explain the necessity of tracrRNA in directing PAM-dependent spacer acquisition. Surprisingly, while the 3' hairpins of the tracrRNA have been shown to provide nearly all of the binding energy and specificity for Cas9, the repeat-anti-repeat region of the sgRNA as well as the seed sequence were required to induce the conformational rearrangement (Briner et al., 2014; Jiang et al., 2015; Wright et al., 2015). The seed sequence of the sgRNA was also found to be pre-ordered in an A-form helix, analogous to the pre-ordered seed region of guide RNA observed in eukaryotic Argonaute structures and the Type I and Type III effector complexes, where the entire crRNA is pre-arranged in a target-binding-competent state (Jackson et al., 2014; Kuhn and Joshua-Tor, 2013; Mulepati et al., 2014; Osawa et al., 2015; Taylor et al., 2015; Zhao et al., 2014). The observed pre-ordering of the guide RNA provides an energetic compensation for the unwinding of the target duplex to facilitate binding.

Cas9 from the Type II-C CRISPR system of *S. aureus* was crystallized in complex with sgRNA and a single-stranded DNA target sequence, providing insight into the structural variation between more distantly related Cas9 (Nishimasu et al., 2015). *S. aureus* Cas9 is significantly smaller than the Cas9 of *S. pyogenes* (1,053 versus 1,368 amino acids) and recognizes a significantly different guide RNA and PAM site. The *S. aureus* Cas9 structure revealed a smaller  $\alpha$ -helical lobe, with domains in the middle and PAM-proximal side notably absent, while the nuclease lobe is largely conserved (Nishimasu et al., 2015). The authors proposed a new domain designation, the wedge domain, which diverges significantly between the two proteins and appears integral to determining guide RNA orthogonality. Another small Cas9, that from *Actinomyces naeslundii*, was previously crystallized in the apo form, but the absence of bound substrate and significant disordered regions limited detailed exploration of the differences between the orthologs (Jinek et al., 2014). Other recent work with Type II-C Cas9 proteins from *N. meningitidis* and *Corynebacterium diphtheriae*, among other Type II-C orthologs, revealed that these enzymes have a reduced ability to unwind dsDNA compared to *S. pyogenes* Cas9 and exhibit efficient PAM-independent and in some cases tracrRNA-independent cleavage of ssDNA (Ma et al., 2015; Zhang et al., 2015). This activity may allow for more efficient interference with ssDNA plasmid or phage or represent a more ancestral activity that predates the expansion of the  $\alpha$ -helical lobe to facilitate more robust DNA unwinding.

### Type III Interference

Type III systems are classified into Type III-A and Type III-B based on their effector complexes (Type III-C and III-D have also been identified, but not yet characterized) (Makarova et al., 2015). The former is constituted by the Csm complex, and the latter by the Cmr complex (Makarova et al., 2011b). Phylogenetic studies suggested that some *csm* and *cmr* genes are distant homologs of *cas* genes that compose the Cascade complex of Type I systems, and subsequent structural studies have revealed a striking structural conservation between Cascade and the Csm and Cmr complexes (Hochstrasser et al., 2014; Jackson et al., 2014; Makarova et al., 2013; Mulepati et al., 2014; Osawa et al., 2015; Staals et al., 2014; Taylor et al., 2015; Zhao et al., 2014). For a detailed discussion of the structural similarities between these complexes, refer to Jackson and Wiedenheft (2015). Briefly, Csm3 (in III-A systems) or Cmr4 (in III-B) polymerizes along the crRNA as a helical backbone, analogously to Cas7, while Csm2 or Cmr5 take the role of Cse2 as the small subunit (Figure 3B) (Jackson and Wiedenheft, 2015). Similar to Cascade, the crRNA is pre-arranged for binding with kinks every six nucleotides. The target nucleic acid (RNA in all solved Type III structures) binds in a distorted manner, forming five-nucleotide helical stretches with the sixth base flipped out to allow for the extreme deviation from helical nucleic acid observed in all structures (Osawa et al., 2015; Taylor et al., 2015). Cmr3 and Csm4 bind the 5' crRNA handle, while Cas10 (also referred to as Csm1 and Cmr2) serves as the large subunit (Makarova et al., 2011a; Osawa et al., 2015; Staals et al., 2014; Taylor et al., 2015). Csm5, Cmr6, and Cmr1 also share homology with Cas7 and cap the helical backbone at the 3' end of the crRNA. In Type III-B systems, two major crRNA species are

generally observed, differing by six nucleotides (Juranek et al., 2012; Staals et al., 2014). Cryo-electron microscopy captured two Cmr complexes of different sizes, with one complex having one fewer Cmr4 and Cmr5 subunit, suggesting that the different crRNA lengths are the result of different complex sizes, or vice versa (Taylor et al., 2015).

Despite the structural similarities, the Type III interference complexes function quite distinctly from Cascade. The substrate specificity of Csm and Cmr complexes has only recently been clarified. Early in vivo genetic experiments suggested Csm targeted DNA, while in vitro studies of Cmr showed binding and cleavage activity against RNA only (Hale et al., 2009; Marraffini and Sontheimer, 2008), leading to a model wherein the two subtypes had evolved distinct and complementary substrate preferences. This simple model was soon complicated by the observation that Csm complexes in vitro also bind and cleave RNA while exhibiting no activity against DNA (Staals et al., 2014; Tamulaitis et al., 2014). Meanwhile, the in vivo DNA-targeting activity of III-A systems was shown to depend on transcription at the target site, in contrast to the transcription-independent targeting seen in Type I and Type II systems, and a similar activity was observed for a III-B system in vivo (Deng et al., 2013; Goldberg et al., 2014). These observations were reconciled by the discovery that the Csm complex from *Staphylococcus epidermidis* exhibits both RNA cleavage and DNA cleavage when directed against the non-template strand of actively transcribed DNA (Samai et al., 2015).

DNA and RNA interference are carried out by distinct subunits of the Type III complexes. RNA interference is mediated by the backbone subunit Csm3 (or Cmr4 in III-B systems), which cleaves the target every six nucleotides in the active site of a separate subunit by activating the ribose 2' OH for nucleophilic attack in a manner typical of metal-independent RNases (Osawa et al., 2015; Staals et al., 2014; Tamulaitis et al., 2014; Taylor et al., 2015). Cas10 cleaves DNA exposed by a transcription bubble using a single catalytic site in its palm polymerase domain (Samai et al., 2015). The details of DNA targeting by Cmr have not been independently confirmed, but the conservation of Cas10 and evidence for transcription-dependent plasmid clearing supports a similar mechanism (Deng et al., 2013; Makarova et al., 2011b).

The distinct behavior of Type III systems provides the host microbe with the ability to tolerate temperate phages (Goldberg et al., 2014). While Type I and Type II systems target and degrade any protospacer-containing DNA, Type III systems ignore foreign DNA until transcription begins that poses a threat to the cell. This has the advantage of allowing cells to acquire advantageous genes contained in prophages, such as antibiotic resistance genes, and causing cell suicide in the event that a lysogenic phage becomes lytic and begins transcribing genes with matching spacers (Goldberg et al., 2014). However, the strand-specific nature of both the RNA targeting and transcription-dependent DNA targeting imposes an additional restriction on the integration step of acquisition, as only one direction of integration will yield productive interference. The means by which this apparent limitation is overcome are unclear. Type III systems are also frequently found coexisting with Type I systems, in which case their distinct target specificity might allow

for interference with targets that somehow avoid recognition by Cascade (Makarova et al., 2011b).

Type III systems are also unique in their lack of a PAM. Rather than recognizing a distinct motif to avoid auto-immunity at the CRISPR locus, the Csm and Cmr complexes instead check for complementarity between the repeat-derived region of the crRNA with the target and do not cleave if a full match is detected (Marraffini and Sontheimer, 2010; Samai et al., 2015; Staals et al., 2014; Tamulaitis et al., 2014). The specificity of Type III effector complexes for single-stranded targets might provide a rationale for their distinct mode of target authentication. For Type I and Type II effector complexes, which target dsDNA, PAM recognition allows for an initial binding event to facilitate subsequent unwinding of the target to probe for complementarity to the crRNA (Hochstrasser et al., 2014; Rollins et al., 2015; Sternberg et al., 2014; Szczelkun et al., 2014; Westra et al., 2012). Type III complexes can immediately probe a potential single-stranded target for complementarity to their bound crRNA without a need to license initial unwinding, and the exposed nature of a single-stranded target facilitates the check for complementarity to the repeat-derived region of the guide.

#### Type V Interference

Type V systems have only recently been classified, but initial work demonstrated that these systems are functional for interference (Makarova et al., 2015; Zetsche et al., 2015a). The systems appear most similar to Type II systems, possessing only the acquisition machinery and a single additional protein (Makarova et al., 2015; Schunder et al., 2013; Vestergaard et al., 2014). Three subtypes of Class V systems have been identified with widely varying interference proteins (Shmakov et al., 2015). Type V-A, V-B, and V-C are characterized by the presence of Cpf1, C2c1, and C2c3, respectively (Shmakov et al., 2015). All three proteins are evolved from the same family of transposon-associated TpnB proteins as Cas9 and have a C-terminal RuvC domain and arginine-rich bridge helix (Shmakov et al., 2015). However, the proteins show little similarity to each other, and the phylogenetic grouping of the associated *cas1* genes with various branches of Type I and Type III *cas1* genes suggests that each of these subtypes originated from distinct recombination events between CRISPR systems and *tpnB* genes (Shmakov et al., 2015).

While some Type V-B systems have an identifiable tracrRNA necessary for activity, Type V-A and V-C systems lack both a tracrRNA and Cas6 or Cas5-like endonuclease, making it unclear how the crRNA is processed (Makarova et al., 2015; Shmakov et al., 2015). The crRNA of Type V-A systems has a conserved stem-loop and can be processed to a functional form when transcribed in *E. coli* in the presence of Cpf1 (Zetsche et al., 2015a). Whether Cpf1 is also required for processing and the potential involvement of host factors remains unknown. The Cpf1 from *Francisella novicida* can successfully interfere with transformed plasmids and recognizes a 5'-TTN-3' PAM at the 5' end of the protospacer sequence, similar to the PAM location of Type I systems and in contrast to the 3' PAM observed in Type II systems. The enzyme makes a double-strand break, resulting in five-nucleotide 5' overhangs distal to the PAM site (Figure 4B). Mutation of catalytic residues in the RuvC active site prevents cleavage of either strand (Zetsche et al., 2015a).

The authors propose that Cpf1 might act as a dimer, with each monomer providing a RuvC active site but only one recognizing the target. If this is the case, whether one or both monomers has a bound crRNA is unclear. Alternatively, an as-of-yet undiscovered active site might be present in the protein, in which case its activity must be tightly coupled to that of RuvC to explain the phenotype observed for the RuvC mutant. A C2c1, which also has only one identifiable nuclease domain, has also been shown to be active for cleavage *in vivo* and *in vitro*, where it recognizes a 5'-TTN-3' PAM and requires a tracrRNA (Shmakov et al., 2015). Many mechanisms in these newly discovered systems, both Type V and the essentially uncharacterized Type VI, remain unknown and open for future study.

### Interference Complexes as Genome Editing Tools

Most tool development of Cas proteins has focused on exploiting the programmable sequence-specific DNA recognition of interference complexes. Cas9 from *S. pyogenes* in particular has proven enormously useful for genome engineering. The ability to render Cas9 a two-component system by fusing the crRNA and tracrRNA into a single guide RNA (sgRNA) has allowed for its easy use for genome editing, transcriptional control, RNA targeting, and imaging (for recent reviews, see Jiang and Marraffini, 2015; Sternberg and Doudna, 2015). Cas9 has been used in various cell types and organisms ranging from mice and monkeys to primary human T cells and stem cells, as well as plants, bacteria, and fungi (Jiang and Marraffini, 2015; Sternberg and Doudna, 2015). Recent work has focused on developing various chemical- and light-inducible Cas9 constructs to allow for greater spatiotemporal control and on employing Cas9 orthologs with different PAM sequences and smaller sizes, allowing for easier packaging in adeno-associated virus vectors (Davis et al., 2015; Nihongaki et al., 2015; Polstein and Gersbach, 2015; Ran et al., 2015; Zetsche et al., 2015b).

Other interference complexes have already been used or have the potential to be useful for genome manipulation as well. Although the multi-subunit composition of Cascade makes it less tractable for genome engineering compared to Cas9, its large size and stable binding has been used for transcriptional silencing in *E. coli* (Rath et al., 2015). No published work has shown the application of Csm or Cmr complexes, but either could likely be used for various RNA modulation applications in cells. Two Cpf1 homologs, out of 16 that were tested, have been shown to facilitate genome editing in human cells (Zetsche et al., 2015a). The alternate PAM specificity of Cpf1 may prove useful for targeting sites without an appropriate PAM for Cas9, and the staggered cuts might prove to favor distinct pathways of DNA repair. However, a thorough investigation of the efficiency and off-target editing of Cpf1 will be needed to determine if this protein will see significant use alongside Cas9.

While Cas9 has already seen extensive use in the research setting, challenges remain for its application in the clinic. While making programmed cuts has become largely trivial, biasing DNA repair toward homology-directed repair rather than non-homologous end joining remains a challenge (Chu et al., 2015; Maruyama et al., 2015). Delivery of Cas9, either as an RNP or on a plasmid or viral vector, to particular tissues in whole organisms

is another challenge that must be addressed to enable clinical applications (D'Astolfo et al., 2015; Gori et al., 2015; Howes and Schofield, 2015; Lin et al., 2014; Zuris et al., 2015). As the field continues to advance rapidly, clinical trials may occur within a few years, with therapies possibly following within a decade. Engineering of crop plants with Cas9 is already underway; regulatory rulings have so far considered knockout plants not to be genetically modified organisms, but the regulatory fate of other modifications is currently being considered (Servick, 2015).

### Concluding Remarks

Despite the rapid progress of the field since the first demonstration of CRISPR immunity in 2007, many mechanistic questions remain unanswered. Fundamental aspects of acquisition, such as how substrates for Cas1-Cas2-mediated integration are generated and the mechanism and extent of self- versus non-self-discrimination in different CRISPR subtypes, are still a mystery. While crRNA biogenesis and interference are reasonably well understood for certain model subtypes (Type I-E, Type II-A), the sheer diversity of CRISPR systems means that many subtypes with potentially distinct mechanisms remain unexplored. Type V and VI systems have only begun to be analyzed, and Type IV systems, bearing some familiar cas genes but no identifiable CRISPR locus, have yet to be characterized experimentally and almost certainly rely on mechanisms distinct from those of traditional CRISPR systems (Makarova and Koonin, 2015).

Other aspects of CRISPR-Cas systems lie beyond the scope of this Review. We have not discussed the non-immune functions of CRISPR-Cas systems, some of which appear to have evolved to serve regulatory rather than defense roles (for reviews, see Westra et al., 2014, and Ratner et al., 2015). Phage evasion of CRISPR immunity is another active area of research, with identified mechanisms including DNA modification, specialized anti-CRISPR proteins, and mutational escape (Bondy-Denomy et al., 2013; Bondy-Denomy et al., 2015; Bryson et al., 2015; Deveau et al., 2008; Paez-Espino et al., 2015; Pawluk et al., 2014). The context-dependent regulation of CRISPR-Cas systems in response to phage infection and stress signals has also been explored but requires further study (Bondy-Denomy and Davidson, 2014; Garrett et al., 2015; Kenchappa et al., 2013; Patterson et al., 2015; Pul et al., 2010). The rapid development of technology derived from CRISPR-Cas systems, most notably Cas9 but also Cas6f/Csy4, Cascade, and Cpf1, has fueled intense interest in the field. The arms race between bacteria and bacteriophage has generated powerful molecular biology tools, from restriction enzymes that enabled recombinant DNA technology to Cas9, which started the "CRISPR revolution" in modern genome engineering. CRISPR systems have proven to be both fascinating and enormously useful. Further study of bacterial immune systems, both CRISPR systems and those yet undiscovered, promises to yield further unforeseen discoveries and exciting new technologies.

### ACKNOWLEDGMENTS

This work was funded by US National Science Foundation grant number 1244557 to J.A.D. A.V.W. and J.K.N. are NSF Graduate Research Fellows. Megan Hochstrasser provided valuable input on the manuscript.

## REFERENCES

- Anders, C., Niewoehner, O., Duerst, A., and Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 513, 569–573.
- Arslan, Z., Wurm, R., Brener, O., Ellinger, P., Nagel-Steger, L., Oesterhelt, F., Schmitt, L., Willbold, D., Wagner, R., Gohlke, H., et al. (2013). Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2. *Nucleic Acids Res.* 41, 6347–6359.
- Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., and Pul, Ü. (2014). Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system. *Nucleic Acids Res.* 42, 7884–7893.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712.
- Blackwood, J.K., Rzechorzek, N.J., Bray, S.M., Maman, J.D., Pellegrini, L., and Robinson, N.P. (2013). End-resection at DNA double-strand breaks in the three domains of life. *Biochem. Soc. Trans.* 41, 314–320.
- Blosser, T.R., Loeff, L., Westra, E.R., Vlot, M., Künne, T., Sobota, M., Dekker, C., Brouns, S.J., and Joo, C. (2015). Two distinct DNA binding modes guide dual roles of a CRISPR-Cas protein complex. *Mol. Cell* 58, 60–70.
- Blundell, J.R., and Levy, S.F. (2014). Beyond genome sequencing: Lineage tracking with barcodes to study the dynamics of evolution, infection, and cancer. *Genomics* 104, 417–430.
- Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151, 2551–2561.
- Bondy-Denomy, J., and Davidson, A.R. (2014). To acquire or resist: the complex biological effects of CRISPR-Cas systems. *Trends Microbiol.* 22, 218–225.
- Bondy-Denomy, J., Pawluk, A., Maxwell, K.L., and Davidson, A.R. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* 493, 429–432.
- Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M.F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K.L., and Davidson, A.R. (2015). Multiple mechanisms for CRISPR-Cas inhibition by anti-CRISPR proteins. *Nature* 526, 136–139.
- Borchardt, E.K., Vadoros, L.A., Huang, M., Lackey, P.E., Marzluff, W.F., and Asokan, A. (2015). Controlling mRNA stability and translation with the CRISPR endoribonuclease Csy4. *RNA* 21, 1921–1930.
- Brendel, J., Stoll, B., Lange, S.J., Sharma, K., Lenz, C., Stachler, A.E., Maier, L.K., Richter, H., Nickel, L., Schmitz, R.A., et al. (2014). A complex of Cas proteins 5, 6, and 7 is required for the biogenesis and stability of clustered regularly interspaced short palindromic repeats (crispr)-derived rnas (crnas) in *Haloflex volcanii*. *J. Biol. Chem.* 289, 7164–7177.
- Briner, A.E., Donohoue, P.D., Gomaa, A.A., Selle, K., Slorach, E.M., Nye, C.H., Haurwitz, R.E., Beisel, C.L., May, A.P., and Barrangou, R. (2014). Guide RNA functional modules direct Cas9 activity and orthogonality. *Mol. Cell* 56, 333–339.
- Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuys, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964.
- Bryson, A.L., Hwang, Y., Sherrill-Mix, S., Wu, G.D., Lewis, J.D., Black, L., Clark, T.A., and Bushman, F.D. (2015). Covalent Modification of Bacteriophage T4 DNA Inhibits CRISPR-Cas9. *MBio* 6, e00648.
- Carte, J., Wang, R., Li, H., Terns, R.M., and Terns, M.P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* 22, 3489–3496.
- Charpentier, E., Richter, H., van der Oost, J., and White, M.F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol. Rev.* 39, 428–441.
- Chow, S.A., Vincent, K.A., Ellison, V., and Brown, P.O. (1992). Reversal of integration and DNA splicing mediated by integrase of human immunodeficiency virus. *Science* 255, 723–726.
- Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* 33, 543–548.
- Chylinski, K., Makarova, K.S., Charpentier, E., and Koonin, E.V. (2014). Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res.* 42, 6091–6105.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823.
- D'Astolfo, D.S., Pagliero, R.J., Pras, A., Karthaus, W.R., Clevers, H., Prasad, V., Lebbink, R.J., Rehmann, H., and Geijsen, N. (2015). Efficient intracellular delivery of native proteins. *Cell* 161, 674–690.
- Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K., and Semenova, E. (2012). Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat. Commun.* 3, 945.
- Davis, K.M., Pattanayak, V., Thompson, D.B., Zuris, J.A., and Liu, D.R. (2015). Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat. Chem. Biol.* 11, 316–318.
- Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607.
- Deng, L., Garrett, R.A., Shah, S.A., Peng, X., and She, Q. (2013). A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol. Microbiol.* 87, 1088–1099.
- Deveau, H., Barrangou, R., Garneau, J.E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390–1400.
- Díez-Villaseñor, C., Guzmán, N.M., Almendros, C., García-Martínez, J., and Mojica, F.J. (2013). CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli*. *RNA Biol.* 10, 792–802.
- Dillingham, M.S., and Kowalczykowski, S.C. (2008). RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol. Mol. Biol. Rev.* 72, 642–671.
- Du, P., Miao, C., Lou, Q., Wang, Z., and Lou, C. (2015). Engineering translational activators with CRISPR-Cas system. *ACS Synth. Biol.* <http://dx.doi.org/10.1021/acssynbio.5b00130>.
- Dugar, G., Herbig, A., Förstner, K.U., Heidrich, N., Reinhardt, R., Nieselt, K., and Sharma, C.M. (2013). High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* 9, e1003495.
- Ellinger, P., Arslan, Z., Wurm, R., Tschapek, B., MacKenzie, C., Pfeffer, K., Panjikar, S., Wagner, R., Schmitt, L., Gohlke, H., et al. (2012). The crystal structure of the CRISPR-associated protein Csn2 from *Streptococcus agalactiae*. *J. Struct. Biol.* 178, 350–362.
- Engelman, A., Mizuuchi, K., and Craigie, R. (1991). HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell* 67, 1211–1221.
- Garrett, R.A., Shah, S.A., Erdmann, S., Liu, G., Mousaei, M., León-Sobrinho, C., Peng, W., Gudbergdottir, S., Deng, L., Vestergaard, G., et al. (2015). CRISPR-Cas Adaptive Immune Systems of the Sulfolobales: Unravelling Their Complexity and Diversity. *Life (Basel)* 5, 783–817.
- Garside, E.L., Schellenberg, M.J., Gesner, E.M., Bonanno, J.B., Sauder, J.M., Burley, S.K., Almo, S.C., Mehta, G., and MacMillan, A.M. (2012). Cas5d processes pre-crRNA and is a member of a larger family of CRISPR RNA endonucleases. *RNA* 18, 2020–2028.

- Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* *109*, E2579–E2586.
- Gesner, E.M., Schellenberg, M.J., Garside, E.L., George, M.M., and Macmillan, A.M. (2011). Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* *18*, 688–692.
- Goldberg, G.W., Jiang, W., Bikard, D., and Marraffini, L.A. (2014). Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* *514*, 633–637.
- Gong, B., Shin, M., Sun, J., Jung, C.-H., Bolt, E.L., van der Oost, J., and Kim, J.-S. (2014). Molecular insights into DNA interference by CRISPR-associated nuclease-helicase Cas3. *Proc. Natl. Acad. Sci. USA* *111*, 16359–16364.
- Gori, J.L., Hsu, P.D., Maeder, M.L., Shen, S., Welstead, G.G., and Bumcrot, D. (2015). Delivery and Specificity of CRISPR-Cas9 Genome Editing Technologies for Human Gene Therapy. *Hum. Gene Ther.* *26*, 443–451.
- Hale, C., Kleppe, K., Terns, R.M., and Terns, M.P. (2008). Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* *14*, 2572–2579.
- Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* *139*, 945–956.
- Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J.A. (2010). Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* *329*, 1355–1358.
- Heler, R., Samai, P., Modell, J.W., Weiner, C., Goldberg, G.W., Bikard, D., and Marraffini, L.A. (2015). Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature* *519*, 199–202.
- Hochstrasser, M.L., and Doudna, J.A. (2015). Cutting it close: CRISPR-associated endoribonuclease structure and function. *Trends Biochem. Sci.* *40*, 58–66.
- Hochstrasser, M.L., Taylor, D.W., Bhat, P., Guegler, C.K., Sternberg, S.H., Nogales, E., and Doudna, J.A. (2014). CasA mediates Cas3-catalyzed target degradation during CRISPR RNA-guided interference. *Proc. Natl. Acad. Sci. USA* *111*, 6618–6623.
- Hooton, S.P., and Connerton, I.F. (2014). *Campylobacter jejuni* acquire new host-derived CRISPR spacers when in association with bacteriophages harboring a CRISPR-like Cas4 protein. *Front Microbiol* *5*, 744.
- Howes, R., and Schofield, C. (2015). Genome engineering using Adeno-Associated Virus (AAV). *Methods Mol. Biol.* *1239*, 75–103.
- Huo, Y., Nam, K.H., Ding, F., Lee, H., Wu, L., Xiao, Y., Farchione, M.D., Jr., Zhou, S., Rajashankar, K., Kurinov, I., et al. (2014). Structures of CRISPR Cas3 offer mechanistic insights into Cascade-activated DNA unwinding and degradation. *Nat. Struct. Mol. Biol.* *21*, 771–777.
- Hynes, A.P., Villion, M., and Moineau, S. (2014). Adaptation in bacterial CRISPR-Cas immunity can be driven by defective phages. *Nat. Commun.* *5*, 4399.
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* *169*, 5429–5433.
- Jackson, R.N., and Wiedenheft, B. (2015). A Conserved Structural Chassis for Mounting Versatile CRISPR RNA-Guided Immune Responses. *Mol. Cell* *58*, 722–728.
- Jackson, R.N., Golden, S.M., van Erp, P.B., Carter, J., Westra, E.R., Brouns, S.J., van der Oost, J., Terwilliger, T.C., Read, R.J., and Wiedenheft, B. (2014). Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from *Escherichia coli*. *Science* *345*, 1473–1479.
- Jansen, R., Embden, J.D., Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* *43*, 1565–1575.
- Jiang, W., and Marraffini, L.A. (2015). CRISPR-Cas: New Tools for Genetic Manipulations from Bacterial Immunity Systems. *Annu. Rev. Microbiol.* *69*, 209–228.
- Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* *31*, 233–239.
- Jiang, F., Zhou, K., Ma, L., Gressel, S., and Doudna, J.A. (2015). STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* *348*, 1477–1481.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.
- Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., et al. (2014). Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* *343*, 1247997.
- Jore, M.M., Lundgren, M., van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R., et al. (2011). Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* *18*, 529–536.
- Juraneck, S., Eban, T., Altuvia, Y., Brown, M., Morozov, P., Tuschl, T., and Margalit, H. (2012). A genome-wide view of the expression and processing patterns of *Thermus thermophilus* HB8 CRISPR RNAs. *RNA* *18*, 783–794.
- Kenchappa, C.S., Heidarsson, P.O., Kragelund, B.B., Garrett, R.A., and Poulos, F.M. (2013). Solution properties of the archaeal CRISPR DNA repeat-binding homeodomain protein Cbp2. *Nucleic Acids Res.* *41*, 3424–3435.
- Koo, Y., Jung, D.K., and Bae, E. (2012). Crystal structure of *Streptococcus pyogenes* Csn2 reveals calcium-dependent conformational changes in its tertiary and quaternary structure. *PLoS ONE* *7*, e33401.
- Kuhn, C.-D., and Joshua-Tor, L. (2013). Eukaryotic Argonautes come into focus. *Trends Biochem. Sci.* *38*, 263–271.
- Kunin, V., Sorek, R., and Hugenholtz, P. (2007). Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol.* *8*, R61.
- Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* *8*, 317–327.
- Lee, K.-H., Lee, S.-G., Eun Lee, K., Jeon, H., Robinson, H., and Oh, B.-H. (2012). Identification, structural, and biochemical characterization of a group of large Csn2 proteins involved in CRISPR-mediated bacterial immunity. *Proteins* *80*, 2573–2582.
- Lee, H.Y., Haurwitz, R.E., Apffel, A., Zhou, K., Smart, B., Wenger, C.D., Laderman, S., Bruhn, L., and Doudna, J.A. (2013). RNA-protein analysis using a conditional CRISPR nuclease. *Proc. Natl. Acad. Sci. USA* *110*, 5416–5421.
- Lemak, S., Beloglazova, N., Nocek, B., Skarina, T., Flick, R., Brown, G., Popovic, A., Joachimiak, A., Savchenko, A., and Yakunin, A.F. (2013). Toroidal structure and DNA cleavage by the CRISPR-associated [4Fe-4S] cluster containing Cas4 nuclease SSO0001 from *Sulfolobus solfataricus*. *J. Am. Chem. Soc.* *135*, 17476–17487.
- Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U., and Sorek, R. (2015). CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* *520*, 505–510.
- Li, M., Wang, R., Zhao, D., and Xiang, H. (2014). Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. *Nucleic Acids Res.* *42*, 2483–2492.
- Lin, S., Staahl, B.T., Alla, R.K., and Doudna, J.A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* *3*, e04766.
- Lintner, N.G., Kerou, M., Brumfield, S.K., Graham, S., Liu, H., Naismith, J.H., Sdano, M., Peng, N., She, Q., Copié, V., et al. (2011). Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). *J. Biol. Chem.* *286*, 21643–21656.
- Ma, E., Harrington, L.B., O'Connell, M.R., Zhou, K., and Doudna, J.A. (2015). Single-stranded DNA cleavage by divergent CRISPR-Cas9 enzymes. *Mol. Cell* *60*, 398–407.
- Makarova, K.S., and Koonin, E.V. (2015). Annotation and Classification of CRISPR-Cas Systems. *Methods Mol. Biol.* *1311*, 47–75.

- Makarova, K.S., Aravind, L., Wolf, Y.I., and Koonin, E.V. (2011a). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct* 6, 38.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., et al. (2011b). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477.
- Makarova, K.S., Wolf, Y.I., and Koonin, E.V. (2013). The basic building blocks and evolution of CRISPR-CAS systems. *Biochem. Soc. Trans.* 41, 1392–1400.
- Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736.
- Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845.
- Marraffini, L.A., and Sontheimer, E.J. (2010). Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463, 568–571.
- Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., and Ploegh, H.L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33, 538–542.
- Mizuuchi, K., and Adzuma, K. (1991). Inversion of the phosphate chirality at the target site of Mu DNA strand transfer: evidence for a one-step transesterification mechanism. *Cell* 66, 129–140.
- Mojica, F.J., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60, 174–182.
- Mojica, F.J., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740.
- Mulepati, S., and Bailey, S. (2011). Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J. Biol. Chem.* 286, 31896–31903.
- Mulepati, S., and Bailey, S. (2013). In vitro reconstitution of an *Escherichia coli* RNA-guided immune system reveals unidirectional, ATP-dependent degradation of DNA target. *J. Biol. Chem.* 288, 22184–22192.
- Mulepati, S., Héroux, A., and Bailey, S. (2014). Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* 345, 1479–1484.
- Nam, K.H., Haitjema, C., Liu, X., Ding, F., Wang, H., DeLisa, M.P., and Ke, A. (2012b). Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/DvuII CRISPR-Cas system. *Structure* 20, 1574–1584.
- Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* 33, 755–760.
- Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* 156, 935–949.
- Nishimasu, H., Cong, L., Yan, W.X., Ran, F.A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., and Nureki, O. (2015). Crystal Structure of *Staphylococcus aureus* Cas9. *Cell* 162, 1113–1126.
- Nissim, L., Perli, S.D., Fridkin, A., Perez-Pinera, P., and Lu, T.K. (2014). Multiplexed and programmable regulation of gene networks with an integrated RNA and CRISPR/Cas toolkit in human cells. *Mol. Cell* 54, 698–710.
- Núñez, J.K., Kranzusch, P.J., Noeske, J., Wright, A.V., Davies, C.W., and Doudna, J.A. (2014). Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* 21, 528–534.
- Núñez, J.K., Lee, A.S., Engelman, A., and Doudna, J.A. (2015a). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* 519, 193–198.
- Núñez, J.K., Harrington, L.B., Kranzusch, P.J., Engelman, A.N., and Doudna, J.A. (2015b). Foreign DNA capture during CRISPR-Cas adaptive immunity. *Nature* 527, 535–538. <http://dx.doi.org/10.1038/nature15760>.
- Osawa, T., Inanaga, H., Sato, C., and Numata, T. (2015). Crystal structure of the CRISPR-Cas RNA silencing Cmr complex bound to a target analog. *Mol. Cell* 58, 418–430.
- Paez-Espino, D., Morovic, W., Sun, C.L., Thomas, B.C., Ueda, K., Stahl, B., Barrangou, R., and Banfield, J.F. (2013). Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nat. Commun.* 4, 1430–1437.
- Paez-Espino, D., Sharon, I., Morovic, W., Stahl, B., Thomas, B.C., Barrangou, R., and Banfield, J.F. (2015). CRISPR immunity drives rapid phage genome evolution in *Streptococcus thermophilus*. *MBio* 6, e00262–e00215.
- Patterson, A.G., Chang, J.T., Taylor, C., and Fineran, P.C. (2015). Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res.* 43, 6038–6048.
- Pawluk, A., Bondy-Denomy, J., Cheung, V.H.W., Maxwell, K.L., and Davidson, A.R. (2014). A new group of phage anti-CRISPR genes inhibits the type I-E CRISPR-Cas system of *Pseudomonas aeruginosa*. *mBio* 5, e00896-14.
- Plagens, A., Tjaden, B., Hagemann, A., Randau, L., and Hensel, R. (2012). Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon *Thermoproteus tenax*. *J. Bacteriol.* 194, 2491–2500.
- Plagens, A., Tripp, V., Daume, M., Sharma, K., Klingl, A., Hrle, A., Conti, E., Ur-laub, H., and Randau, L. (2014). In vitro assembly and activity of an archaeal CRISPR-Cas type I-A Cascade interference complex. *Nucleic Acids Res.* 42, 5125–5138.
- Plagens, A., Richter, H., Charpentier, E., and Randau, L. (2015). DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes. *FEMS Microbiol. Rev.* 39, 442–463.
- Polstein, L.R., and Gersbach, C.A. (2015). A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* 11, 198–200.
- Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151, 653–663.
- Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., and Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* 75, 1495–1512.
- Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520, 186–191.
- Rath, D., Amlinger, L., Hoekzema, M., Devulapally, P.R., and Lundgren, M. (2015). Efficient programmable gene silencing by Cascade. *Nucleic Acids Res.* 43, 237–246.
- Ratner, H.K., Sampson, T.R., and Weiss, D.S. (2015). I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. *Curr. Opin. Infect. Dis.* 28, 267–274.
- Redding, S., Sternberg, S.H., Marshall, M., Gibb, B., Bhat, P., Guegler, C.K., Wiedenheft, B., Doudna, J.A., and Greene, E.C. (2015). Surveillance and processing of foreign DNA by the *Escherichia coli* CRISPR-Cas system. *Cell* 163, 854–865.
- Richter, C., and Fineran, P.C. (2013). The subtype I-F CRISPR-Cas system influences pathogenicity island retention in *Pectobacterium atrosepticum* via crRNA generation and Csy complex formation. *Biochem. Soc. Trans.* 41, 1468–1474.
- Richter, C., Gristwood, T., Clulow, J.S., and Fineran, P.C. (2012). In vivo protein interactions and complex formation in the *Pectobacterium atrosepticum* subtype I-F CRISPR/Cas System. *PLoS ONE* 7, e49549.
- Richter, C., Dy, R.L., McKenzie, R.E., Watson, B.N., Taylor, C., Chang, J.T., McNeil, M.B., Staals, R.H., and Fineran, P.C. (2014). Priming in the Type I-F CRISPR-Cas system triggers trans-independent spacer acquisition, bi-directionally from the primed protospacer. *Nucleic Acids Res.* 42, 8516–8526.

- Rollie, C., Schneider, S., Brinkmann, A.S., Bolt, E.L., and White, M.F. (2015). Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *eLife* 4, 4.
- Rollins, M.F., Schuman, J.T., Paulus, K., Bukhari, H.S., and Wiedenheft, B. (2015). Mechanism of foreign DNA recognition by a CRISPR RNA-guided surveillance complex from *Pseudomonas aeruginosa*. *Nucleic Acids Res.* 43, 2216–2222.
- Rutkauskas, M., Sinkunas, T., Songailiene, I., Tikhomirova, M.S., Siksnys, V., and Seidel, R. (2015). Directional R-Loop Formation by the CRISPR-Cas Surveillance Complex Cascade Provides Efficient Off-Target Site Rejection. *Cell Rep.* 10, 1534–1543.
- Salvail-Lacoste, A., Di Tomasso, G., Piette, B.L., and Legault, P. (2013). Affinity purification of T7 RNA transcripts with homogeneous ends using ARIBo and CRISPR tags. *RNA* 19, 1003–1014.
- Samai, P., Pyenson, N., Jiang, W., Goldberg, G.W., Hatoum-Aslan, A., and Marraffini, L.A. (2015). Co-transcriptional DNA and RNA Cleavage during Type III CRISPR-Cas Immunity. *Cell* 161, 1164–1174.
- Sashital, D.G., Jinek, M., and Doudna, J.A. (2011). An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat. Struct. Mol. Biol.* 18, 680–687.
- Sashital, D.G., Wiedenheft, B., and Doudna, J.A. (2012). Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* 46, 606–615.
- Savitskaya, E., Semenova, E., Dedkov, V., Metlitskaya, A., and Severinov, K. (2013). High-throughput analysis of type I-E CRISPR/Cas spacer acquisition in *E. coli*. *RNA Biol.* 10, 716–725.
- Schunder, E., Rydzewski, K., Grunow, R., and Heuner, K. (2013). First indication for a functional CRISPR/Cas system in *Francisella tularensis*. *Int. J. Med. Microbiol.* 303, 51–60.
- Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* 108, 10098–10103.
- Servick, K. (2015). SCIENCE POLICY. U.S. to review agricultural biotech regulations. *Science* 349, 131.
- Shao, Y., and Li, H. (2013). Recognition and cleavage of a nonstructured CRISPR RNA by its processing endoribonuclease Cas6. *Structure* 21, 385–393.
- Shmakov, S., Savitskaya, E., Semenova, E., Logacheva, M.D., Datsenko, K.A., and Severinov, K. (2014). Pervasive generation of oppositely oriented spacers during CRISPR adaptation. *Nucleic Acids Res.* 42, 5907–5916.
- Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., et al. (2015). Discovery and functional characterization of diverse Class 2 CRISPR-Cas systems. *Mol. Cell* 60, 385–397.
- Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 30, 1335–1342.
- Sinkunas, T., Gasiunas, G., Waghmare, S.P., Dickman, M.J., Barrangou, R., Horvath, P., and Siksnys, V. (2013). In vitro reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J.* 32, 385–394.
- Sokolowski, R.D., Graham, S., and White, M.F. (2014). Cas6 specificity and CRISPR RNA loading in a complex CRISPR-Cas system. *Nucleic Acids Res.* 42, 6532–6541.
- Staals, R.H.J., Zhu, Y., Taylor, D.W., Kornfeld, J.E., Sharma, K., Barendregt, A., Koehorst, J.J., Vlot, M., Neupane, N., Varossieau, K., et al. (2014). RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol. Cell* 56, 518–530.
- Sternberg, S.H., and Doudna, J.A. (2015). Expanding the Biologist's Toolkit with CRISPR-Cas9. *Mol. Cell* 58, 568–574.
- Sternberg, S.H., Haurwitz, R.E., and Doudna, J.A. (2012). Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. *RNA* 18, 661–672.
- Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507, 62–67.
- Sternberg, S.H., LaFrance, B., Kaplan, M., and Doudna, J.A. (2015). Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature* 527, 110–113.
- Swarts, D.C., Mosterd, C., van Passel, M.W., and Brouns, S.J. (2012). CRISPR interference directs strand specific spacer acquisition. *PLoS ONE* 7, e35888.
- Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. USA* 111, 9798–9803.
- Tamulaitis, G., Kazlauskienė, M., Manakova, E., Venclovas, Č., Nwokeoji, A.O., Dickman, M.J., Horvath, P., and Siksnys, V. (2014). Programmable RNA shredding by the type III-A CRISPR-Cas system of *Streptococcus thermophilus*. *Mol. Cell* 56, 506–517.
- Taylor, D.W., Zhu, Y., Staals, R.H.J., Kornfeld, J.E., Shinkai, A., van der Oost, J., Nogales, E., and Doudna, J.A. (2015). Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* 348, 581–585.
- Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576.
- Tsui, T.K.M., and Li, H. (2015). Structure Principles of CRISPR-Cas Surveillance and Effector Complexes. *Annu. Rev. Biophys.* 44, 229–255.
- van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M., and Brouns, S.J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* 34, 401–407.
- van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 12, 479–492.
- van Erp, P.B.G., Jackson, R.N., Carter, J., Golden, S.M., Bailey, S., and Wiedenheft, B. (2015). Mechanism of CRISPR-RNA guided recognition of DNA targets in *Escherichia coli*. *Nucleic Acids Res.* 43, 8381–8391.
- Vestergaard, G., Garrett, R.A., and Shah, S.A. (2014). CRISPR adaptive immune systems of Archaea. *RNA Biol.* 11, 156–167.
- Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M., and Wang, Y. (2015). Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems. *Cell* 163, 840–853.
- Wei, Y., Chesne, M.T., Terns, R.M., and Terns, M.P. (2015a). Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in *Streptococcus thermophilus*. *Nucleic Acids Res.* 43, 1749–1758.
- Wei, Y., Terns, R.M., and Terns, M.P. (2015b). Cas9 function and host genome sampling in Type II-A CRISPR-Cas adaptation. *Genes Dev.* 29, 356–361.
- Westra, E.R., van Erp, P.B.G., Künne, T., Wong, S.P., Staals, R.H.J., Seegers, C.L.C., Bollen, S., Jore, M.M., Semenova, E., Severinov, K., et al. (2012). CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* 46, 595–605.
- Westra, E.R., Buckling, A., and Fineran, P.C. (2014). CRISPR-Cas systems: beyond adaptive immunity. *Nat. Rev. Microbiol.* 12, 317–326.
- Wiedenheft, B., Lander, G.C., Zhou, K., Jore, M.M., Brouns, S.J., van der Oost, J., Doudna, J.A., and Nogales, E. (2011a). Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* 477, 486–489.
- Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J., and Doudna, J.A. (2011b). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* 108, 10092–10097.
- Wright, A.V., Sternberg, S.H., Taylor, D.W., Stahl, B.T., Bardales, J.A., Kornfeld, J.E., and Doudna, J.A. (2015). Rational design of a split-Cas9 enzyme complex. *Proc. Natl. Acad. Sci. USA* 112, 2984–2989.

- Yosef, I., Goren, M.G., and Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* *40*, 5569–5576.
- Yosef, I., Shitrit, D., Goren, M.G., Burstein, D., Pupko, T., and Qimron, U. (2013). DNA motifs determining the efficiency of adaptation into the *Escherichia coli* CRISPR array. *Proc. Natl. Acad. Sci. USA* *110*, 14396–14401.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015a). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* *163*, 759–771.
- Zetsche, B., Volz, S.E., and Zhang, F. (2015b). A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.* *33*, 139–142.
- Zhang, Y., Heidrich, N., Ampattu, B.J., Gunderson, C.W., Seifert, H.S., Schoen, C., Vogel, J., and Sontheimer, E.J. (2013). Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell* *50*, 488–503.
- Zhang, Y., Rajan, R., Seifert, H.S., Mondragón, A., and Sontheimer, E.J. (2015). DNase H Activity of *Neisseria meningitidis* Cas9. *Mol. Cell* *60*, 242–255.
- Zhao, H., Sheng, G., Wang, J., Wang, M., Bunkoczi, G., Gong, W., Wei, Z., and Wang, Y. (2014). Crystal structure of the RNA-guided immune surveillance Cascade complex in *Escherichia coli*. *Nature* *515*, 147–150.
- Zuris, J.A., Thompson, D.B., Shu, Y., Gullinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* *33*, 73–80.

# The Basis of Oncoimmunology

A. Karolina Palucka<sup>1,2,\*</sup> and Lisa M. Coussens<sup>3,\*</sup>

<sup>1</sup>The Jackson Laboratory for Genomic Medicine, 10 Discovery Drive, Farmington, CT 06032, USA

<sup>2</sup>Baylor Institute for Immunology Research, Dallas, TX 75204, USA

<sup>3</sup>Department of Cell, Developmental & Cancer Biology, Knight Cancer Institute, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Mail Code L215, Room 5508, Richard Jones Hall, Portland, OR 97239-3098, USA

\*Correspondence: karolina.palucka@jax.org (A.K.P.), cousseni@ohsu.edu (L.M.C.)

<http://dx.doi.org/10.1016/j.cell.2016.01.049>

**Cancer heterogeneity, a hallmark enabling clonal survival and therapy resistance, is shaped by active immune responses. Antigen-specific T cells can control cancer, as revealed clinically by immunotherapeutics such as adoptive T-cell transfer and checkpoint blockade. The host immune system is thus a powerful tool that, if better harnessed, could significantly enhance the efficacy of cytotoxic therapy and improve outcomes for cancer sufferers. To realize this vision, however, a number of research frontiers must be tackled. These include developing strategies for neutralizing tumor-promoting inflammation, broadening T-cell repertoires (via vaccination), and elucidating the mechanisms by which immune cells organize tumor microenvironments to regulate T-cell activity. Such efforts will pave the way for identifying new targets for combination therapies that overcome resistance to current treatments and promote long-term cancer control.**

## Introduction

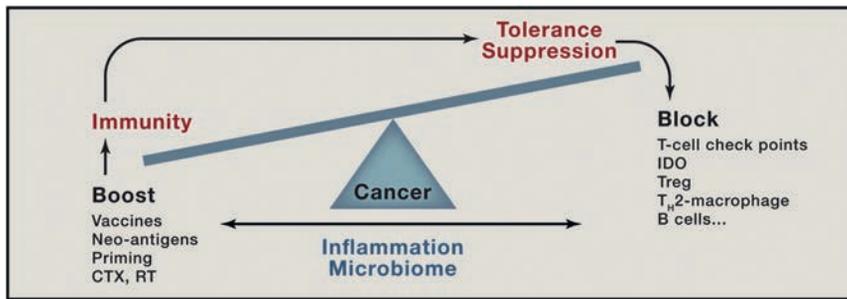
Cancer is an insidious disease traditionally classified by cell and tissue type of origin. Cancer has historically been treated according to a “one size fits all” approach based on broad pathologic criteria and involving various regimens of cytotoxic therapy. With the advent of modern sequencing methodologies, however, we now appreciate that significant genomic, transcriptomic, and epigenetic heterogeneity exists within individual tumor types; this recognition has enabled subclassification of tumors of common origin. This, in turn, has led to improved outcomes for some cancer types, as response rates to targeted and cytotoxic therapies increase when patients are stratified based on the molecular characteristics of their tumors. Examples include imatinib in chronic myelogenous leukemia (Druker et al., 2006), HER2-targeted therapies for HER2-positive breast cancer (Shepard et al., 1991), and estrogen antagonists for estrogen-receptor-positive breast cancers (Heiser et al., 2012). These molecular advances helped to usher in a new era of precision medicine that is reshaping clinical treatment across the cancer spectrum. However, there remain significant fractions of patients that do not respond to “designer” therapies even when their tumors are classified based on molecular and pathologic criteria. Additional tumor or systemic characteristic(s) are thus unaccounted for that not only impact neoplastic growth and dissemination, but also impact response to therapy.

Recent seminal *in vivo* studies revealed that neoplastic cells rely on the diversity of normal resident and recruited accessory cells to support their evolution (Hanahan and Coussens, 2012). Accessory cells are now recognized as “neoplastic cell-extrinsic hallmarks of cancer” and include those forming the hematogenous and lymphatic vasculature, tissue-specific mesenchymal support cells, and myeloid and lymphoid-lineage immune cells. Accessory cells integrate with the dynamic soluble and insoluble matrices

constituting the “tumor stroma”; collectively, they fuel neoplastic evolution (Hanahan and Coussens, 2012). In other words, reciprocal interactions between accessory cells, their mediators, structural components of the extracellular matrix (ECM), and genetically altered neoplastic cells regulate all aspects of tumorigenicity. These realizations fueled the development of anti-cancer agents targeting the vasculature (Kerbel, 2011) and, more recently, propelled clinical investigations into the efficacy of immune therapeutic approaches that neutralize tumor-promoting chronic inflammation and/or embolden or unleash cytotoxic activities of antigen-specific T cells (Coussens et al., 2013; Pardoll, 2012).

Indeed, cancer is visible to the immune system, *i.e.*, immunogenic, during early neoplasia. Classic studies from Schreiber and colleagues in mice with carcinogen-initiated sarcomas revealed that the immune system could recognize and reject cancerous cells (Dunn et al., 2004). Elimination can be explained by cytotoxic antigen-specific T cells responding to relatively high mutational burdens induced by carcinogens and thus providing neo-antigens for T-cell priming; these findings established the principles of elimination, equilibrium, and eventually escape when neoplastic cells become invisible to the immune system (Dunn et al., 2004). Neoplastic cells in part escape when tumor arises out of chronically inflamed tissues—there, chronic infiltration of tissue by leukocytes (*e.g.*, type 2 cytokine-activated myeloid cells and immune-suppressive B, T, and myeloid subsets) subvert T-cell-directed elimination and thus aid tissue-based programs, *e.g.*, angiogenesis, lymphangiogenesis, matrix remodeling, etc., supporting neoplastic progression (Coussens et al., 2013).

Mounting observations in humans support the concept that cancer initiation and progression are significantly impacted by altered or misled immune responses (Figure 1). Individuals suffering from chronic inflammatory conditions are at increased



**Figure 1. The Makings of Tumor Immunity**

The communication between cancer and the immune system is a dynamic process, reminiscent of a balance. When immunity to cancer is “up” and the suppressive processes are “down,” cancer is under control. However, a strong anti-tumor immune response will trigger largely physiological processes designed to dampen effector T cells to prevent tissue damage and maintain tissue homeostasis. Given that the immunity might have evolved mainly to maintain self, to establish coexistence with environment, and to occasionally protect self from external threats, the suppression prevails. Multiple pathways of suppression are at play in tumor microenvironments, including cells

such as  $T_H2$ -polarized macrophages, immature and suppressive monocytes, regulatory B cells, and regulatory T cells, as well as molecules such as checkpoints that control T-cell differentiation (for example, CTLA-4 and IDO) and effector function (such as PD-1). Pharmacological blockade of these inhibitory pathways can tip the balance toward anti-cancer effector T cells. The latter ones can be primed or boosted by antigen-presenting cells (DCs) and/or by co-stimulatory signals (for example, CD137 ligands). Recent studies demonstrate that thymus-independent neo-antigens generated in adult life by somatic mutation or post-translational regulation (for example, phosphorylation) might be more immunogenic (or perhaps linked with less suppression) than shared tumor antigens. Neo-antigens can occur as random results of somatic mutation, as well as a by-product of anticancer treatments, e.g., chemotherapy (CTX) or radiation therapy (RT), or by targeting epigenetic control mechanisms or drugs intervening with DNA repair pathways. They can be presented to T cells in exogenous vaccines, as well as endogenously via DCs that captured dying neoplastic cells. When T cells specific to defined antigens kill neoplastic cells, such a process can enable generation of responses to other antigens, so called epitope spreading. A critical factor in the balance between immunogenicity and suppression is inflammation (which, in turn, is impacted by the microbiome); indeed, the type of inflammation (tumor-destructing  $T_H1$  or tumor-promoting  $T_H2$  and  $T_H17$ ) should become a part of TNM grading, along with pathology, microbiome phenotype, and immune infiltrate assessment.

risk for developing cancer (Thun et al., 2004). Incidence of viral (DNA tumor virus) and carcinogen-associated cancers is increased in immune-compromised individuals, even as the relative risk of cancer types lacking viral or carcinogen etiology is diminished (reviewed in de Visser et al., 2006). Age-related immunosenescence likely plays a role in increased incidence of malignancy in aged individuals (Campisi et al., 2011). The advent of some biologic therapies impacting how tissues activate and resolve inflammation, e.g., tumor necrosis factor (TNF) blockade (Bongartz et al., 2006), also skews cancer incidence metrics. However, the role(s) that immune pathways play in driving malignancy remains to be clarified. How does the immune system recognize tissue-specific mediators triggering and maintaining chronic inflammatory responses? What oncogenic events and altered metabolic states lead to the generation of neo-antigens that in turn induce T-cell responses? What physiological mechanisms regulate immune homeostasis such that (acute) inflammation can be resolved as rapidly as it is activated (a critical control program to thwart autoimmunity)? What is the role of the host microbiota in regulating systemic immune responses to neoplasia? How do neoplastic cells survive immune attack by T cells? These questions are in need of answering to effectively move cancer research and cancer medicine forward.

A common feature of all cancers, regardless of origin, is prominent presence of diverse assemblages of immune cells (Coussens et al., 2013). The consequences of such infiltrates on the fate of cancerous cells are diverse (Figure 2). For example, under continual immune pressure, i.e., antigen presentation to T cells, neoplastic cells become “immune-edited” to escape immune surveillance (Dunn et al., 2004) and instead co-opt immune cells to favor their sustained proliferation (Balkwill et al., 2005). Nonetheless, recent studies demonstrate that the presence of lymphoid aggregates is linked with improved responses to cancer therapies—for example, standard cytotoxic therapies, vaccine-based treatments, or immune checkpoint blockade (Topalian et al., 2015). Such “hot” tumors are thus

more amenable to control than “cold” tumors, i.e., tumors with diminished T-cell infiltrates, thus driving modern cancer medicine to investigate how to reprogram the tumor microenvironment (TME) to attract the right type of immune infiltrate. This topic, along with other open questions in the field of oncoimmunology, are discussed here.

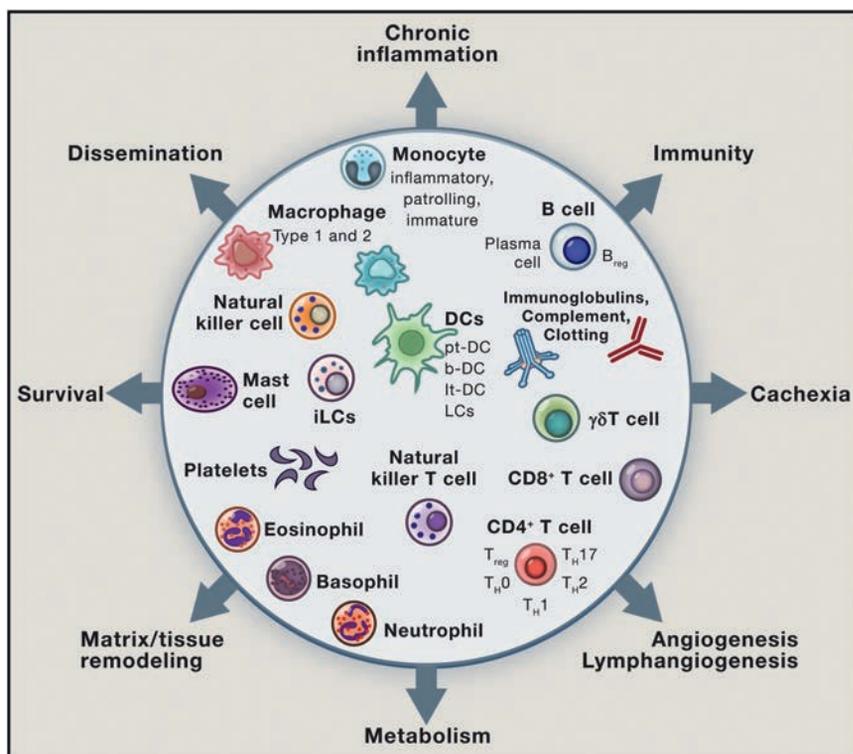
### The Makings of the Immune Response to Cancer

Tumors are organized tissues with numerous reciprocal local and systemic connections with immune cell populations of both the myeloid and lymphoid lineages. Here, we summarize the key myeloid and lymphoid populations regulating the immune response to cancer and how the fundamental physiological processes that they govern are harnessed for neoplastic progression and tumor formation.

#### The Myeloid Compartment

Myeloid cells have multiple homeostatic functions that are co-opted by evolving neoplasms; these can be roughly summarized as: (1) antigen capture for degradation (macrophages) or presentation (dendritic cells [DCs]); (2) tissue repair (macrophages), and (3) effector functions (mast cells, monocytes, and granulocytes). Neoplastic cells can alter the steady-state activity of all myeloid cells present in the TME, including tissue-resident and blood-derived cells, by secreting factors such as interleukin (IL)-6 or granulocyte-macrophage colony-stimulating factor (GM-CSF), that increase recruitment and proliferation of immature myeloid cells atypical under physiological conditions (Gabrilovich et al., 2012).

An important feature of myeloid cells is their functional plasticity in response to environmental signals. This property can dictate such opposite outcomes as antigen degradation or antigen presentation when macrophages acquire DC capabilities (Banchereau et al., 2000), tissue repair rather than inflammation when macrophages are polarized toward type 2 states, and protective or non-protective T-cell immunity when programmed by cancer-derived factors (Balkwill et al., 2005). Thus, plasticity



**Figure 2. Immune-Mediated Landscape**

The yin and yang implications of tumor-immune system communications form the basis for disease pathophysiology and, at the same time, targets for therapeutic intervention. The disease landscape emerging from these multi-factorial interactions is orchestrated by the three compartments, i.e., the cancer, the immune system, and the host. The outputs are numerous and dramatically opposite, as well as both local and systemic, and include: immunity that might control cancer; chronic inflammation that can be linked with tissue remodeling processes and metabolic changes that support neoplastic cell survival and primary tumor development; angiogenesis and lymphangiogenesis that can also support metastatic dissemination; as well as systemic consequences for the host including cachexia. Clearly, therapy going forward will require a well-timed and orchestrated combination of therapies, targeting multiple modes of communication and effect, to combat this multi-factorial disease, taking into account the patient's steady-state commensal bacteria complexity and load and how that is impacted by therapy.

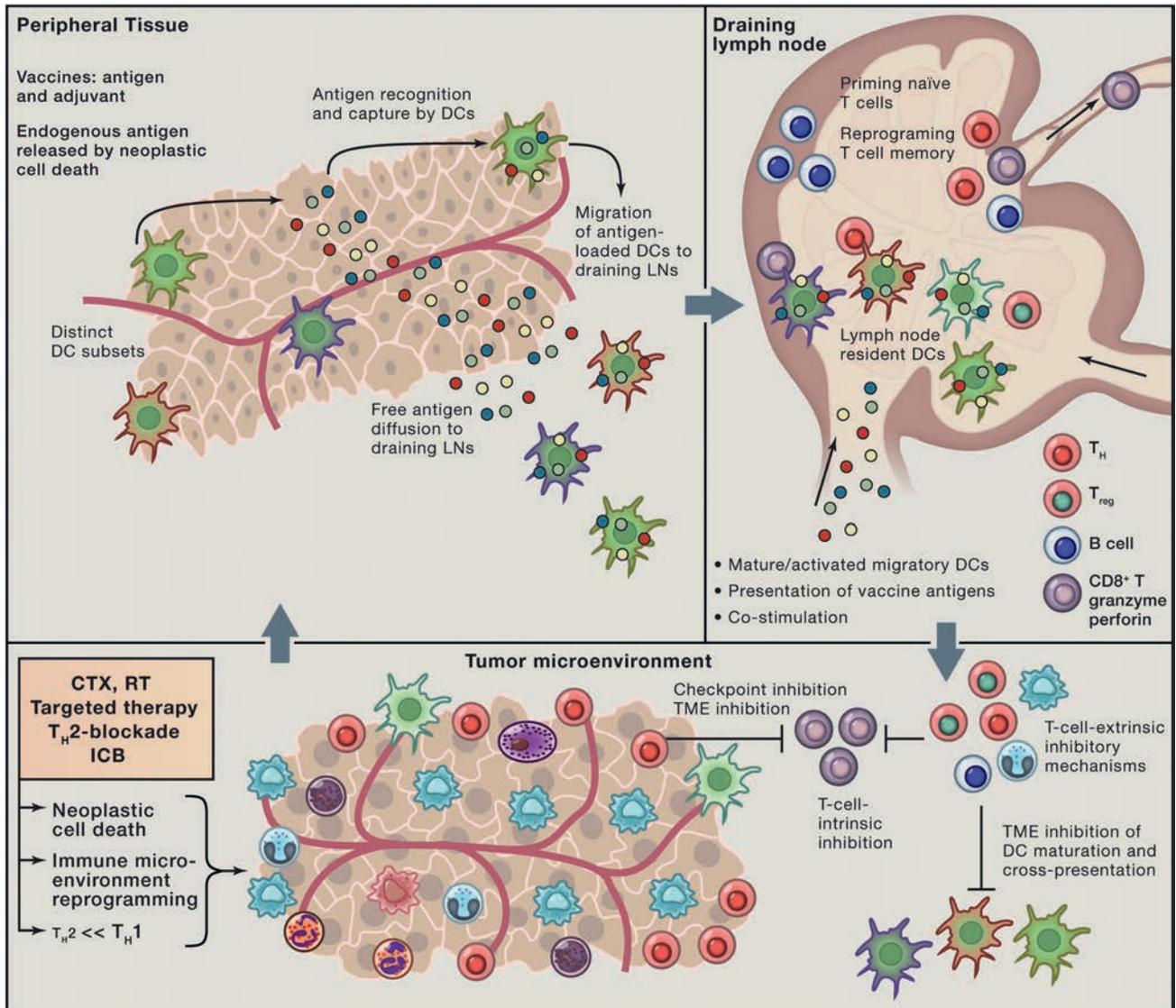
DCs express numerous pattern recognition receptors, including lectins, Toll-like receptors (TLRs), NOD-like receptors (NLRs), and helicases, through which they can sense microbes and tissue damage

and communication within the myeloid compartment and between myeloid and other immune cells and stromal components is critical for tumor formation.

**Cancer Antigen Presentation and Dendritic Cells.** Cancer antigens are presented to T cells either at tumor sites or in draining lymph nodes by DCs (Figure 3). Cancer antigens, soluble and cell borne, are transported to lymph nodes via lymphatic vessels. Soluble antigen is captured by lymph-node-resident DCs, while tissue-resident DCs capture antigen at tumor sites; either population can present antigen locally or migrate through lymphatic vessels to present in lymph nodes (Steinman, 2011). DCs display protein antigens in the context of classical major histocompatibility (MHC) class I and MHC class II molecules or lipid antigens in the context of non-classical CD1 molecules that allow selection of rare antigen-specific T lymphocytes, including CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, and NK T cells. Compared with other antigen-presenting cells (APCs), DCs are extremely efficient in their ability to induce antigen-specific T-cell responses, justifying their name “professional APCs” (Lanzavecchia and Sallusto, 2001). Naive CD8<sup>+</sup> T cells differentiate into cytotoxic T lymphocytes (CTLs) in lymphoid organs upon encounter with DCs, presenting tumor-derived peptides in the context of co-stimulation through CD80, CD70, and 4-1BB, as well as through DC-derived cytokines such as IL-12, type I interferon, and IL-15 (Steinman, 2012). The priming of new T-cell repertoires during tumorigenesis may be critical for clinical success of therapeutic agents aiming to unleash antigen-specific CTL activities. Naive CD4<sup>+</sup> T cells can give rise to helper cells with distinct cytokine profiles or to FoxP3<sup>+</sup> regulatory T cells (Treg), whose role is to dampen CTL activity and avoid autoimmune responses (Zhu and Paul, 2008).

(cancer) such as increased pericellular nucleic acids (Pulendran, 2015). If DCs do not receive maturation signals, such as when exposed to high levels of IL-10 (Ruffell et al., 2014), they remain immature and antigen presentation instead leads to T-cell suppression. DC plasticity in response to extrinsic signals, together with the existence of discrete subsets with unique functions, empowers DCs as key initiators and regulators of the immune response (Pulendran, 2015). We will illustrate this point briefly; mouse and human DC subset biology was recently reviewed elsewhere (Merad et al., 2013).

The diversity of human DC subsets was revealed by studies of blood and skin DCs. Three main cell-surface markers distinguished human-blood-circulating DC subsets: CD303 (BDCA-2) on plasmacytoid DCs (pDCs), CD1c (or BDCA-1) expressed on the majority of circulating DCs, and CD141/BDCA-3 expressed on a small fraction (Merad et al., 2013). Human CD141<sup>+</sup>CD1c<sup>-</sup> DCs uniquely express TLR3, produce IL-12, and efficiently cross-prime CD8<sup>+</sup> T cells when activated with poly I:C (Joffre et al., 2012). However, other human DCs, such as epidermal Langerhans cells and CD1c<sup>+</sup> DCs, also cross-present antigens to CD8<sup>+</sup> T cells. Indeed, our studies have unraveled the basic principles by which human DC subsets differentially regulate CD8<sup>+</sup> T cells (Klechevsky et al., 2008). Thus, human Langerhans cells are highly efficient at priming cytotoxic CD8<sup>+</sup> T cells, while CD14<sup>+</sup> dermal DCs prime type 2 cytokine-secreting CD8<sup>+</sup> T cells (Klechevsky et al., 2008). Blood- and tissue-resident CD1c<sup>+</sup> DCs, but not CD141<sup>+</sup> DCs, exposed to live-attenuated influenza virus promote CD103 (αE integrin) expression on CD8<sup>+</sup> T cells and their accumulation in epithelia (Yu et al., 2013).



**Figure 3. The Priming of Cancer Immunity**

The cycle of anti-tumor immunity starts presumably with presentation of cancer antigens liberated in the process of cell turnover; this same pathway can be followed for vaccination, as illustrated herein. Antigens are sensed and captured either by tissue-resident DCs or by DCs in draining lymph nodes (LNs). DCs initiate an immune response by presenting captured antigens, in the form of peptide-major histocompatibility complex (MHC) molecule complexes, to naive (that is, antigen inexperienced) T cells in lymphoid tissues. When compared with other APCs, such as macrophages, DCs are extremely efficient and can elicit very low numbers of T cells to respond. Naive CD8<sup>+</sup> T cells differentiate into CTLs in lymphoid organs upon encounter with DCs presenting tumor-derived peptides in the context of co-stimulation through CD8, CD70, and 4-1BB, as well as DC-derived cytokines such as IL-12 and IL-15. Naive CD4<sup>+</sup> T cells can give rise to helper cells (e.g., T<sub>H</sub>) with distinct cytokine profiles or to regulatory T cells (T<sub>reg</sub>) whose role is to dampen the immune response. T cells migrate through blood and lymphatics. Upon arrival in tumor beds, CD8<sup>+</sup> T cells must confront numerous barriers including: (1) intrinsic regulators, for example, CD28-CTLA-4, PD1-PDL1, and ILTs, as well as extrinsic regulators cells such as Tregs, Bregs, or myeloid cells; (2) a corrupted TME with pro-tumor inflammation; (3) impaired cross-presentation due to TME-based DC inhibition; (4) antigen loss and immune evasion of tumor target; and (5) tissue-specific alterations such as fatty cells in breast cancer or desmofibrosis in pancreatic cancer stroma. Killing of tumor cells either via T cells or by standard therapy can lead to endogenous antigen release and DC activation, so called “endogenous vaccination,” thereby closing the cycle. Inevitable to this is the induction of tissue resistance mechanisms, for example, expression of PD-L1 on neoplastic or other immune cells, as the result of powerful effector immunity, including actions of IFN $\gamma$ . Thus, future immunotherapy approaches will be based on combinations of different therapeutics targeting distinct components of this cycle, for example, via intratumoral delivery of activating agents able to reprogram the function of infiltrating leukocytes.

**The Lymphoid Compartment**

The lymphoid compartment in tumors includes natural killer (NK) cells,  $\gamma\delta$  T cells, NK T cells, CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, and B cells. Their functional activity depends upon expression of restriction elements, including peptide-MHC complexes (pMHC;

for T cells), the MHC class I molecule (for NK cells), or surface proteins (for B-cell products, i.e., antibodies) that can be recognized in a specific manner. In addition, lymphoid cells can be induced to secrete different types of cytokines based on effector functions. For example, following an activating stimulus,

$T_H1$ -polarized  $CD4^+$  T cells secrete IL-2,  $TNF\alpha$ , and  $IFN\gamma$ ; in conjunction with cytotoxic  $CD8^+$  T cells, they promote macrophage cytotoxic activity (Stout and Bottomly, 1989) and can induce upregulation of antigen processing and expression of MHC I and II molecules in professional APCs (i.e., macrophages and DCs). In contrast, expression of IL-4, -5, -6, -10, and -13 by  $T_H2$ -polarized  $CD4^+$  T cells can induce T-cell anergy and loss of T-cell-mediated cytotoxicity, enhance humoral immunity, and regulate the tumor-promoting activities of macrophages (DeNardo et al., 2009).

$CD8^+$  T cells are considered the major anti-cancer effector cells, as they can give rise to CTLs that kill neoplastic cells presenting a specific pMHC complex (Appay et al., 2008). CTLs can be generated through either the priming of naive T cells or reprogramming of memory T cells. Naive  $CD8^+$  T cells differentiate into CTLs in lymphoid organs upon encounter with APCs presenting tumor-derived peptides in the context of appropriate co-stimulation and cytokine help. The ideal properties of anti-cancer  $CD8^+$  T cells include: high affinity for pMHC on tumor cells; high levels of cytotoxic mediators, e.g., granzymes A and B and perforin; expression of surface molecules, allowing trafficking into the tumor; and extended longevity and memory, thus enabling CTL generation upon antigen re-exposure (Appay et al., 2008).

Memory T cells have long been described as two circulating populations: (1) central memory T cells that migrate between the secondary lymphoid organs and are capable of mounting proliferative responses on pathogen re-encounter and (2) effector memory T cells that traffic between blood and extralymphoid compartments for peripheral immune surveillance (Mueller et al., 2012). Tissue-resident memory T cells are a third and phenotypically distinct category. Studies in mice and humans have revealed that this latter population can be superior to circulating central memory T cells at providing rapid long-term protection against re-infection (Sheridan and Lefrançois, 2011). Therefore, an active mechanism of peripheral T-cell retention likely exists not only to facilitate clearance of infected cells, but also to promote accumulation at sites having cleared an infectious virus.  $CD103/\beta7$  integrin endows peripheral  $CD8^+$  T cells with a unique capacity to access epithelial compartments. Expression of  $CD103$  on CTLs mediates adherence to E-cadherin and appears to be important in the final stages of neoplastic cell lysis and rejection (Le Floc'h et al., 2007). Indeed, for mucosal cancer vaccines, homing to and retention of  $CD8^+$  T cells in mucosa is critical for efficacy (Sandoval et al., 2013).

Upon arrival in tumor beds,  $CD8^+$  T cells must confront numerous barriers, including intrinsic checkpoint regulators, such as  $CD28$ - $CTLA-4$ ,  $PD1$ - $PD-L1$ , and immunoglobulin-like transcript receptors (ILTs) (Pardoll, 2012); extrinsic checkpoint regulators, such as Treg cells (Fehérvári and Sakaguchi, 2004) or myeloid cells (Gabrilovich et al., 2012); a corrupted TME with protumor inflammation (Coussens et al., 2013); antigen loss and immune evasion of tumor targets (Klebanoff et al., 2011); and tissue-specific alterations, such as fatty cells in breast cancer or desmofibrosis in pancreatic cancer stroma. Defining strategies for bypassing these obstacles and improving the clinical efficacy of T-cell therapies is the object of intense study.

An important concept recently proposed by Mellman and colleagues is the cancer-immunity cycle (Chen and Mellman, 2013). It becomes apparent that any effective immune response against cancer will generate resistance via physiological pathways that evolved to protect tissue homeostasis. Here, we discuss how this cycle is altered in cancer pathogenesis and how it can be harnessed therapeutically. Clearly, combination therapies that intervene at several distinct pathways within the cancer-immunity cycle are needed to achieve cancer control.

## Chronic Inflammation and Alterations of Leukocyte Compartments in Cancer

### Basic Principles

Unabated inflammation is a hallmark of cancer and is mediated by immune cells attracted to or residing at sites of neoplastic transformation (Balkwill et al., 2005). Indeed, immune cells are selectively recruited into early neoplastic tissues, likely in response to hard-wired pathways utilized by all tissues to resist/repair damage caused by bacterial, viral, or other pathogenic assaults. When successful, "initiated" pre-neoplastic cells are purged by the immune system (Dunn et al., 2004). When the immune system fails, neoplastic cells are retained in "damaged" TMEs and provide a survival advantage resulting from abundant bioavailable mediators liberated as a function of tissue remodeling (Hanahan and Coussens, 2012). Ensuing neoplastic progression requires sustained presence of select immune subtypes that, combined with ongoing host-derived programs (angiogenesis, matrix, and tissue remodeling, etc.), contribute to cancer progression (Hanahan and Coussens, 2012) (Figure 2).

The classic view that immune cells merely facilitate tumor rejection has been supplanted by a more complex view of leukocytes having both tumor-promoting and tumor-inhibiting properties (Coussens et al., 2013). This is best explained by the existence of (at least) two types of inflammation with opposing effects on tumors: chronic inflammation, which promotes neoplastic cell survival, angiogenesis, tissue remodeling, and metastasis, and acute inflammation that triggers neoplastic cell destruction. While chronic inflammation is often linked with the presence of  $T_H2$  responses, acute inflammation associated with cancer destruction is linked to  $T_H1$  responses.

As neoplastic cells escape elimination, some become less immunogenic by downregulating MHC molecules; however, most if not all also activate intrinsic gene-expression programs that are inherently T-cell suppressive and myelo-stimulatory, e.g.,  $T_H2$  responses. Cytokines implicated in these scenarios include transforming growth factor  $\beta$  ( $TGF\beta$ ); IL-4, -13, -8, and -10; thymic stromal lymphopoietin (TSLP); and indoleamine 2,3-dioxygenase (IDO) (Coussens et al., 2013). This enables recruitment of  $FoxP3^+CD4^+$  Treg cells,  $T_H2$ - $CD4^+$  T cells,  $T_H2$ -polarized macrophages and monocytes, and B regulatory cells (Bregs). In response to  $T_H2$ -mediated activation, myeloid cells commonly increase synthesis of angiogenic (e.g., VEGF), growth and/or survival (e.g., EGF,  $TNF\alpha$ ) factors that directly regulate epithelial cell proliferation, as well as tissue-remodeling enzymes (e.g., metallo-, cysteine, and serine proteases). These activities are remarkably pro-tumorigenic in that they nurture a TME favoring neoplastic cell survival and sustained proliferation (Balkwill et al., 2005). Simultaneous  $T_H2$  activation of

macrophages and monocytes also increases expression of molecules, e.g., inducible nitric oxide synthase or Arginase 1, that directly and indirectly suppress CD8<sup>+</sup> T-cell proliferation and cytokines such as IL-10 that inhibit DC maturation and antigen cross-presentation to T cells (Ruffell et al., 2014). Thus, T<sub>H</sub>2-type immune microenvironments are both tumor promoting and immune suppressive. Notably, in the colon, tumor-promoting immunity via IL-17 (T<sub>H</sub>17)-mediated activation of myeloid and lymphoid cells has been reported (Wang et al., 2009; Wu et al., 2009).

### **Tumor-Promoting Activities of the Myeloid Compartment**

Owing to their established role in wound healing, we investigated the ability of myeloid cells infiltrating early benign tissues to foster malignancy. In mice prone to squamous carcinogenesis, mast cells and macrophages activate pro-neoplastic angiogenic and tissue-remodeling programs (Coussens et al., 1999). In other studies of mice bearing mammary carcinomas, macrophages could regulate neoplastic cell dissemination and metastasis via EGF-mediated paracrine interactions with neoplastic epithelial cells (Lin et al., 2001). In human cancers, multiple studies have reported that the presence of macrophages in stroma correlates with aggressive disease and outcome (Komohara et al., 2014). Macrophages are recruited into tumors following activation of colony-stimulating factor 1 receptor (CSF1R) by either CSF1 or IL-34, two high-affinity ligands for CSF1R; the chemokine CCL2 may also facilitate macrophage recruitment (Qian et al., 2011). A CSF1-response gene-expression signature has been identified in 17%–25% of breast cancers associated with decreased estrogen receptor and progesterone receptor expression (Beck et al., 2009); serum concentrations of CSF-1 correlate positively with breast tumor size and predict poor survival (Aharinejad et al., 2013). In addition, in two independent breast cancer cohorts, intratumoral macrophage presence was correlated with potentially prognostic tumor features (high-grade, hormone receptor negativity; basal-like subtype; and increased risk of death) (Komohara et al., 2014). Macrophages therefore serve as promising targets for novel therapeutic interventions, particularly for patients with high-risk disease. Conversely, favorable prognosis has been associated with some tumor types exhibiting increased macrophage infiltration, e.g., non-small-cell lung cancer, prostate, colorectal, and gastric cancers (Komohara et al., 2014). Whether these distinctions reflect true differences in macrophage biology and function or arise due to discordant detection methodologies is unclear.

Neutrophils, on the other hand, are typically less abundant than macrophages in solid tumors, but their presence correlates with reduced survival in head and neck and breast cancers, and similar to macrophages, neutrophils develop polarized phenotypes that either favor or restrict tumor progression (Fridlender and Albelda, 2012). Recent studies identified granule products that suppress T-cell function (Sippel et al., 2011). Neutrophil expansion in mammary carcinomas of mice bearing mutant p53 alleles is driven by T-cell-derived IL-17; this results in systemic granulocyte colony-stimulating factor (G-CSF)-dependent expansion and polarization toward a T-cell-suppressive phenotype that facilitates metastatic dissemination and colonization (Coffelt et al., 2015). In contrast, neutrophils create a tumor-restrictive microen-

vironment in the lung that resists neoplastic progression and metastatic dissemination (Fridlender and Albelda, 2012).

Eosinophils, like other myeloid lineage cells, can exert cytotoxic immune-effector activities. Tumor-associated tissue eosinophilia (TATE) is associated with improved prognosis for a number of malignancies, including gastrointestinal, bladder, and prostate cancers; in contrast, TATE is associated with poor outcome in Hodgkin's lymphoma, cervical carcinoma, and oral squamous cell carcinoma (Davis and Rothenberg, 2014). Eosinophils have been associated with degranulation and release of cytotoxic proteins that mediate tumor rejection; recent results also reveal their role in normalizing the vasculature to improve CD8<sup>+</sup> T-cell trafficking associated with tumor regression (Carretero et al., 2015).

Monocytes, once in tissues, can differentiate into macrophages and DCs. Two circulating monocyte populations have been identified: classical inflammatory monocytes that are CCR2<sup>HIGH</sup> and non-classical patrolling monocytes that are CX3CR1<sup>HIGH</sup> (Geissmann et al., 2003). Recruitment of inflammatory monocytes into tissues is normally guided by the CCR2-CCL2 axis in response to parasitic or bacterial infections; in tumors, when CCR2<sup>HIGH</sup> monocytes are recruited, they can promote neoplastic cell survival and extravasation through VEGF and CSF1 production (Qian et al., 2011). CCR2<sup>HIGH</sup> monocytes promote survival of metastatic cells through a CCL3-dependent mechanism (Kitamura et al., 2015). CX3CR1<sup>HIGH</sup> monocytes instead patrol capillaries in response to the CXCR3-CX3CL1 axis; in these locales, they are positioned to scavenge particles and debris and thus are more likely to be found in wounds when inflammation is resolving. At sites of metastasis, CX3CR1<sup>HIGH</sup> monocytes recruit NK cells that, in turn, kill metastatic cells, thereby providing a potent survival advantage (Hanna et al., 2015). In pancreatic adenocarcinomas, activation of the Ras oncogene leads to increased expression of GM-CSF and recruitment of immature monocytes that subsequently suppress CD8<sup>+</sup> T-cell proliferation to enhance tumor progression (Pylayeva-Gupta et al., 2012), analogous to other tumor systems (Gabrilovich et al., 2012).

Mast cells, present in all vascularized tissues, respond to diverse stimuli by either secreting or releasing (via degranulation) biologically active compounds, e.g., proteolytic enzymes, heparin, histamine, prostaglandins, cytokines, and chemokines. Mast cells are key for maintaining tissue homeostasis and are best known for their effector functions following IgE-stimulated allergic responses and anaphylaxis (Metz et al., 2007). Mast cells have been implicated in the vascularization of a multitude of solid human tumor types, likely owing to their proteolytic products and high VEGF expression following activation (Coussens et al., 1999; Marichal et al., 2013) following CCL2-mediated recruitment where their bioactive mediators promote neoplastic progression.

### **Tissue Specificity of Myeloid Programming**

While it is conceptually unclear how some myeloid cells adopt a T<sub>H</sub>2 or protumorigenic state to support neoplastic progression, some clues have emerged in recent genetic studies. Several groups revealed that lymphocytes drive initial myeloid cell programming to foster chronic inflammation in a tissue-specific manner. For example, during mammary branching morphogenesis and ductal carcinogenesis, cytokines derived from

$T_H2$ -CD4<sup>+</sup> T cells, e.g., IL-4 and -13, activate macrophages and monocytes infiltrating mammary tissue (DeNardo et al., 2009; Plaks et al., 2015). In neoplastic scenarios, signaling downstream of IL-4 receptors on monocytes and macrophages triggers protumorigenic  $T_H2$  gene-expression programs that activate tissue-remodeling cascades via expression and activation of cathepsin proteases and immune-suppressive programs via upregulation of IL-10 and immune-checkpoint molecules (DeNardo et al., 2009; Gocheva et al., 2010; Mitchem et al., 2013; Ruffell et al., 2014). Mast cells and macrophages (as well as other myeloid cell types) are  $T_H2$  programmed in early squamous and pancreatic carcinomas by a diversity of pathways, which also include activation of immunoglobulin receptors (Fc $\gamma$ Rs) by immune complexes (ICs) (Affara et al., 2014; Andreu et al., 2010). ICs are composed of antigen-specific antibodies and complement proteins that variably activate FcR and complement receptors depending on composition of IC and status of the myeloid cell being activated (Karsten and Köhl, 2012). While these humoral immune-regulated paracrine programs were known to shape outcomes in chronic inflammatory diseases, recognition of their significance in solid tumors was paradigm shifting and highlighted the significance of hard-wired tissue-specific programs shaping host response to disease. These data illustrate the diversity of pathways utilized by innate immune cells to propel cancer by directly enhancing tissue-based programs favoring survival of neoplastic cells, in concert with direct and indirect activation of programs to extinguish cytotoxic immune responses aiding immune escape (Figure 2).

### **$T_H2$ -Based Targets for Anti-cancer Therapy**

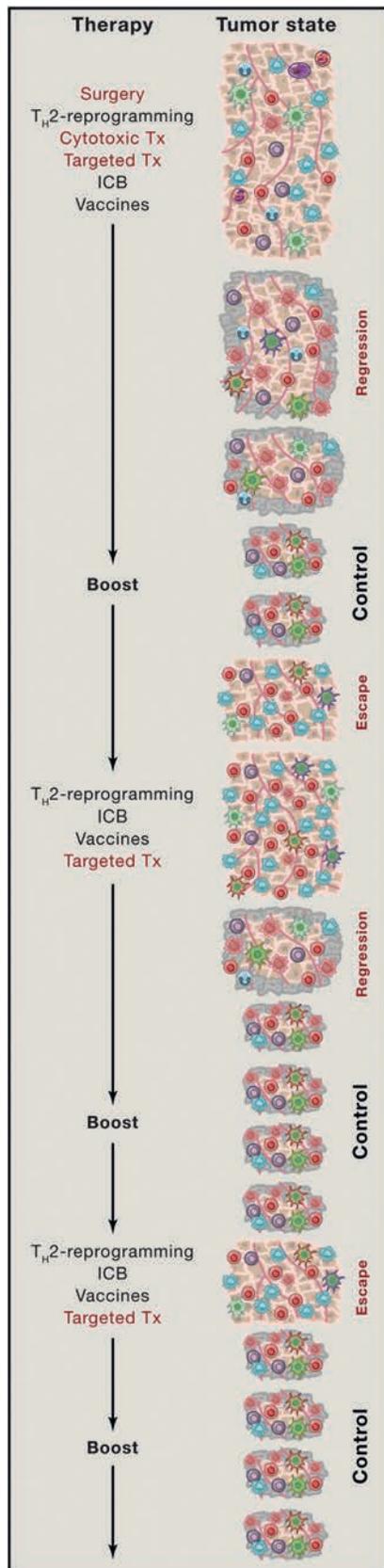
The collective evidence not only supports a protumorigenic role for chronic inflammation in cancer, but also indicates that inflammation is malleable, akin to the healing of acute wounds during which immune cells toggle between  $T_H1$  and  $T_H2$  states. Thus, the hypothesis that  $T_H2$ -driven myeloid cells could be re-programmed, or at least neutralized, to reduce the presence or immunosuppressive status of macrophages, trigger anti-tumor immunity, and/or suppress tumor growth has been tested in several tissue-specific cancer models. We and others have evaluated CSF1-neutralizing monoclonal antibodies ( $\alpha$ CSF1 mAb) and small-molecule CSF1R inhibitors for their ability to suppress macrophage survival and/or presence in tumors, in combination with chemotherapy (CTX) or radiation therapy (RT) (DeNardo et al., 2011; Ruffell et al., 2014; Shiao et al., 2015). These studies reveal increased chemo- and radiation sensitivity associated with anti-tumor immune responses directed by CD8<sup>+</sup> T-cell infiltration of tumors, culminating in reduced primary tumor growth and metastasis with increased survival. Other preclinical studies revealed that CSF1/CSF1R blockade, as monotherapy or combined with CTX/RT, improved outcomes for glioma, prostate, and pancreatic adenocarcinoma, as well as melanoma, whereas CSF1R blockade improved antitumor efficacy of immune checkpoint blockade and adoptive T-cell therapy (reviewed in Ruffell and Coussens, 2015). Importantly, administration of RG7155, a CSF1R-blocking mAb, in patients with diffuse-type giant cell tumors, reduced CSF1R<sup>+</sup>CD163<sup>+</sup> macrophage levels; this translated into objective clinical responses (Ries et al., 2014). Treatment of tenosynovial giant-cell tumors with a small-molecule

inhibitor of CSF1R kinase increased progression-free survival and improved outcomes as a monotherapy (Tap et al., 2015). The macrophage presence in tumors has also been therapeutically manipulated by targeting the macrophage signaling protein acting through its transmembrane receptor kinase RON, wherein activation of RON in macrophages favors conversion of micro-metastatic lesions to overt metastases by suppressing antitumor immune responses. Functional RON blockade in preclinical models potentiates tumor-specific CD8<sup>+</sup> T-cell responses, indicating that RON inhibitors may also improve outcomes for cancer patients (Eyob et al., 2013).

Bruton's tyrosine kinase (BTK) is an attractive target, as BTK is activated downstream of the B-cell receptor and Fc $\gamma$ R and PI3K $\gamma$  in some myeloid subsets (Smith et al., 2001). In vitro, neoplastic cell challenge via co-culture with splenic cells from B-cell-deficient versus B-cell-proficient mice revealed that IFN $\gamma$  release from CD8<sup>+</sup> and NK cells is increased when B cells were absent, whereas presence of B cells or B-cell-derived IL-10 was associated with reduced IFN $\gamma$  (Inoue et al., 2006). Though these in vitro studies indicate that B cells can direct T-cell responses, the role of myeloid cells as mediators of these responses is now clear and indicates that therapies targeting common pathways in B cells and/or myeloid cells, such as SYK, BTK, PI3K $\gamma$ , may be efficacious in solid tumors, analogous to efficacy observed for BTK and PI3K $\delta$  inhibitors in B-cell malignancies (Hendriks et al., 2014). This concept was recently validated preclinically, whereby BTK inhibition enhanced survival of mice bearing pancreatic adenocarcinoma (Gunderson et al., 2016; Massó-Vallés et al., 2015), neuroendocrine cancers (Soucek et al., 2011), and other subcutaneous tumors (Sagiv-Barfi et al., 2015) in which a common feature was reduced inflammation and inflammatory desmoplasia with evidence of macrophage repolarization.

If these preclinical findings are any indication, immune therapies targeting macrophages and/or other protumorigenic immune cells could alter the human tumor immune microenvironment in a way that fosters cytotoxic properties of CD8<sup>+</sup> T cells. As immune-checkpoint inhibitors of pathways regulating T-cell activity are proving efficacious for subsets of cancer patients, we predict that combining these two immune-based approaches represents a compelling clinical opportunity. However, it is likely that not all tumors will respond; thus, identifying predictive biomarkers and correlates of therapeutic response is a top priority. Based on preclinical data evaluating macrophage antagonists and checkpoint inhibitors, we predict that biomarkers of response will also be reflected by changes in peripheral blood. Such biomarkers will form the basis for simple, non-invasive diagnostic and/or prognostic screens aiding early detection in susceptible populations (Figure 4).

In preclinical models, regardless of tumor or approach,  $T_H1$  immunity emerges when dominant  $T_H2$ -driver pathways are attenuated; when concomitant with cytotoxic therapy, tumor growth stalls or regresses by CD8<sup>+</sup> T-cell-dependent mechanisms. These findings highlight the importance of neutralizing pro-tumor inflammation as a therapeutic strategy and indicate that tumor-infiltrating CTLs can be mobilized in tumors with low mutational burdens. These data also highlight the clinical need for biomarkers that identify tissue-specific programs driving  $T_H2$  immune responses; such data is needed to inform



**Figure 4. Multi-Modal Biomarker-Based Approach for Optimal Immune-Mediated Tumor Control**

Cancer medicine is evolving. Going forward, individuals with cancer will be evaluated for biomarkers enabling stratification to determine most optimal combinations for therapy based on tumor-based and systemic biomarkers. Eventually, all patients with cancer will be treated with checkpoint inhibitors, either directly or after interventions targeting inflammation (for example, with  $T_H2$ -blockade therapies, radiation therapy, or epigenetic modulation), or vaccination via DCs to boost T-cell repertoires, or adoptive T cell transfer. Based on the known tissue-embedded programs empowered to control autoimmunity, it is reasonable to anticipate that a majority of patients will develop acquired resistance followed by immune escape; this will lead to the next cycle of treatments incorporating multi-modal biomarkers (e.g., based on microbiome phenotype, circulating cell-free DNA [cfDNA], circulating cytokine levels) and perhaps NK cells recognizing loss of MHC class I by neoplastic cells, thus rendering them invisible to T cells. Cytotoxic treatments such as with NK cells or standard cytotoxic therapy (CTX or RT) or oncolytic viruses will release neo-antigens that can be used for generation of the next round of effector T cells. Whole-exome sequencing (WES) of tumor samples as well as cfDNA will yield information on mutational load that can, in turn, be used as one class of neo-antigens for vaccination and priming of new T-cell repertoires. T-cell receptors (TCR) can be assessed using genomic approaches enabling sequencing of TCR $\beta$  chains to assess repertoire diversity. Given the importance of T-cell specificity for relevant antigens, strategies enabling paired sequencing of  $\alpha$  and  $\beta$  TCR chains will be invaluable, as will high-throughput tetramer analysis. In addition, RNA-seq and epigenetic analysis of tumors and their infiltrates will enable assessment of the type and flavor of inflammation. Future studies will incorporate metabolomics to this biomarker portfolio.

precision medicine strategies employing  $T_H2$  blockade, in concert with other immune, targeted, or cytotoxic approaches (Figure 4).

### Immune-Targeted Therapies Focused on T Cells

#### Basic Principles

Cancer immunotherapy historically relied on two principal mechanisms of action: (1) “passive” immunotherapy via provision of anti-tumor antibodies, e.g., Trastuzumab ( $\alpha$ HER2 mAB) or Rituximab ( $\alpha$ CD20 mAB), or adoptive transfer of cytotoxic T and NK cells; and (2) “active” immunotherapy that mobilizes the patient’s immune cells via checkpoint blockade, i.e., administration of antibodies directed against immune-regulatory checkpoint molecules expressed on T cells or via vaccines that expand antigen-specific T cells. In all circumstances, T cells are the drug; we are learning that T cells have the ability to clinically control some cancers (Postow et al., 2015). T cells can be targeted in three major ways: (1) by being liberated by checkpoint inhibitors; (2) through adoptive transfer when T cells are missing, as validated by the clinical success of genetically engineered T-cell therapies; and (3) through induction in vivo by vaccination or endogenous mechanisms subsequent to other anti-cancer therapies (e.g., CTX, targeted therapies, or anticancer antibodies) (Palucka and Banchereau, 2013).

#### Immune Checkpoint Blockade

Immune checkpoint blockade (ICB) can unleash the power of naturally occurring T cells by eliminating negative signals that block T-cell function (Pardoll, 2012). ICB has produced durable clinical responses and improved survival across a variety of cancers (Topalian et al., 2015). CTL expansion and function are carefully regulated by cytotoxic-T-lymphocyte-associated protein 4 (CTLA-4), programmed cell death protein 1 (PD-1), and other molecules so as to maintain a delicate balance between the resolution of infection, the elimination of infected cells, the protection of tissue homeostasis, and the prevention of autoimmune

attack. CTLA-4 is a cell-surface receptor expressed by activated T cells with homology to the T-cell costimulatory molecule CD28. Although CD28 and CTLA-4 are both ligands for B7-1 (CD80) and B7-2 (CD86), they serve opposing roles in regulating T-cell activation. CD28 provides costimulatory signals required for T-cell activation, whereas CTLA-4 negatively modulates T-cell responses by raising the activation threshold for T-cell priming; thus, CTLA-4 is likely most important during priming. PD-1 binds programmed death ligand 1 (PD-L1; a.k.a., B7-H1 or CD274) expressed by neoplastic cells, various immune cells, mesenchymal support cells, and vascular cells; this interaction negatively regulates T-cell activation when engaged with an APC and/or effector function when engaged with other PD-L1-positive cells. Indeed, binding of PD-L1 to its receptors suppresses T-cell migration and proliferation and restricts cancer cell killing (Topalian et al., 2015); thus, PD-1 is important in regulating effector functions after CD8<sup>+</sup> T cells are activated. PD-1 and CTLA-4 regulate distinct phases of T-cell differentiation and function, and their inhibition might need to be optimally phased for maximum efficacy. This concept needs to be incorporated in the next generation of clinical trials of combination therapy regimens, especially when combined with vaccines or agents that reprogram myeloid cells to foster a T<sub>H</sub>1-type activation state.

Indeed, combination therapies targeting the two checkpoints, i.e., CTLA-4 and PD-1, further increase progression-free survival in patients with metastatic melanoma (Larkin et al., 2015); however, in other cancers, these responses are present in fewer patients. Resolving the natural and acquired resistance to checkpoint inhibition therapy represents the next frontier in basic research and clinical development (Figure 4). As the effector arm of checkpoint inhibition, T cells could underpin the major resistance mechanisms for checkpoint blockade. Thus, non-responding patients might actually lack naturally occurring T cells with specificity against neoplasias, and/or their T cells could be held hostage and rendered dysfunctional in TMEs via pathways other than checkpoints, such as immune-suppressive microenvironments directed by T<sub>H</sub>2-activated myeloid, Treg, or Breg cells (Coussens et al., 2013). Links between treatment resistance and T-cell shortage are supported by recent findings that tumor-specific mutations generate neo-antigens that, in turn, may drive anti-tumor responses. Indeed, whole-exome sequencing of malignant melanomas from patients treated with CTLA-4 blockers demonstrated an association between mutational load and degree of clinical benefit (Snyder et al., 2015); however, in other melanoma cohorts, recurrent neo-antigen peptide sequences were not found to predict responder populations (Van Allen et al., 2015). In non-small-cell lung cancers treated with  $\alpha$ PD-1 mAb, higher mutation burden in tumors was associated with durable clinical benefit and progression-free survival (Rizvi et al., 2015). Neo-antigens arising as products of somatic mutations are not presented in the thymus; thus, they can be recognized by the immune system as foreign antigens, similar to viral antigens or organ transplants, because the T cells have not been eliminated or tolerated. These concepts were put forward early (Srivastava, 2000) but validated only recently (Schumacher and Schreiber, 2015), thanks to the availability of massively parallel sequencing.

### Cancer Vaccines

Patients may fail or resist checkpoint therapy owing to a lack of pre-existing T cell infiltrates. Therefore, vaccination and adoptive transfer strategies to first induce and expand the breadth of endogenous T-cell responses could prove useful. Vaccines are composed of antigens and adjuvants. Responses to vaccination and adjuvants involve DCs that capture and present vaccine antigens, thereby facilitating differentiation of lymphocytes and subsequent immunity (Figure 3). DCs also integrate the adjuvant signals and determine the quality of induced immune responses.

Several phase III clinical trials testing various cancer vaccine platforms, including DC-based and viral-vector-based vaccines, are ongoing. These exogenous vaccine platforms will need to be accompanied by high-throughput genomics to incorporate personalized cancer-specific mutations and candidate peptide antigens. Indeed, proof-of-concept trials in patients with advanced melanoma demonstrated that naturally occurring neo-antigen-specific immunity was enhanced by vaccination with DCs loaded with patient-specific peptides (Carreno et al., 2015). Another concept is endogenous vaccination based on antigen released upon standard CTX/RT or oncolytic viruses (i.e., viruses that preferentially infect and kill cancer cells); this, however, requires effective antigen presentation to generate therapeutic T-cell immunity (Palucka and Banchereau, 2013). DCs are skewed by tumors toward pro-tumor immunity; thus, reprogramming to foster T<sub>H</sub>1-skewed mature functionality in vivo is critical for success of endogenous vaccination. Our understanding of cancer genomics, the biology of antigen presentation, and T-cell biology has advanced impressively and continues to increase; this knowledge will feed into the development of the next-generation cancer vaccines that, combined with checkpoint inhibitors, hold promise for improving outcomes for cancer patients (Figure 4).

DC vaccines can be exploited as anti-cancer therapeutics through various strategies, including: non-targeted peptide/protein and nucleic-acid-based vaccines captured by DCs in vivo; vaccines composed of antigens directly coupled to anti-DC-antibodies; or vaccines composed of ex-vivo-generated DCs loaded with tumor antigens (Palucka and Banchereau, 2013). DCs are also engaged in response to complex vaccine preparations such as GVAX, a tumor-cell-based vaccine in which neoplastic cells genetically modified to express GM-CSF attract and activate DCs (Le et al., 2010). Other vaccine platforms are based on recombinant-attenuated *Listeria monocytogenes* (*Lm*), an intracellular bacterium that targets DCs in vivo and utilizes MHC I and II antigen-processing pathways (Le et al., 2012), as well as intratumoral delivery of oncolytic viruses; these can be modified to express GM-CSF to attract DCs and lymphocytes at the lysed tumor site (Russell et al., 2012). Finally, pioneering studies from Ralph Steinman and Michel Nussenzweig demonstrated the principle of targeting antigens to DCs in vivo through coupling of antigens to antibodies specific to DC surface receptors such as DEC205 or DCIR (Bonifaz et al., 2002; Soares et al., 2007). Importantly, in the absence of adjuvants, targeting antigens to DEC205<sup>+</sup> DCs in vivo induces antigen-specific tolerance (Hawiger et al., 2001). Administration of these complex vaccines with DC activators such as TLR3, TLR7-8, or CD40 agonists enables maturation of DCs and

consequent establishment of immunity rather than tolerance (Bonifaz et al., 2002; Soares et al., 2007). It remains to be seen which vaccine platform will be most effective at priming and boosting T cells in patients; this clearly represents the next frontier in research.

### T-Cell-Dependent Nature of Cytotoxic and Targeted Therapy

Cancer medicine evolved largely based on the principle that rapidly proliferating malignant cells can be eradicated by cytotoxic regimens (CTX or RT) or by targeted drugs attacking attributes of mutationally corrupted cells. As discussed above, the recent advent and remarkable efficacy of immune-checkpoint inhibitors revealed the clinical potential of harnessing endogenous mechanisms of anti-tumor immunity in tumors harboring significant mutational burdens. Upon reflection, however, it is appreciated that conventional cytotoxic approaches modulate the composition and functional bioactivity of intratumoral leukocytes, in addition to effects on neoplastic cells (Galluzzi et al., 2015). Furthermore, in some scenarios, the efficacy of neoadjuvant CTX correlates with increased presence of intratumoral immune-effector T cells (Galluzzi et al., 2015). These correlations are not limited to cytotoxic regimens—the tyrosine kinase inhibitor imatinib also leads to increased presence of CTLs and NK cells in gastrointestinal tumors in a manner that correlates with disease outcome (Kroemer et al., 2013), while efficacy against chronic myelogenous leukemia can be reversed by co-administration of type I IFN (Galluzzi et al., 2015). A recent study in breast cancer also revealed that efficacy of trastuzumab emtansine is linked to elicitation of anti-tumor immune responses (Müller et al., 2015).

Malignant cells can emit danger signals, albeit distinct from those of normal tissue, that are sensed by immune cells and thus are antigenic. Increased antigenicity is linked to either mutational burden, where peptides from mutant proteins are presented by MHC molecules (Gubin et al., 2014), or to ectopic expression of cancer testis or oncofetal antigens typically only expressed during embryonic or fetal development (Whitehurst, 2014). The increased adjuvanticity of neoplastic cells is linked to metabolic stress caused by their sustained proliferation and to their ability to adapt and survive in hypoxic TMEs (Krysko et al., 2012). Furthermore, preclinical data have emerged supporting the proposition that tumors treated with conventional CTX engage antigenic and adjuvant immune-mediated mechanisms. In murine tumor models, the anti-neoplastic effects of anthracyclines are significantly reduced when either  $\gamma\delta$  or CD8<sup>+</sup> T cells are depleted, but not when B or NK cells are absent, DC infiltration is blocked or corrupted, immune-stimulatory type I IFNs or IL-17 are lacking, or DAMP-mediated recruitment and activation of effector cells is thwarted (Kroemer et al., 2013). Cyclophosphamide, oxaliplatin, and bortezomib similarly rely on immune-mediated mechanisms for their efficacy; these commonly used cytotoxics elicit effector cell activity via plasma membrane exposure of calreticulin and release of the chromatin-binding protein high-mobility group box 1 (HMGB1). This, in turn, fosters DC maturation and TLR4 and RAGE activation (Apetoh et al., 2007), thus increasing adjuvanticity of malignant cells.

Taxanes, broadly used microtubule inhibitors, and vinca alkaloids promote polyploidization due to mitotic interference, thus leading to endoplasmic reticulum stress responses favoring calreticulin exposure and immune-mediated elimination (Senovilla et al., 2012). Clinically, docetaxel, vinorelbine, and cisplatin all lead to increased abundance of circulating CTLs and decreased presence of Treg cells and immature myeloid cells harboring T-cell-suppressive activity; this latter effect is also shared by gemcitabine, a common CTX for pancreatic adenocarcinomas, and 5-fluorouracil. Interestingly, paclitaxel is also a TLR4 ligand and thus enhances T-cell priming by DCs (Pfannenstiel et al., 2010).

Cyclophosphamide also provokes relocalization of intestinal gram-positive bacteria to secondary lymphoid organs, resulting in generation of T<sub>H</sub>17 cells secreting IL-17 and IFN $\gamma$  that promote anti-tumor immune responses (Viaud et al., 2013). In murine tumor models, therapies targeting T<sub>H</sub>2-based programs (e.g., CSF1R or RON antagonists, BTK or SYK inhibitors, B-cell depletion,  $\alpha$ IL-4 or  $\alpha$ IL-13 mAbs) enhance efficacy of either CTX or RT by T-cell-dependent mechanisms (Ruffell and Coussens, 2015). Perhaps the most compelling evidence is that provided by recent data revealing that immune-checkpoint blockade, when combined with CTX, improves overall survival in several cancer types beyond CTX alone (Topalian et al., 2015). The ability of these agents to activate adaptive stress-response pathways and send danger signals operative as immunologic adjuvants inherently increases the antigenicity of tumors even when mutational burden is low. These untoward effects can be capitalized upon to improve outcomes for individuals with cancer.

### A Role for the Microbiome in Regulating Systemic Cancer Risk and Response to Therapy

If the precision medicine equation wasn't sufficiently complicated by neoplastic cell genomics, epigenomics, host immune responses, and the TME, mounting evidence points to an additional consideration when attempting to stratify patients and predict therapy response: the host microbiome. The context and composition of common microorganisms living in the gut not only shapes local immune responses, but also regulates systemic immunity and thus impacts the risk of and progression to malignancy and the response to anti-cancer therapies. Intra-abdominal infections and use of antibiotics has long been associated with increased incidence of colorectal cancer (Wang et al., 2014). In mouse models, attenuating or selectively altering the composition of gut microorganisms influences both the incidence and progression of cancer (Zitvogel et al., 2015). Intestinal microorganisms not only impact local risk of tumorigenesis, but also influence neoplastic progression distally by altering inflammatory and metabolic circuitry. These experimental results correlate with epidemiologic data revealing increased incidence of breast cancer in women with significant history of antibiotic use (Zitvogel et al., 2015).

Gut microbiota composition is dramatically impacted by common anti-neoplastic drugs, including RT, allogeneic stem-cell transplantation, and select CTXs, notably 5-fluorouracil and irinotecan (Zitvogel et al., 2015). Along these lines, the gut microbiota affects the amenability of some tumor types to therapy by impacting regulatory aspects of the immune response.

Examples include translocation of gut microbiota across the intestinal epithelium in response to lympho-depleting irradiation in which DCs are inadvertently activated, leading to altered serum cytokines and improved responses to adoptively transferred CTLs; these beneficial effects are abated by antibiotics (Paulos et al., 2007). Similarly, cyclophosphamide alters composition of gut microbiota, resulting in translocation of gram-positive bacteria into secondary lymphoid organs wherein pathogenic  $T_H17$  and memory  $T_H1$  cells are activated; tumors grown in germ-free mice, or antibiotics tropic for gram-positive bacteria exhibit reduced  $T_H17$  responses and tumor resistance to cyclophosphamide (Viaud et al., 2013). Antibiotic eradication of gram-positive bacteria also impairs the efficacy of CpG-oligonucleotide immunotherapy and platinum CTX by altering myeloid cells within the TME (Iida et al., 2013). *Bifidobacterium* occupancy supports anti-tumor immunity against melanoma and improves the efficacy of  $\alpha$ PD-L1 and  $\alpha$ CTLA-4 mAb therapy by altering DC activity, leading to improved antigen-specific CD8<sup>+</sup> T cell function—these effects were reduced by ampicillin, colistin, or streptomycin but were enhanced by vancomycin due to preferential enhancement of *Bacteroidales* colonization (Vétizou et al., 2015). These data underscore the impact of gut commensals on therapeutic responsiveness.

Could selectively manipulating the gut microbiota impact risk of developing cancer, limit incidence of select tumor types, and/or improve activity of some anti-cancer therapies? Zitvogel and colleagues have proposed four distinct approaches for manipulating the gut microbiota to boost cancer therapy: (1) preferential use of antibiotics selective for untoward bacterial species; (2) increased use of probiotics; (3) increased use of prebiotics to stimulate healthy gut colonization; and (4) use of postbiotics, nonviable products of microbiota that exert biological activities in hosts (Zitvogel et al., 2015). Prospective stool analysis and monitoring in cancer patients receiving therapy will surely reveal biomarkers that, if harnessed, could improve patient stratification and/or support new microbiota-based strategies for boosting therapeutic responsiveness, e.g., fecal transplant of beneficial species.

### Multi-Modal Tumor and Systemic Biomarkers for Stratification and Resistance Monitoring

A major clinical goal is to understand the multi-modal tissue-based and systemic pathways regulating therapy responses so as to minimize resistance and maximize efficacy of cancer medicine (Figure 4). Whether therapies target tumor-intrinsic pathways, host pathways, or commensal microbiota, it is clear that understanding non-genomic mechanisms of resistance from an integrated standpoint is critical.

Understanding which immune cell types are present in and around a tumor currently provides invaluable retrospective information regarding tumor ecology and/or tumor response to therapy. However, we must improve our ability to integrate information on not only the complexity of leukocytes in tumors, but also their geography in tumor nests and stroma. Immune cells are scattered in tumor core and within tumor stroma, in invasive margins, and in organized lymphoid structures often distant from neoplastic cells. Investigating the mechanisms governing formation of tertiary lymphoid structures (TLSs) found in numerous

cancers represents a new frontier for biomedical research. Such topology has been reported by Galon and colleagues to be clinically meaningful in colorectal cancer, where a statistically significant correlation between immune cell density and patient outcome was revealed (Galon et al., 2006). Moreover, development of TLS in individuals with pancreatic adenocarcinoma treated with vaccines correlated with improved clinical outcomes (Le et al., 2015). Furthermore, compared with single-region analysis, combined analysis of the tumor core and invasive margins improved the accuracy of survival prediction in different patient groups (Galon et al., 2006). These early results form the basis for immune stratification of patients, or the so-called Immunoscore, and its coordinated assessment in the clinic (Ascierto et al., 2013). An international consortium has been initiated to validate and promote the Immunoscore in routine clinical settings (Ascierto et al., 2013); results of this international effort may lead to implementation of the Immunoscore as a new classification metric, designated TNM-I (TNM-Immune).

Will the Immunoscore provide enough additional information to prospectively predict response to therapy? Likely not. We predict that integrating the Immunoscore with additional metrics will be critical for guiding patient stratification and phasing of combinatorial therapies. Such metrics will include genomic and exomic features of neoplastic cells (through sequencing of neoplastic cells themselves or cfDNA, tracking the expansion of tumor-specific CD4<sup>+</sup> and CD8<sup>+</sup> T cells, monitoring serum cytokine fluxes, and evaluating the composition and health of commensal bacteria (Figure 4). Serum cytokines have long been enigmatic due to their labile nature and the detection limitations of conventional methodologies. That said, serum biomarker signatures are now able to discern asymptomatic early stage pancreatic cancer from healthy controls with 96% accuracy (Ghatnekar et al., 2013) and can be used to monitor the pharmacodynamics of CSF1R-targeted therapies (Butowski et al., 2015). Moreover, transcriptional profiling of blood monocytes in renal cell carcinoma identifies biomarkers correlating with tumor staging (Chittezhath et al., 2014), and mRNA sequencing of tumor-educated blood platelets distinguishes cancer patients from healthy individuals with 96% accuracy (Best et al., 2015). Thus, multi-modal functional diagnostic strategies integrating the tumor, host, and commensals will likely forge the advent of next-generation precision bioinformatics to match patients with appropriately combined and phased anti-cancer therapies.

### Oncoimmunology Treatment Paradigm

Future immunotherapies will be based on cycles of interventions designed to boost and modulate anti-cancer immunity (Figure 4). Indeed, as we rediscover and refine the fundamental principles of tumor immunology, it is increasingly clear that curing cancer might not be a realistic goal. Rather, aiming for a continuum of treatment cycles designed and based on mechanistic *in vivo* studies and *in-depth* analysis of each patient's tumor will be necessary for optimizing outcomes. Clinical trials with checkpoint inhibitors teach us that *in situ* immune infiltration is critical for tumor regression; however, not all immune infiltrates are equal, and as discussed throughout this article, the quality of immune response is a critical factor for therapeutic success. This,

in turn, is determined by underlying inflammation, which we assert must become a staging parameter, along with classical pathology-based schemas and the Immunoscore. It will also need to be established to what extent inflammation, which clearly plays a role in epithelial tumors, impacts other tumor types, e.g., melanoma or sarcomas. Additional parameters pertain to the specificity of infiltrating T cells against cancer antigens, as again, the infiltrate with passenger T cells might not be therapeutically useful and should be tested. Eventually all patients will be treated with checkpoint inhibitors, either directly or after interventions targeting inflammation, by vaccination to boost T-cell repertoires or by adoptive T-cell transfer. The majority of patients will subsequently develop acquired resistance followed by immune escape; this will lead to the next cycle of treatments incorporating multi-modal biomarkers (e.g., based on microbiome phenotype, cfDNA, circulating cytokine levels) and perhaps NK cells recognizing loss of MHC class I by neoplastic cells, thus rendering them invisible to T cells. Cytotoxic treatments, such as with NK cells, standard CTX/RT, or oncolytic viruses, will release neo-antigens that can be used to generate the next round of effector T cells. To this latter point, we must fully understand the rules of T-cell priming in vivo in humans, identify the most effective ways to utilize DCs for priming, and develop strategies for mobilizing the naive T-cell repertoire from the thymus in adults (Sportès et al., 2008). In later rounds of therapy, the scope of neo-antigens will likely be broadened as, in addition to somatic mutations, neo-antigens can be generated via epigenetic and post-translational regulation. Last but not least, the role of Tregs, so well established in murine cancer, will need to be redefined in humans. Resolving all of these challenges will surely keep us busy for a long while.

#### ACKNOWLEDGMENTS

The authors thank members of the Palucka and Coussens laboratories for critical insight and discussions; Amanda Lund and Jeffrey Nolz for critical feedback; Anna Lisa Lucido for help with editing the manuscript; and all authors contributing to studies discussed herein but not mentioned due to space consideration. A.K.P. thanks Jacques Banchereau, a long-time collaborator. The authors acknowledge support from the NIH/NCI and Susan B. Komen Foundation (A.K.P. and L.M.C.), Baylor Research Foundation (A.K.P.), DOD BCRP Era of Hope Scholar Expansion Award, Breast Cancer Research Foundation, a Stand Up To Cancer–Lustgarten Foundation Pancreatic Cancer Convergence Dream Team Translational Research Grant (SU2C-AACR-DT14-14), and Brenden-Colson Center for Pancreatic Care (to L.M.C.).

#### REFERENCES

- Affara, N.I., Ruffell, B., Medler, T.R., Gunderson, A.J., Johansson, M., Bornstein, S., Bergsland, E., Steinhoff, M., Li, Y., Gong, Q., et al. (2014). B cells regulate macrophage phenotype and response to chemotherapy in squamous carcinomas. *Cancer Cell* 25, 809–821.
- Aharinejad, S., Salama, M., Paulus, P., Zins, K., Berger, A., and Singer, C.F. (2013). Elevated CSF1 serum concentration predicts poor overall survival in women with early breast cancer. *Endocr. Relat. Cancer* 20, 777–783.
- Andreu, P., Johansson, M., Affara, N.I., Pucci, F., Tan, T., Junankar, S., Korets, L., Lam, J., Tawfik, D., DeNardo, D.G., et al. (2010). FcγR activation regulates inflammation-associated squamous carcinogenesis. *Cancer Cell* 17, 121–134.
- Apetoh, L., Ghiringhelli, F., Tesniere, A., Criollo, A., Ortiz, C., Lidereau, R., Mariette, C., Chaput, N., Mira, J.P., Delaloge, S., et al. (2007). The interaction between HMGB1 and TLR4 dictates the outcome of anticancer chemotherapy and radiotherapy. *Immunol. Rev.* 220, 47–59.
- Appay, V., Douek, D.C., and Price, D.A. (2008). CD8+ T cell efficacy in vaccination and disease. *Nat. Med.* 14, 623–628.
- Ascierto, P.A., Capone, M., Urba, W.J., Bifulco, C.B., Botti, G., Lugli, A., Marincola, F.M., Ciliberto, G., Galon, J., and Fox, B.A. (2013). The additional facet of immunoscore: immunoprofiling as a possible predictive tool for cancer treatment. *J. Transl. Med.* 11, 54.
- Balkwill, F., Charles, K.A., and Mantovani, A. (2005). Smoldering and polarized inflammation in the initiation and promotion of malignant disease. *Cancer Cell* 7, 211–217.
- Banchereau, J., Briere, F., Caux, C., Davoust, J., Lebecque, S., Liu, Y.J., Puelandran, B., and Palucka, K. (2000). Immunobiology of dendritic cells. *Annu. Rev. Immunol.* 18, 767–811.
- Beck, A.H., Espinosa, I., Edris, B., Li, R., Montgomery, K., Zhu, S., Varma, S., Marinelli, R.J., van de Rijn, M., and West, R.B. (2009). The macrophage colony-stimulating factor 1 response signature in breast carcinoma. *Clin. Cancer Res.* 15, 778–787.
- Best, M.G., Sol, N., Kooi, I., Tannous, J., Westerman, B.A., Rustenburg, F., Schellen, P., Verschuere, H., Post, E., Koster, J., et al. (2015). RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. *Cancer Cell* 28, 666–676.
- Bongartz, T., Sutton, A.J., Sweeting, M.J., Buchan, I., Matteson, E.L., and Montori, V. (2006). Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. *JAMA* 295, 2275–2285.
- Bonifaz, L., Bonnyay, D., Mahnke, K., Rivera, M., Nussenzweig, M.C., and Steinman, R.M. (2002). Efficient targeting of protein antigen to the dendritic cell receptor DEC-205 in the steady state leads to antigen presentation on major histocompatibility complex class I products and peripheral CD8+ T cell tolerance. *J. Exp. Med.* 196, 1627–1638.
- Butowski, N., Colman, H., De Groot, J.F., Omuro, A.M., Nayak, L., Wen, P.Y., Cloughesy, T.F., Marimuthu, A., Haidar, S., Perry, A., et al. (2015). Orally administered colony stimulating factor 1 receptor inhibitor PLX3397 in recurrent glioblastoma: an Ivy Foundation Early Phase Clinical Trials Consortium phase II study. *Neuro. Oncol.* Published online October 8, 2015. <http://dx.doi.org/10.1093/neuonc/nov245>.
- Campisi, J., Andersen, J.K., Kapahi, P., and Melov, S. (2011). Cellular senescence: a link between cancer and age-related degenerative disease? *Semin. Cancer Biol.* 21, 354–359.
- Carreno, B.M., Magrini, V., Becker-Hapak, M., Kaabinejadian, S., Hundal, J., Petti, A.A., Ly, A., Lie, W.R., Hildebrand, W.H., Mardis, E.R., and Linette, G.P. (2015). Cancer immunotherapy. A dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science* 348, 803–808.
- Carretero, R., Sektioglu, I.M., Garbi, N., Salgado, O.C., Beckhove, P., and Hämmerling, G.J. (2015). Eosinophils orchestrate cancer rejection by normalizing tumor vessels and enhancing infiltration of CD8(+) T cells. *Nat. Immunol.* 16, 609–617.
- Chen, D.S., and Mellman, I. (2013). Oncology meets immunology: the cancer-immunity cycle. *Immunity* 39, 1–10.
- Chittezhath, M., Dhillon, M.K., Lim, J.Y., Laoui, D., Shalova, I.N., Teo, Y.L., Chen, J., Kamaraj, R., Raman, L., Lum, J., et al. (2014). Molecular profiling reveals a tumor-promoting phenotype of monocytes and macrophages in human cancer progression. *Immunity* 41, 815–829.
- Coffelt, S.B., Kersten, K., Doornebal, C.W., Weiden, J., Vrijland, K., Hau, C.S., Versteegen, N.J., Ciampicotti, M., Hawinkels, L.J., Jonkers, J., et al. (2015). 17-producing gammadelta T cells and neutrophils conspire to promote breast cancer metastasis (IL: Nature).

- Coussens, L.M., Raymond, W.W., Bergers, G., Laig-Webster, M., Behrendtsen, O., Werb, Z., Caughey, G.H., and Hanahan, D. (1999). Inflammatory mast cells up-regulate angiogenesis during squamous epithelial carcinogenesis. *Genes Dev.* *13*, 1382–1397.
- Coussens, L.M., Zitvogel, L., and Palucka, A.K. (2013). Neutralizing tumor-promoting chronic inflammation: a magic bullet? *Science* *339*, 286–291.
- Davis, B.P., and Rothenberg, M.E. (2014). Eosinophils and cancer. *Cancer Immunol. Res.* *2*, 1–8.
- de Visser, K.E., Eichten, A., and Coussens, L.M. (2006). Paradoxical roles of the immune system during cancer development. *Nat. Rev. Cancer* *6*, 24–37.
- DeNardo, D.G., Barreto, J.B., Andreu, P., Vasquez, L., Tawfik, D., Kolhatkar, N., and Coussens, L.M. (2009). CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages. *Cancer Cell* *16*, 91–102.
- DeNardo, D.G., Brennan, D.J., Rexhepaj, E., Ruffell, B., Shiao, S.L., Madden, S.F., Gallagher, W.M., Wadhvani, N., Keil, S.D., Junaid, S.A., et al. (2011). Leukocyte complexity predicts breast cancer survival and functionally regulates response to chemotherapy. *Cancer Discov.* *1*, 54–67.
- Druker, B.J., Guilhot, F., O'Brien, S.G., Gathmann, I., Kantarjian, H., Gattermann, N., Deininger, M.W., Silver, R.T., Goldman, J.M., Stone, R.M., et al.; IRIS Investigators (2006). Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. *N. Engl. J. Med.* *355*, 2408–2417.
- Dunn, G.P., Old, L.J., and Schreiber, R.D. (2004). The three Es of cancer immunoeediting. *Annu. Rev. Immunol.* *22*, 329–360.
- Eyob, H., Ekiz, H.A., Deroose, Y.S., Waltz, S.E., Williams, M.A., and Welm, A.L. (2013). Inhibition of ron kinase blocks conversion of micrometastases to overt metastases by boosting antitumor immunity. *Cancer Discov.* *3*, 751–760.
- Fehérvari, Z., and Sakaguchi, S. (2004). CD4+ Tregs and immune control. *J. Clin. Invest.* *114*, 1209–1217.
- Fridlender, Z.G., and Albelda, S.M. (2012). Tumor-associated neutrophils: friend or foe? *Carcinogenesis* *33*, 949–955.
- Gabrilovich, D.I., Ostrand-Rosenberg, S., and Bronte, V. (2012). Coordinated regulation of myeloid cells by tumours. *Nat. Rev. Immunol.* *12*, 253–268.
- Galluzzi, L., Buqué, A., Kepp, O., Zitvogel, L., and Kroemer, G. (2015). Immunological Effects of Conventional Chemotherapy and Targeted Anticancer Agents. *Cancer Cell* *28*, 690–714.
- Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., et al. (2006). Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* *313*, 1960–1964.
- Geissmann, F., Jung, S., and Littman, D.R. (2003). Blood monocytes consist of two principal subsets with distinct migratory properties. *Immunity* *19*, 71–82.
- Ghatnekar, O., Andersson, R., Svensson, M., Persson, U., Ringdahl, U., Zeilon, P., and Borrebaeck, C.A. (2013). Modelling the benefits of early diagnosis of pancreatic cancer using a biomarker signature. *Int. J. Cancer* *133*, 2392–2397.
- Gocheva, V., Wang, H.W., Gadea, B.B., Shree, T., Hunter, K.E., Garfall, A.L., Berman, T., and Joyce, J.A. (2010). IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion. *Genes Dev.* *24*, 241–255.
- Gubin, M.M., Zhang, X., Schuster, H., Caron, E., Ward, J.P., Noguchi, T., Ivanova, Y., Hundal, J., Arthur, C.D., Krebber, W.J., et al. (2014). Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. *Nature* *515*, 577–581.
- Gunderson, A.J., Kaneda, M.M., Tsujikawa, T., Nguyen, A.V., Affara, N.I., Ruffell, B., Gorjestani, S., Liudahl, S.M., Truitt, M., Olson, P., et al. (2016). Bruton's Tyrosine Kinase (BTK)-dependent immune cell crosstalk drives pancreas cancer. *Cancer Discov.* Published online December 29, 2015. <http://dx.doi.org/10.1158/2159-8290>.
- Hanahan, D., and Coussens, L.M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* *21*, 309–322.
- Hanna, R.N., Cekic, C., Sag, D., Tacke, R., Thomas, G.D., Nowyhed, H., Herrley, E., Rasquinha, N., McArdle, S., Wu, R., et al. (2015). Patrolling monocytes control tumor metastasis to the lung. *Science* *350*, 985–990.
- Hawiger, D., Inaba, K., Dorsett, Y., Guo, M., Mahnke, K., Rivera, M., Ravetch, J.V., Steinman, R.M., and Nussenzweig, M.C. (2001). Dendritic cells induce peripheral T cell unresponsiveness under steady state conditions in vivo. *J. Exp. Med.* *194*, 769–779.
- Heiser, L.M., Sadanandam, A., Kuo, W.L., Benz, S.C., Goldstein, T.C., Ng, S., Gibb, W.J., Wang, N.J., Ziyad, S., Tong, F., et al. (2012). Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc. Natl. Acad. Sci. USA* *109*, 2724–2729.
- Hendriks, R.W., Yuvaraj, S., and Kil, L.P. (2014). Targeting Bruton's tyrosine kinase in B cell malignancies. *Nat. Rev. Cancer* *14*, 219–232.
- Iida, N., Dzutsev, A., Stewart, C.A., Smith, L., Bouladoux, N., Weingarten, R.A., Molina, D.A., Salcedo, R., Back, T., Cramer, S., et al. (2013). Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* *342*, 967–970.
- Inoue, S., Leitner, W.W., Golding, B., and Scott, D. (2006). Inhibitory effects of B cells on antitumor immunity. *Cancer Res.* *66*, 7741–7747.
- Joffre, O.P., Segura, E., Savina, A., and Amigorena, S. (2012). Cross-presentation by dendritic cells. *Nat. Rev. Immunol.* *12*, 557–569.
- Karsten, C.M., and Köhl, J. (2012). The immunoglobulin, IgG Fc receptor and complement triangle in autoimmune diseases. *Immunobiology* *217*, 1067–1079.
- Kerbel, R. (2011). Anti-angiogenesis in cancer; met and unmet goals - an interview with Robert Kerbel by Francesco Bertolini. *Int. J. Dev. Biol.* *55*, 395–398.
- Kitamura, T., Qian, B.Z., and Pollard, J.W. (2015). Immune cell promotion of metastasis. *Nat. Rev. Immunol.* *15*, 73–86.
- Klebanoff, C.A., Acquavella, N., Yu, Z., and Restifo, N.P. (2011). Therapeutic cancer vaccines: are we there yet? *Immunol. Rev.* *239*, 27–44.
- Klechevsky, E., Morita, R., Liu, M., Cao, Y., Coquery, S., Thompson-Snipes, L., Briere, F., Chaussabel, D., Zurawski, G., Palucka, A.K., et al. (2008). Functional specializations of human epidermal Langerhans cells and CD14+ dermal dendritic cells. *Immunity* *29*, 497–510.
- Komohara, Y., Jinushi, M., and Takeya, M. (2014). Clinical significance of macrophage heterogeneity in human malignant tumors. *Cancer Sci.* *105*, 1–8.
- Kroemer, G., Galluzzi, L., Kepp, O., and Zitvogel, L. (2013). Immunogenic cell death in cancer therapy. *Annu. Rev. Immunol.* *31*, 51–72.
- Krysko, D.V., Garg, A.D., Kaczmarek, A., Krysko, O., Agostinis, P., and Vandenabeele, P. (2012). Immunogenic cell death and DAMPs in cancer therapy. *Nat. Rev. Cancer* *12*, 860–875.
- Lanzavecchia, A., and Sallusto, F. (2001). The instructive role of dendritic cells on T cell responses: lineages, plasticity and kinetics. *Curr. Opin. Immunol.* *13*, 291–298.
- Larkin, J., Chiarion-Sileni, V., Gonzalez, R., Grob, J.J., Cowey, C.L., Lao, C.D., Schadendorf, D., Dummer, R., Smylie, M., Rutkowski, P., et al. (2015). Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N. Engl. J. Med.* *373*, 23–34.
- Le, D.T., Pardoll, D.M., and Jaffee, E.M. (2010). Cellular vaccine approaches. *Cancer J.* *16*, 304–310.
- Le, D.T., Dubenksy, T.W., Jr., and Brockstedt, D.G. (2012). Clinical development of Listeria monocytogenes-based immunotherapies. *Semin. Oncol.* *39*, 311–322.
- Le, D.T., Wang-Gillam, A., Picozzi, V., Greten, T.F., Crocenzi, T., Springett, G., Morse, M., Zeh, H., Cohen, D., Fine, R.L., et al. (2015). Safety and survival with GVAX pancreas prime and Listeria Monocytogenes-expressing mesothelin (CRS-207) boost vaccines for metastatic pancreatic cancer. *J. Clin. Oncol.* *33*, 1325–1333.
- Le Floc'h, A., Jalil, A., Vergnon, I., Le Maux Chansac, B., Lazar, V., Bismuth, G., Chouaib, S., and Mami-Chouaib, F. (2007). Alpha E beta 7 integrin interaction with E-cadherin promotes antitumor CTL activity by triggering lytic granule polarization and exocytosis. *J. Exp. Med.* *204*, 559–570.

- Lin, E.Y., Nguyen, A.V., Russell, R.G., and Pollard, J.W. (2001). Colony-stimulating factor 1 promotes progression of mammary tumors to malignancy. *J. Exp. Med.* *193*, 727–740.
- Marichal, T., Tsai, M., and Galli, S.J. (2013). Mast cells: potential positive and negative roles in tumor biology. *Cancer Immunol. Res.* *1*, 269–279.
- Massó-Vallés, D., Jauset, T., Serrano, E., Sodir, N.M., Pedersen, K., Affara, N.I., Whitfield, J.R., Beaulieu, M.E., Evan, G.I., Elias, L., et al. (2015). Ibrutinib exerts potent antifibrotic and antitumor activities in mouse models of pancreatic adenocarcinoma. *Cancer Res.* *75*, 1675–1681.
- Merad, M., Sathe, P., Helft, J., Miller, J., and Mortha, A. (2013). The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annu. Rev. Immunol.* *31*, 563–604.
- Metz, M., Grimbaldston, M.A., Nakae, S., Piliponsky, A.M., Tsai, M., and Galli, S.J. (2007). Mast cells in the promotion and limitation of chronic inflammation. *Immunol. Rev.* *217*, 304–328.
- Mitchem, J.B., Brennan, D.J., Knolhoff, B.L., Belt, B.A., Zhu, Y., Sanford, D.E., Belaygorod, L., Carpenter, D., Collins, L., Piwnica-Worms, D., et al. (2013). Targeting tumor-infiltrating macrophages decreases tumor-initiating cells, relieves immunosuppression, and improves chemotherapeutic responses. *Cancer Res.* *73*, 1128–1141.
- Mueller, S.N., Gebhardt, T., Carbone, F.R., and Heath, W.R. (2012). Memory T cell subsets, migration patterns, and tissue residence. *Annu. Rev. Immunol.* *31*, 137–161.
- Müller, P., Kreuzaler, M., Khan, T., Thommen, D.S., Martin, K., Glatz, K., Savic, S., Harbeck, N., Nitz, U., Gluz, O., et al. (2015). Trastuzumab emtansine (T-DM1) renders HER2+ breast cancer highly susceptible to CTLA-4/PD-1 blockade. *Sci. Transl. Med.* *7*, 315ra188.
- Palucka, K., and Banchereau, J. (2013). Dendritic-cell-based therapeutic cancer vaccines. *Immunity* *39*, 38–48.
- Pardoll, D.M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* *12*, 252–264.
- Paulos, C.M., Wrzesinski, C., Kaiser, A., Hinrichs, C.S., Chieppa, M., Cassard, L., Palmer, D.C., Boni, A., Muranski, P., Yu, Z., et al. (2007). Microbial translocation augments the function of adoptively transferred self/tumor-specific CD8+ T cells via TLR4 signaling. *J. Clin. Invest.* *117*, 2197–2204.
- Pfannenstiel, L.W., Lam, S.S., Emens, L.A., Jaffee, E.M., and Armstrong, T.D. (2010). Paclitaxel enhances early dendritic cell maturation and function through TLR4 signaling in mice. *Cell. Immunol.* *263*, 79–87.
- Plaks, V., Boldajipour, B., Linnemann, J.R., Nguyen, N.H., Kersten, K., Wolf, Y., Casbon, A.J., Kong, N., van den Bijgaart, R.J., Sheppard, D., et al. (2015). Adaptive Immune Regulation of Mammary Postnatal Organogenesis. *Dev. Cell* *34*, 493–504.
- Postow, M.A., Callahan, M.K., and Wolchok, J.D. (2015). Immune Checkpoint Blockade in Cancer Therapy. *J. Clin. Oncol.* *33*, 1974–1982.
- Pulendran, B. (2015). The varieties of immunological experience: of pathogens, stress, and dendritic cells. *Annu. Rev. Immunol.* *33*, 563–606.
- Pylayeva-Gupta, Y., Lee, K.E., Hajdu, C.H., Miller, G., and Bar-Sagi, D. (2012). Oncogenic Kras-induced GM-CSF production promotes the development of pancreatic neoplasia. *Cancer Cell* *21*, 836–847.
- Qian, B.Z., Li, J., Zhang, H., Kitamura, T., Zhang, J., Campion, L.R., Kaiser, E.A., Snyder, L.A., and Pollard, J.W. (2011). CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature* *475*, 222–225.
- Ries, C.H., Cannarile, M.A., Hoves, S., Benz, J., Wartha, K., Runza, V., Rey-Giraud, F., Pradel, L.P., Feuerhake, F., Klamann, I., et al. (2014). Targeting tumor-associated macrophages with anti-CSF-1R antibody reveals a strategy for cancer therapy. *Cancer Cell* *25*, 846–859.
- Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., et al. (2015). Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* *348*, 124–128.
- Ruffell, B., and Coussens, L.M. (2015). Macrophages and therapeutic resistance in cancer. *Cancer Cell* *27*, 462–472.
- Ruffell, B., Chang-Strachan, D., Chan, V., Rosenbusch, A., Ho, C.M., Pryer, N., Daniel, D., Hwang, E.S., Rugo, H.S., and Coussens, L.M. (2014). Macrophage IL-10 blocks CD8+ T cell-dependent responses to chemotherapy by suppressing IL-12 expression in intratumoral dendritic cells. *Cancer Cell* *26*, 623–637.
- Russell, S.J., Peng, K.W., and Bell, J.C. (2012). Oncolytic virotherapy. *Nat. Biotechnol.* *30*, 658–670.
- Sagiv-Barfi, I., Kohrt, H.E., Czerwinski, D.K., Ng, P.P., Chang, B.Y., and Levy, R. (2015). Therapeutic antitumor immunity by checkpoint blockade is enhanced by ibrutinib, an inhibitor of both BTK and ITK. *Proc. Natl. Acad. Sci. USA* *112*, E966–E972.
- Sandoval, F., Terme, M., Nizard, M., Badoual, C., Bureau, M.F., Freyburger, L., Clement, O., Marcheteau, E., Gey, A., Fraise, G., et al. (2013). Mucosal imprinting of vaccine-induced CD8+ T cells is crucial to inhibit the growth of mucosal tumors. *Sci. Transl. Med.* *5*, 172ra20.
- Schumacher, T.N., and Schreiber, R.D. (2015). Neoantigens in cancer immunotherapy. *Science* *348*, 69–74.
- Senovilla, L., Vitale, I., Martins, I., Tailler, M., Pailleret, C., Michaud, M., Galuzzi, L., Adjemian, S., Kepp, O., Niso-Santano, M., et al. (2012). An immunosurveillance mechanism controls cancer cell ploidy. *Science* *337*, 1678–1684.
- Shepard, H.M., Lewis, G.D., Sarup, J.C., Fendly, B.M., Maneval, D., Mordenti, J., Figari, I., Kotts, C.E., Palladino, M.A., Jr., Ullrich, A., et al. (1991). Monoclonal antibody therapy of human cancer: taking the HER2 protooncogene to the clinic. *J. Clin. Immunol.* *11*, 117–127.
- Sheridan, B.S., and Lefrançois, L. (2011). Regional and mucosal memory T cells. *Nat. Immunol.* *12*, 485–491.
- Shiao, S.L., Ruffell, B., DeNardo, D.G., Faddegon, B.A., Park, C.C., and Coussens, L.M. (2015). TH2-Polarized CD4(+) T Cells and Macrophages Limit Efficacy of Radiotherapy. *Cancer Immunol. Res.* *3*, 518–525.
- Sippel, T.R., White, J., Nag, K., Tsvankin, V., Klaassen, M., Kleinschmidt-DeMasters, B.K., and Waziri, A. (2011). Neutrophil degranulation and immunosuppression in patients with GBM: restoration of cellular immune function by targeting arginase I. *Clin. Cancer Res.* *17*, 6992–7002.
- Smith, C.I., Islam, T.C., Mattsson, P.T., Mohamed, A.J., Nore, B.F., and Vihinen, M. (2001). The Tec family of cytoplasmic tyrosine kinases: mammalian Btk, Bmx, Itk, Tec, Txk and homologs in other species. *BioEssays* *23*, 436–446.
- Snyder, A., Wolchok, J.D., and Chan, T.A. (2015). Genetic basis for clinical response to CTLA-4 blockade. *N. Engl. J. Med.* *372*, 783.
- Soares, H., Waechter, H., Glaichenhaus, N., Mougneau, E., Yagita, H., Mizelina, O., Dudziak, D., Nussenzweig, M.C., and Steinman, R.M. (2007). A subset of dendritic cells induces CD4+ T cells to produce IFN-gamma by an IL-12-independent but CD70-dependent mechanism in vivo. *J. Exp. Med.* *204*, 1095–1106.
- Soucek, L., Buggy, J.J., Kortlever, R., Adimoolam, S., Monclús, H.A., Allende, M.T., Swigart, L.B., and Evan, G.I. (2011). Modeling pharmacological inhibition of mast cell degranulation as a therapy for insulinoma. *Neoplasia* *13*, 1093–1100.
- Sportès, C., Hakim, F.T., Memon, S.A., Zhang, H., Chua, K.S., Brown, M.R., Fleisher, T.A., Krumlauf, M.C., Babb, R.R., Chow, C.K., et al. (2008). Administration of rhIL-7 in humans increases in vivo TCR repertoire diversity by preferential expansion of naive T cell subsets. *J. Exp. Med.* *205*, 1701–1714.
- Srivastava, P.K. (2000). Immunotherapy of human cancer: lessons from mice. *Nat. Immunol.* *1*, 363–366.
- Steinman, R. (2011). Cancer therapeutics: time to swim downstream? *Oncologist* *16*, 1479–1480.
- Steinman, R.M. (2012). Decisions about dendritic cells: past, present, and future. *Annu. Rev. Immunol.* *30*, 1–22.
- Stout, R.D., and Bottomly, K. (1989). Antigen-specific activation of effector macrophages by IFN-gamma producing (TH1) T cell clones. Failure of IL-4-producing (TH2) T cell clones to activate effector function in macrophages. *J. Immunol.* *142*, 760–765.

- Tap, W.D., Wainberg, Z.A., Anthony, S.P., Ibrahim, P.N., Zhang, C., Healey, J.H., Chmielowski, B., Staddon, A.P., Cohn, A.L., Shapiro, G.I., et al. (2015). Structure-Guided Blockade of CSF1R Kinase in Tenosynovial Giant-Cell Tumor. *N. Engl. J. Med.* *373*, 428–437.
- Thun, M.J., Henley, S.J., and Gansler, T. (2004). Inflammation and cancer: an epidemiological perspective. *Novartis Found. Symp.* *256*, 6–21, discussion 22–28, 49–52, 266–269.
- Topalian, S.L., Drake, C.G., and Pardoll, D.M. (2015). Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* *27*, 450–461.
- Van Allen, E.M., Miao, D., Schilling, B., Shukla, S.A., Blank, C., Zimmer, L., Sucker, A., Hillen, U., Foppen, M.H., Goldinger, S.M., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* *350*, 207–211.
- Vétizou, M., Pitt, J.M., Daillère, R., Lepage, P., Waldschmitt, N., Flament, C., Rusakiewicz, S., Routy, B., Roberti, M.P., Duong, C.P., et al. (2015). Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* *350*, 1079–1084.
- Viaud, S., Saccheri, F., Mignot, G., Yamazaki, T., Daillère, R., Hannani, D., Enot, D.P., Pfirschke, C., Engblom, C., Pittet, M.J., et al. (2013). The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. *Science* *342*, 971–976.
- Wang, L., Yi, T., Kortylewski, M., Pardoll, D.M., Zeng, D., and Yu, H. (2009). IL-17 can promote tumor growth through an IL-6-Stat3 signaling pathway. *J. Exp. Med.* *206*, 1457–1464.
- Wang, J.L., Chang, C.H., Lin, J.W., Wu, L.C., Chuang, L.M., and Lai, M.S. (2014). Infection, antibiotic therapy and risk of colorectal cancer: a nationwide nested case-control study in patients with Type 2 diabetes mellitus. *Int. J. Cancer* *135*, 956–967.
- Whitehurst, A.W. (2014). Cause and consequence of cancer/testis antigen activation in cancer. *Annu. Rev. Pharmacol. Toxicol.* *54*, 251–272.
- Wu, S., Rhee, K.J., Albesiano, E., Rabizadeh, S., Wu, X., Yen, H.R., Huso, D.L., Brancati, F.L., Wick, E., McAllister, F., et al. (2009). A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* *15*, 1016–1022.
- Yu, C.I., Becker, C., Wang, Y., Marches, F., Helft, J., Leboeuf, M., Anguiano, E., Pourpe, S., Goller, K., Pascual, V., et al. (2013). Human CD1c+ dendritic cells drive the differentiation of CD103+ CD8+ mucosal effector T cells via the cytokine TGF- $\beta$ . *Immunity* *38*, 818–830.
- Zhu, J., and Paul, W.E. (2008). CD4 T cells: fates, functions, and faults. *Blood* *112*, 1557–1569.
- Zitvogel, L., Galluzzi, L., Viaud, S., Vétizou, M., Daillère, R., Merad, M., and Kroemer, G. (2015). Cancer and the gut microbiota: an unexpected link. *Sci. Transl. Med.* *7*, 271ps1.

# Modeling Development and Disease with Organoids

Hans Clevers<sup>1,\*</sup>

<sup>1</sup>Hubrecht Institute/Royal Netherlands Academy of Arts and Sciences, Princess Maxima Centre and University Medical Centre Utrecht, 3584CT Utrecht, The Netherlands

\*Correspondence: h.clevers@hubrecht.eu

<http://dx.doi.org/10.1016/j.cell.2016.05.082>

Recent advances in 3D culture technology allow embryonic and adult mammalian stem cells to exhibit their remarkable self-organizing properties, and the resulting organoids reflect key structural and functional properties of organs such as kidney, lung, gut, brain and retina. Organoid technology can therefore be used to model human organ development and various human pathologies ‘in a dish.’ Additionally, patient-derived organoids hold promise to predict drug response in a personalized fashion. Organoids open up new avenues for regenerative medicine and, in combination with editing technology, for gene therapy. The many potential applications of this technology are only beginning to be explored.

In 1975, James Rheinwald and Howard Green described the first long-term culture of normal human cells (Rheinwald and Green, 1975). For this, they combined freshly isolated keratinocytes with irradiated mouse 3T3 fibroblasts, established in the same lab years earlier. As in stratified skin, cell division was confined to the basal layer of the growing clones, while superficial layers consisted of terminally differentiating keratinocytes that gradually developed a cornified cell envelope. Successive improvements in the methodology allowed the cultivation of large confluent sheets of epidermis grown from relatively small numbers of primary proliferative keratinocytes (for which the term “stem cell” was not applied). Green and co-workers performed the first successful treatment of two third-degree burn patients with cultured autologous keratinocyte sheets at the Peter Bent Brigham Hospital in 1980 (O’connor et al., 1981). In a particularly dramatic demonstration of the potential of the method, they showed, in the summer of 1983, that this approach was life-saving for the 5-year-old Jamie Selby and his 6-year-old brother Glen, who had both sustained burns over >95% of their body surface (Gallico et al., 1984).

In his own lab, Rheinwald built on this work to establish a comparable method for culturing another stratified squamous epithelium, the cornea (Lindberg et al., 1993). De Luca and Pellegrini applied this technology for the treatment of corneal blindness with a high rate of success, as reported upon long-term follow-up of 112 patients. Their procedure was straightforward: a 1–2 mm biopsy from the limbal region of the healthy eye was grown in culture on 3T3 feeder cells, and the resulting sheet was grafted onto the injured eye (Pellegrini et al., 1997; Rama et al., 2010). While the term “organoid” was not used in these pioneering studies, Rheinwald and Green were the first to reconstitute 3D tissue structure from cultured human stem cells.

Organoids revealed their first popularity in the years 1965–1985, shown by an increase in the PubMed search term “organoids” (Figure 1), mostly in classic developmental biology experiments that sought to describe organogenesis by cell dissociation and reaggregation experiments (for an overview, see Lancaster and Knoblich, 2014). The past 7–8 years have

witnessed a revival of the organoid, yet in a somewhat different guise: an organoid is now defined as a 3D structure grown from stem cells and consisting of organ-specific cell types that self-organizes through cell sorting and spatially restricted lineage commitment (after Eiraku and Sasai, 2012; Lancaster and Knoblich, 2014).

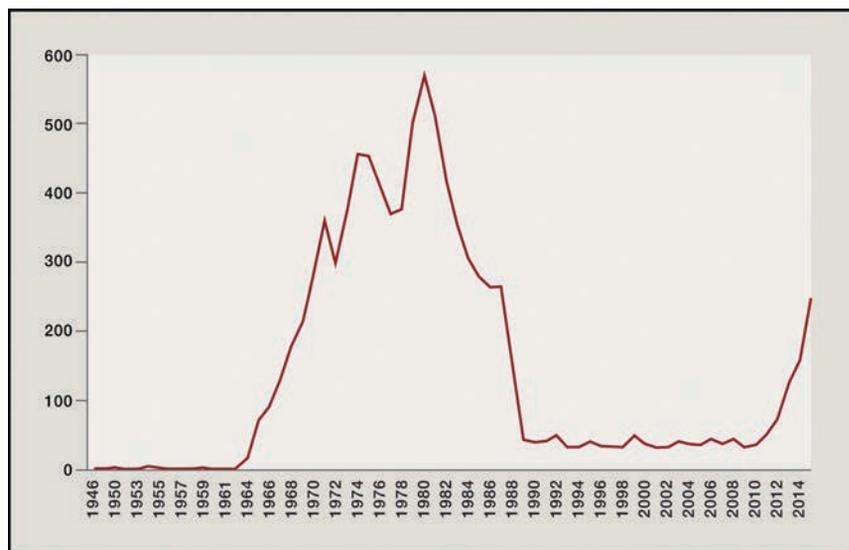
Organoids can be initiated from the two main types of stem cells: (1) pluripotent embryonic stem (ES) cells and their synthetic induced pluripotent stem (iPS) cell counterparts and (2) organ-restricted adult stem cells (aSCs). Both approaches exploit the seemingly infinite expansion potential of normal stem cells in culture. For ES and iPS cells, here collectively termed pluripotent stem cells or PSCs, this potential has been an essential prerequisite for their discovery. By contrast, aSCs—with the exception of Green’s skin cells—were long believed to be incapable of significant proliferation outside of the body. Yet, recent years have witnessed the rapid development of growth factor cocktails that mimic the various organ stem cell niches. When PSCs and aSCs are allowed to differentiate in culture, they display an uncanny capacity to self-organize into structures that reflect crucial aspects of the tissues to which they are fated.

## Organoids Derived from Pluripotent Stem Cells

Ever since pluripotent ES and iPS cell lines were established, scientists have applied insights from developmental biology to derive differentiated cell types from these stem cells (Chen et al., 2014; Cherry and Daley, 2012) (Figure 2). Yoshiki Sasai and his colleagues were the first to take this one step further by asking whether such an *in vitro* system could recapitulate some of the robust regulatory systems of organogenesis—in terms of not only cell differentiation, but also spatial patterning and morphogenesis. In a remarkable tour de force, they developed methods to generate brain structures, retina, and pituitary ‘in a dish’ (Eiraku and Sasai, 2012).

## Brain Organoids

The central nervous system derives from the neural ectoderm. Set up first as the neural plate, it is then shaped into the neural tube through folding and fusion. Morphogen gradients in this



**Figure 1. Citations to the Search Term “Organoids” in PubMed**

organoids containing representations of several different brain regions (Lancaster et al., 2013). Like the Sasai method, the approach starts with floating embryoid bodies, but growth factors are not added to drive particular brain region identities. Instead, the aggregates are embedded in a laminin-rich extracellular matrix secreted by the Engelbreth-Holm-Swarm tumor cell line (Matrigel). This allows outgrowth of large neuroepithelial buds, which spontaneously develop into various brain regions. Cerebral organoids can reach sizes of up to a few millimeters when grown in a spinning bioreactor.

A spectacular variety of brain regions,

tube establish a dorsal-ventral axis (roof, alar, basal, and floor plate) and a rostral-caudal axis (tel-, di-, mes-, and rhombencephalon and spinal cord). Neurons are generally generated from neural stem cells that reside near the ventricles. These stem cells initially increase their numbers through symmetric divisions. During neurogenesis, stem cells switch to asymmetric divisions to yield temporal waves of distinct self-renewing progenitors and differentiated cell types, such as neurons and intermediate progenitors, that migrate outward to generate region-specific stratified structures such as the medulla, the optic tectum, and the cerebral cortex (see Eiraku and Sasai, 2012; Lancaster and Knoblich, 2014).

Spontaneous neural differentiation occurs in ES culture in the absence of inhibitors of neural differentiation (such as BMP, Nodal, and Wnts), consistent with a neural-default state for ES cells. Based on this notion, Sasai and colleagues designed SFEBq: serum-free floating culture of embryoid body-like aggregates with quick reaggregation (Eiraku et al., 2008). In this culture system, ES cells isolated from growth factor-free 2D cultures are reaggregated in 96-well non-adhesive culture plates. The re-aggregates are kept in serum-free medium containing no or minimal growth factors for 7 days, after which they are replated in adhesion plates. When lumens form, the ES cells differentiate and polarize to form a continuous neuroectoderm-like epithelium that subsequently generates stratified cortical tissues containing cortical progenitors, deep cortical-layer neurons, superficial cortical-layer neurons, and so called Cajal-Retzius cells. In the absence of growth factors, the generated cortical tissue spontaneously adopts a rostral hypothalamic fate. Regional identity (e.g., olfactory bulb, rostral and caudal cortices, hem, and choroid plexus) can be selectively controlled by addition of specific patterning factors such as Fgf, Wnt, and BMP. It was thus shown that, under controlled conditions, ES cells can recapitulate some of the spatial and temporal events leading to the formation of layered neural structures in the brain.

Lancaster and Knoblich took this approach to a next level by generating cerebral organoids, or “mini-brains”: single neural

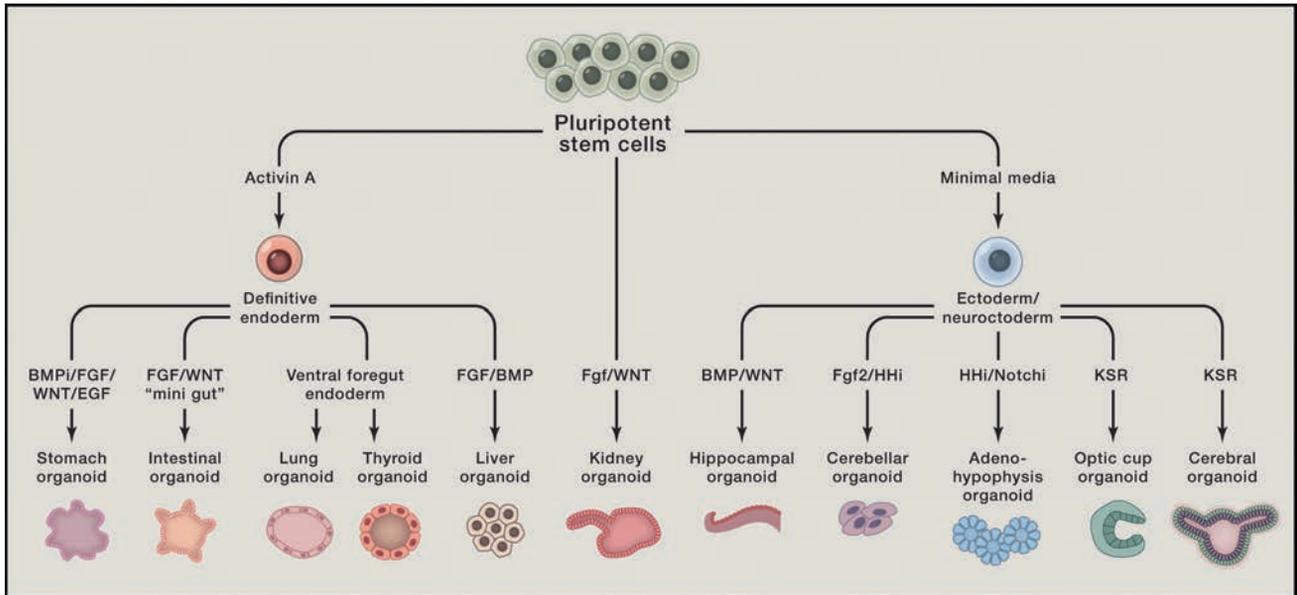
including retina, dorsal cortex, ventral forebrain, midbrain-hind-brain boundary, choroid plexus, and hippocampus, is observed in these cultures (Figure 4A).

A subsequent study applied single-cell RNA sequencing to compare gene expression programs of cells within cerebral organoids to those of fetal human neocortex development (Camp et al., 2015). It was thus found that gene expression programs of cortical cells in organoids are remarkably similar to those of the corresponding fetal tissue, underscoring that aspects of human cortical development can be studied in organoid culture.

### **Retinal Organoids**

The retina is of neuroectodermal origin and constitutes the light-receptive neural region of the eye. The optic vesicle forms as a pseudostratified, cystic outgrowth of the diencephalon. The front of the vesicle then moves inward to form the two-layered optic cup, consisting of the outer retinal pigment epithelium and the inner neural retina. The neural retina continues to stratify into layers of photoreceptors and supportive cell types, such as horizontal cells, bipolar cells, and amacrine cells.

To mimic this process in vitro (Eiraku et al., 2011), Sasai and colleagues again generated floating embryoid bodies from re-aggregated murine ES cells in growth factor-free medium to generate neuroectoderm. Matrigel, dissolved in the medium, allowed the formation of more rigid neuro-epithelial tissues. This resulted in the formation of buds of retinal primordial tissue resembling the optic vesicle. Isolated buds were then maintained in a medium supporting retinal tissue identity. The morphological tissue shape changes were reminiscent of the stepwise evagination and invagination of the optic cup in vivo. Retinal stratification with proper apical-basal polarity occurred, and markers of neural retina and pigment epithelium were expressed in a spatially correct manner. More recently, optic cup organoids were generated from human PSCs (Nakano et al., 2012). These human retinal organoids resembled their mouse counterparts but encouragingly also displayed human-specific features: they are larger in size (yet still small relative to the “real thing”), and the formed neural



**Figure 2. Schematic of the Various Organoids that Can Be Grown from PSCs and the Developmental Signals that Are Employed**  
Adapted from Lancaster and Knoblich, 2014.

retina grows into a thick multi-layered tissue containing both rods and cones, whereas cones were rarely observed in mouse organoid cultures.

#### **Adenohypophysis Organoids**

The adenohypophysis secretes multiple systemic hormones. During early mammalian development, its anlage originates as a placode in the non-neural head ectoderm near the anterior neural plate. The thickened placode invaginates and detaches from the oral ectoderm, forming a hollowed epithelial vesicle, Rathke's pouch. This process depends on poorly defined cross-signaling between ectoderm and developing neural tube. Sasai's group sought to recapitulate the inductive microenvironment of this morphogenetic field in order to promote the simultaneous generation of both tissues within the same aggregate of SFEBq-cultured ES cells. Three-fold larger cell aggregates were required, compared to the above protocols. Hedgehog and Notch antagonists were added to block neural fate in the outer layers and to allow the subsequent development of all major hormone-producing lineages, respectively. Under these conditions, ES cells differentiated into head ectoderm and hypothalamic neuroectoderm in adjacent layers within the aggregate. Rathke's-pouch-like structures arose at the interface of these two epithelia, and the various endocrine cell types were subsequently formed. Upon transplantation under the kidney capsule of hypophysectomized mice, the aggregates partially rescued systemic glucocorticoid level and prolonged survival of the mice.

#### **Cerebellar Organoids**

The initial phase of cerebellar development depends on the function of the isthmic organizer located at the midbrain-hind-brain boundary. Sasai and colleagues focused on the induction of isthmic development in an attempt to create functional Purkinje cells, the beautiful key output cells of the cerebellum. Again, they started from a mouse SFEBq culture. In order to produce

caudal brain structures, Fgf2 was added soon after initiation of the culture. To dorsalize the caudalized brain organoids, a Hedgehog inhibitor was added during the second week. These conditions recapitulated early cerebellar plate development, eventually leading to the formation of mature Purkinje cells (Muguruma et al., 2010). In a subsequent study, the investigators reported that the addition of Fgf19 and SDF1 to this protocol allows human ES cells to generate a polarized structure reminiscent of the first trimester cerebellum (Muguruma et al., 2015).

#### **Hippocampus**

The hippocampus develops from the dorsomedial telencephalon through a precursor structure termed the medial pallium. A final protocol developed by Sasai and coworkers involved the *in vitro* generation of a reliable source of hippocampal tissue from human ES cells (Sakaguchi et al., 2015). SFEBq served once again as the starting material. Stimulation by BMP and Wnt induced choroid plexus, the dorsomedial-most part of the telencephalon. Careful titration of BMP and Wnt exposure allowed the self-organization of tissue resembling the medial pallium, located adjacent to choroid plexus in the developing brain. Following long-term dissociation culture, granule neurons and pyramidal neurons were formed, both of which were electrically functional within connected networks.

In addition to these CNS organoids, protocols have also been developed to grow various endodermal organoids from PSCs. Formation of the endoderm germ layer during gastrulation requires Nodal signaling. Definitive endoderm presents as a 2D sheet of cells, which is subsequently patterned along the anterior-posterior axis and folded into a primitive gut tube, from which all endodermal organs arise. The foregut forms the anterior section of this tube and generates, e.g., the thyroid, lungs, stomach, liver, and pancreas. The mid- and hindgut develop into small intestine, colon, and rectum. Insights into the signals

that control these fate decisions *in vivo* can be exploited *in vitro*. Exposure to Nodal or its mimetic Activin directs differentiation of PSCs into definitive endoderm and serves as a common starting point of these protocols. Exposure to subsequent inductive signals can then induce the various endodermal organ identities (reviewed in Sinagoga and Wells, 2015).

### **Stomach Organoids**

The stomach develops from the posterior foregut. Wells and colleagues used activin treatment of human PSCs to generate definitive endoderm (McCracken et al., 2014). Subsequent addition of BMP inhibitors and of FGF and Wnt activators instructed the cells toward a foregut fate. When retinoic acid was applied, the organoids were specified toward a posterior foregut fate. Finally, high concentrations of EGF then converted these into human gastric organoids, progressing through molecular and morphogenetic stages that resembled those of the developing antrum of the mouse stomach. Organoids contained primitive gastric gland- and pit-like domains, proliferative zones with Lgr5<sup>+</sup> stem cells, mucous cells, and a host of gastric endocrine cells.

### **Lung and Thyroid Organoids**

The lung and the thyroid arise from Nkx2-1<sup>+</sup> progenitors in the developing ventral foregut endoderm. An initial study demonstrated the directed differentiation of primordial lung and thyroid progenitors from ESCs. The protocol involves activin-induced definitive endoderm and treatment with TGF $\beta$  /BMP inhibitors, followed by BMP/FGF stimulation, and results in a relatively pure population of progenitors that recapitulate early developmental milestones of lung/thyroid development (Longmire et al., 2012). This study has been the stepping stone for subsequent attempts to create organoids representing mature versions of the two organs.

A first description of the generation of lung organoids from iPS cells was reported by Rossant and colleagues and involved at its last stage air-liquid interphase culture. The protocol was applied to CFTR mutant iPS cells as a proof of concept for modeling cystic fibrosis (Wong et al., 2012). Snoeck and colleagues designed an improved four-stage, 50-day protocol (Huang et al., 2014). First, definitive endoderm was induced by Activin A. Subsequently, anterior foregut endoderm was induced by sequential inhibition of BMP, TGF- $\beta$ , and Wnt signaling. The cells were then ventralized by Wnt, BMP, FGF, and RA to obtain lung and airway progenitors. Finally, epithelial cell types (basal, goblet, Clara, ciliated, type I and type II alveolar epithelial cells) were matured using Wnt, FGF, c-AMP, and glucocorticoids. Spence and colleagues (Dye et al., 2015) similarly started from Activin-treated human PSCs but then followed a slightly different trajectory. Subsequent addition of TGF $\beta$ /BMP inhibitors, FGF4, and Wnt activators instructed the cells toward an anterior foregut fate. When the Hedgehog pathway was simultaneously activated, organoids were ventrally specified toward a lung fate. Upon embedding in Matrigel and prolonged exposure to Fgf10, mature lung organoids arose. The cultures could be maintained for several months and resembled proximal airways, containing basal cells, ciliated cells, and Clara cells. The endodermal airway tissues were found to be often surrounded by smooth muscle actin (SMA)-positive mesenchymal cells. Early markers of the distal (alveolar) airways were expressed early in culture but were lost later.

Initial attempts to create thyroid organoids involved forced expression of the lineage-specific transcription factors NKX2.1 and PAX8 and encouragingly resulted in the formation of mouse and human thyroid follicles *in vitro* and upon transplantation (Antonica et al., 2012; Ma et al., 2015). Kotton and colleagues applied an improved version of their “all soluble factor” protocol (Longmire et al., 2012), followed by sequential treatment by BMP4/FGF2 and induced maturation by 3D plating in the presence of thyroid-stimulating hormone in Matrigel. The resulting fully mature murine thyroid follicular organoids secreted thyroid hormones *in vivo* upon transplantation and rescued hypothyroid mice. The same protocol allowed the derivation of human thyroid progenitors from iPS cells (Kurmann et al., 2015).

### **Small Intestinal Organoids**

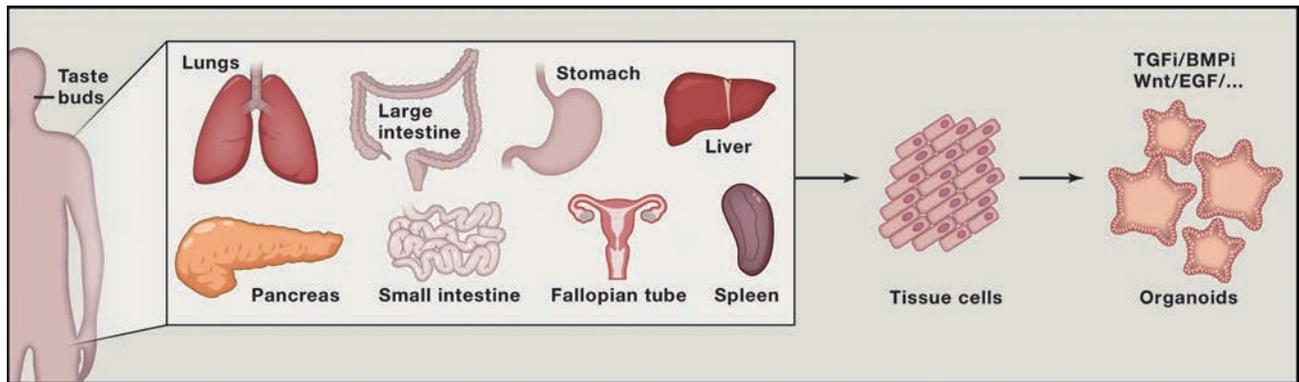
Wnt and FGF signals are known to specify definitive endoderm toward mid-/hindgut fates (“posteriorization”). To generate intestinal organoids (McCracken et al., 2014; Spence et al., 2011), Activin-treated human PSCs were cultured with FGF4 and WNT3a. Mid/hindgut spheroids budded off from the 2D monolayer epithelium and were further cultured in Matrigel along with a pro-intestinal growth factor cocktail, defined previously for expansion of adult crypt cultures (Sato et al., 2009; see below). The organoids expanded over 1–3 months to give rise to a polarized intestinal epithelium patterned into villus-like structures and crypt-like proliferative zones and containing all major epithelial cell types. Intestinal mesenchyme (presumably derived from mesodermal remnants after endoderm induction) surrounded the epithelial structures and consisted of myofibroblasts and smooth muscle cells (McCracken et al., 2014; Spence et al., 2011). Transplantation of these organoids into immunodeficient mice yielded human epithelium and laminated human mesenchyme, supported by mouse vasculature. The transplanted tissue was functional, as shown by permeability and peptide uptake tests (Watson et al., 2014).

### **Liver Organoids**

During early hepatogenesis, progenitor cells delaminate from the foregut endoderm to form a condensed tissue mass termed the liver bud, which is vascularized soon thereafter. Taniguchi and co-workers exploited cross-signaling between endodermal epithelial, mesenchymal, and endothelial progenitors in an effort to generate tissues reminiscent of the human liver bud. Human PSCs were induced into hepatic endodermal cells in 2D culture (a protocol involving activin treatment, followed by bFGF/BMP4). The human PSC-derived hepatic cells were mixed with mesenchymal stem cells and endothelial cells. Plated at high density on a layer of Matrigel, 3D aggregates spontaneously formed (Takebe et al., 2013). These liver bud-like aggregates contained blood vessels that, upon transplantation into mice, connected to the host vessels within 48 hr. Liver-specific functions such as protein production and human-specific drug metabolism became evident over time. Furthermore, mesenteric liver bud transplantation rescued recipient mice from drug-induced lethal liver failure.

### **The Mesodermal Kidney**

The kidney, with its more than 20 specialized cell types, exhibits the highest architectural complexity of all organs outside of the CNS. The adult kidney, or metanephric kidney, arises from the



**Figure 3. Schematic of the Various Regions of the Body that Can Be Cultured as aSC-Derived Organoids**

posterior end of the embryonic intermediate mesoderm, which in turn derives from the primitive streak (presomitic mesoderm). The intermediate mesoderm generates the two key kidney progenitor populations: the ureteric epithelium and the metanephric mesenchyme. Through reciprocal interactions, these form the collecting ducts and nephrons (i.e., the epithelia of glomeruli and proximal and distal renal tubules), respectively. Until recently, the complex spatial and temporal control of organogenesis has stood in the way of a detailed molecular understanding of specification of individual cell types. Despite this, rapid progress has been made in establishing protocols for differentiation of human PSCs into virtually complete “mini-kidneys.”

First, in 2013, it was shown how to induce intermediate mesoderm from PSCs under defined media conditions (Mae et al., 2013). One of the renal precursor tissues that derives from the intermediate mesoderm, the ureteric epithelium, can be generated from human PSCs in 2D via a similar mesodermal specification step (Xia et al., 2013). Upon aggregation with dissociated mouse embryonic kidney, these progenitors self-organize into 3D ureteric bud structures. The second renal precursor tissue, the metanephric mesenchyme, can be created from human and mouse embryoid bodies through sequential exposure to defined soluble factors. Coculturing of the resulting metanephric mesenchyme with spinal cord tissue, a nephric inducer, produces well-organized nephric tubules and nascent glomeruli (Taguchi et al., 2014).

Little and colleagues managed to balance the two divergent commitment paths to produce both principal lineages of the kidney simultaneously (Takasato et al., 2014). Their original protocol involves the application of Activin A and Bmp4 to human PSCs cultured in 2D to generate primitive streak identity. Fgf9 drives these cells toward an intermediate mesoderm identity, after which they spontaneously develop further into ureteric bud and metanephric mesenchyme. The cells display 3D morphologies when grown at low density in 2D or when cocultured with mouse kidney reaggregates. In both cases, structures resembling ureteric epithelium and proximal tubules appear. In a spectacular follow-up study, the protocol was further refined and simplified: human PSCs are cultured in 2D in the presence of Wnt signals for 4 days followed by 3-day exposure to Fgf9. After

this, the cells are pelleted and cultured as 3D organoids for up to an additional 3 weeks. Numbers of nephrons are strongly increased upon a brief (1 hr) exposure to a Wnt agonist at the start of organoid culture. A complex multicellular kidney organoid results that contains fully segmented nephrons and is surrounded by endothelia and renal interstitium (Figure 4B). Kidney organoids may contain >500 nephrons with defined glomeruli comprising a Bowman’s capsule with podocytes and connected to proximal tubules. Occasionally, glomeruli show evidence of endothelial invasion.

While remarkably complete, further improvements of the protocol will focus on tubular functional maturation, more extensive glomerular vascularization, and the formation of a contiguous collecting ductal tree “with a single exit path for urine” (Takasato et al., 2015).

### Organoids Derived from Adult Stem Cells

While PSC-based organoids exploit developmental processes for their establishment, aSCs can be coerced to form organoids by creating conditions that mimic the stem cell niche environment during physiological tissue self-renewal or during damage repair (Figure 3). As first described for gut stem cells (Korinek et al., 1998), the Wnt pathway has emerged as the major driver of epithelial aSCs (Clevers et al., 2014). Lgr5 (a receptor for the secreted Wnt-amplifying R-spondins and itself encoded by a Wnt target gene) marks active aSCs in many, if not all, epithelia. It is not surprising that Wnt activators (Wnt3A, R-spondins, or the small molecule GSK3 inhibitor CHIR) are key components of most aSC culture protocols and that Lgr5<sup>+</sup> stem cells invariably appear in such cultures. Below, I discuss the establishment of feeder layer/serum-free, fully defined 3D culture conditions for a rapidly growing list of epithelial tissues.

### Small Intestine and Colon

The small intestinal epithelium displays an extremely short turnover time of ~5 days. Actively proliferating Lgr5<sup>+</sup> intestinal stem cells reside at the crypt base (Barker et al., 2007). Their rapidly dividing, transit-amplifying (TA) daughter cells occupy the remainder of the crypts and, upon differentiation, move onto the flanks of the villi to eventually die at the villus tips. Differentiated cell types include absorptive enterocytes, multiple secretory cell types (Paneth cells, goblet cells, enteroendocrine

cells, and tuft cells), and the M cells that cover Peyer's patches (Clevers, 2013).

Crypt stem cells are tightly controlled by four signaling pathways. Wnt constitutes the key pathway to maintain stem cell fate and drive proliferation of stem and TA cells. Notch helps to maintain the undifferentiated state of proliferative stem and TA cells: when Notch signaling is blocked, the cells instantly differentiate into goblet cells. Epidermal growth factor (EGF) signals exert strong mitogenic effects on stem and TA cells. And finally, BMP signals are active in the villus compartment, and their inhibition is crucial to create a crypt-permissive environment (Clevers, 2013).

Encouraged by the observation that Lgr5 crypt stem cells can go through thousands of cell divisions *in vivo*, we established a culture system that allows growth of epithelial organoids ("mini-guts") from a single Lgr5 stem cell (Sato et al., 2009). Whole crypts or single Lgr5 stem cells are suspended in Matrigel and are cultured in serum-free medium supplemented with three recombinant proteins: R-spondin-1 (a Wnt signal amplifier and ligand of Lgr5), EGF, and the BMP inhibitor Noggin. For colon crypt culture, Wnt3a is additionally required because colon epithelium itself makes little, if any, Wnt. The organoids strictly consist of a simple highly polarized epithelium, tightly closing off a central lumen. Crypt-like structures project outward. The basal side of the cells is oriented toward the surrounding Matrigel. Enterocyte brush borders form the luminal surface, while secretion by Paneth and goblet cells occurs toward the lumen. All cell types of the epithelium are represented at normal ratios (Grün et al., 2015; Sato et al., 2009). The organoids can be passaged weekly at a 1:5 ratio for years and are remarkably stable, both genetically and phenotypically (Sato and Clevers, 2013).

Addition of Wnt3A to the combination of growth factors allowed seemingly indefinite growth of mouse colon organoids. Addition of nicotinamide, along with a small molecule inhibitor of Alk and an inhibitor of p38, was required for long-term culture of human small intestine and colon organoids (Jung et al., 2011; Sato et al., 2011). Given that Lgr5 protein expression is vanishingly low, other stem cell markers have been explored to initiate intestinal organoid cultures, including CD24 (von Furstenberg et al., 2011), EphB2 (Jung et al., 2011), and CD166+/GRP78– (Wang et al., 2013). As proof of stability upon culture, a large batch of organoids was grown from a single Lgr5 colon stem cell and transplanted *per anum* into multiple mice with experimental colitis. The organoids readily integrated as functional epithelial patches that were indiscernible from the surrounding host epithelium (Yui et al., 2012).

In a different approach, fragments of neonatal mouse intestine containing epithelial and mesenchymal elements were grown in serum-containing medium (without specific growth factors) in collagen with air-liquid interface. The expanding cystic structures consisted of a simple epithelium in which all cell types were discernible. The structures were surrounded by myofibroblasts and were responsive to R-spondin and to Notch inhibition (see above) (Ootani et al., 2009).

The mini-gut culture system has since been adapted for the generation of organoids representing the epithelial compartments of a series of mouse and human tissues of ecto-, meso-

and endodermal origin. The essential components appear to be: (1) a potent source of Wnt, (2) a potent activator of tyrosine kinase receptor signaling like EGF, (3) inhibition of BMP/Tgf $\beta$  signals, and (4) Matrigel. It is not essential to start from purified Lgr5<sup>+</sup> aSCs. Small fragments of primary tissue serve well as starting material, possibly due to the fact that the culture conditions mimic a damage response, which in many tissues can recruit committed cells back to a stem cell state (Clevers, 2015)

### **Stomach**

Rapidly proliferating Lgr5 stem cells are located at the base of pyloric glands of the adult mouse stomach. With slight modifications to the mini-gut culture system, single Lgr5 cells efficiently generated long-term, continuously expanding organoids closely resembling mature pyloric epithelium (Barker et al., 2010). At the base of glands of the gastric corpus, Troy marks specialized Chief cells. In a remarkable example of cellular plasticity, these Chief cells spontaneously dedifferentiate to act as multipotent epithelial stem cells *in vivo*, particularly upon damage. Single Troy<sup>+</sup> chief cells can be cultured to generate long-lived gastric organoids, containing the various cell types of corpus glands (Stange et al., 2013). Very similar conditions have allowed long-term culturing of human stomach organoids that maintain many characteristics of the original tissue (Bartfeld et al., 2015).

### **Liver and Pancreas**

The two cell types of the liver proper (the hepatocyte and the bile duct cell) turn over at a slow, pedestrian rate. In agreement, Lgr5 is not expressed at appreciable levels in the healthy adult mouse liver. Yet, we observed that small Lgr5<sup>+</sup> cells appear near bile ducts upon toxic damage, coinciding with robust Wnt pathway activation. These damage-induced Lgr5<sup>+</sup> cells generated hepatocytes and bile ducts *in vivo*. When cultured in a modified version of the mini-gut medium, single Lgr5<sup>+</sup> cells could be clonally expanded as organoids, consisting largely of progenitor cells expressing early bile duct and hepatocyte markers. Removal of mitotic stimuli and simultaneous inhibition of Notch signals led to hepatocyte lineage differentiation. Upon transplantation, these organoids matured into functional hepatocytes (Huch et al., 2013b). In a follow-up study, we defined conditions for long-term expansion of adult bipotent progenitor cells from human liver. Somewhat surprisingly, one-third of all mature bile duct cells could initiate clonal liver organoid growth. Deep sequencing of clonal organoids derived at different intervals of culture revealed a highly stable genome at the structural level, while single base changes occurred at very low rates. Again, the cells could be converted into functional hepatocytes *in vitro* and upon transplantation into mice (Huch et al., 2015). The same protocols allowed long-term expansion of canine liver progenitor cells that could be differentiated toward functional hepatocytes (Nantasanti et al., 2015).

Like their liver counterparts, the exocrine/acinar, ductal, and endocrine cell types of the adult pancreas turn over very slowly. Wnt signaling is inactive and Lgr5 is not expressed under physiological conditions, yet the Wnt pathway is robustly activated upon injury, concomitant with induced Lgr5 expression in regenerating pancreatic ducts. Under modified mini-gut conditions, single isolated duct cells could be cultured long-term as pancreatic progenitor organoids. Clonal pancreas organoids differentiated along ductal and endocrine lineages when grafted

in vivo in a developing pancreas, indicative of bipotentiality (Huch et al., 2013a). Similar observations were made for human pancreatic organoids (Boj et al., 2015). Grompe and colleagues addressed the identity of the organoid-initiating epithelial cell from mouse pancreas and liver using a set of cell surface markers and found that the transcriptomes of the two populations overlapped extensively. Pancreatic organoid cells had the unexpected capacity to generate hepatocyte-like cells upon transplantation in a mouse liver damage model, indicative of the close kinship of these two progenitor populations (Dorrell et al., 2014).

### **Prostate**

The pseudostratified prostate epithelium consists of basal and luminal cells. We developed a mini-gut-based 3D culture protocol that supports long-term expansion of primary mouse and human prostate organoids, composed of fully differentiated basal and luminal cells. Single human luminal as well as basal cells gave rise to organoids, yet luminal-cell-derived organoids more closely resembled prostate glands. Stimulation with R-spondin/Wnt was not essential for continued growth of the organoids but strongly induced luminal cells, leading to a prostate-like pseudostratified structure of the organoids. Long-term cultured organoids were genetically stable and reconstitute prostate glands in recombination assays (Karthaus et al., 2014). Independently, Shen and colleagues developed a Matrigel/EGF-based culture system supplemented with androgens and reported very similar observations (Chua et al., 2014).

### **Mammary Gland**

This pseudostratified epithelium consists of two major cell lineages. The inner (luminal) cells secrete milk, while the contractile outer layer of myoepithelial (basal) cells ejects the milk. No long-term organoid protocol has been reported yet. However, encouragingly, freshly isolated human mammary epithelial cells have been cultured for two to three passages in floating collagen gels in the presence of a Rho-associated kinase inhibitor to form branching ducts with alveoli at their tips. Basal and luminal markers were expressed at correct positions, and the ducts displayed contractility. Thus, the organoids resembled terminal ductal-lobular units, the functional units of the mammary gland (Linnemann et al., 2015). Since Wnt signals and Lgr5 have been implied in mammary stem cell biology (Plaks et al., 2013; Rios et al., 2014), it will be of interest to test the effects of the addition of Wnt/R-spondin to these cultures.

### **Fallopian Tube**

The fallopian tube of the uterus is lined by a simple columnar epithelium in which secretory cells produce tubular fluid, while ciliated cells facilitate transport of gametes. Since the epithelium is exposed to cyclical hormonal changes, self-renewal mechanisms are of critical importance for its integrity. Notably, recent evidence has indicated that the fallopian tube epithelium is the tissue of origin for ovarian cancer. Based on the mini-gut-protocol, long-term, stable 3D organoid cultures were established from human fallopian tubes. Single epithelial stem cells gave rise to clonal organoids containing both ciliated and secretory cells, thus establishing an experimental system for the study of the human fallopian tube epithelium in health and disease (Kessler et al., 2015).

### **Taste Buds**

Previous studies had shown that Lgr5 marks adult stem cells in the rapidly self-renewing taste buds of the tongue. Using the original mini-gut culture protocol, single isolated stem cells from taste tissue generated continuously expanding 3D organoids, which phenotypically contained mature taste receptor cells (Ren et al., 2014). To assay functionality of these cells, cultured organoids were reseeded in 2D onto laminin-coated coverslips in the same culture medium. By calcium imaging assays, dose-dependent responses to tastants were readily documented, demonstrating that functional taste cells can be generated ex vivo from single Lgr5+ taste bud stem cells. Moreover, it could be concluded that single stem cells generate all taste cell types and that the formation of taste cells does not require innervation.

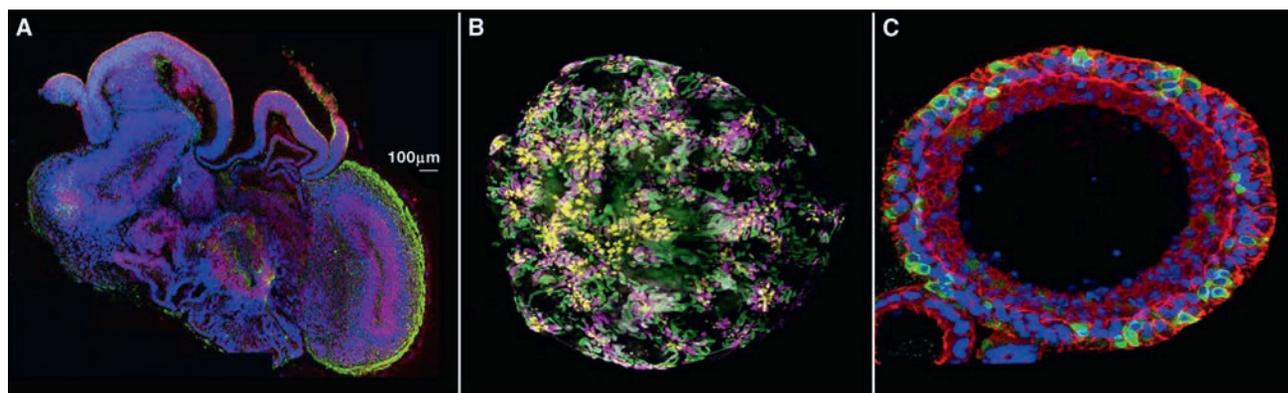
### **Lung**

Hogan and colleagues reported an early bronchiolar lung organoid culture protocol, involving Matrigel supplemented with EGF, e.g. Single basal cells isolated from the trachea grew into “tracheospheres” consisting of a pseudostratified epithelium with basal cells and ciliated luminal cells. These organoids could be passaged at least twice. No mature Clara-, neuroendocrine- or mucus-producing cells were observed (Rock et al., 2009). In a later study, this clonal 3D organoid assay was used to screen for factors controlling generation of ciliated versus secretory cells from basal cells. It was thus found that IL-6 treatment resulted in the formation of multiciliated cells at the expense of secretory and basal cells (Tadokoro et al., 2014). Figure 4C depicts a human airway organoid.

Organoids representing the distal airways (“alveolospheres”) have been more recently established. The alveoli consist of gas-exchanging type I and surfactant-secreting type II cells. While both cell types originally derive from a common progenitor, it appears that, later in life, a rare self-renewing type II cell acts as the stem cell to regenerate the alveolar epithelium. Indeed, sorted type II cells remained proliferative in short-term culture and could generate type I cells (Desai et al., 2014; Treutlein et al., 2014). Alternative culture conditions allowed establishment of mouse and human alveolospheres from single type I as well as type 2 alveolar cells, containing both cell types in the same organoid. Having said this, these alveolosphere culture conditions are as yet not fully defined, requiring co-culture with non-epithelial cells (e.g., mouse lung fibroblasts) (Barkauskas et al., 2013; Jain et al., 2015).

### **Salivary Gland**

Coppes and colleagues have exploited organoid culture to expand single salivary gland cells in vitro into distinct lobular or ductal/lobular organoids, containing some salivary gland lineages. The original short-term culture technology depended on FGF, EGF, and Matrigel. The cultured cells were able to efficiently restore radiation-damaged salivary gland function in transplanted mice (Nanduri et al., 2014). In a follow-up study, robust Wnt pathway activation through the addition of Wnt3A and R-spondin allowed long-term expansion of the organoids, containing all differentiated salivary gland cell types. Transplantation of these cells into submandibular glands of irradiated mice robustly restored saliva secretion and increased the number of functional acini in vivo (Maimets et al., 2016). Since post-radiation hyposalivation often leads to irreversible and untreatable



**Figure 4. A “Mini-Brain” Generated from PSCs**

(A) A complex morphology with heterogeneous regions containing neural progenitors (SOX2, red) and neurons (TUJ1, green) is apparent (Lancaster et al., 2013). Courtesy of Madeline Lancaster.

(B) Immunofluorescent image of an entire kidney organoid grown from PSCs with patterned nephrons. Podocytes of the forming glomeruli (NPHS1, yellow), early proximal tubules (lotus tetragonolobus lectin, pink), and distal tubules/collecting ducts (E-Cadherin, green). Courtesy of Melissa Little.

(C) 3D reconstruction of the midsection of a human aSC-derived lung organoid stained for intermediate filaments of basal cells (green), the actin cytoskeleton (red), and nuclei (blue) and imaged by confocal microscopy (N. Sachs and H.C., unpublished data).

xerostomia, this condition may present an early opportunity for the development of organoid technology-based cell therapy.

### Esophagus

All examples above represent simple or two-layered epithelia. Lagasse and colleagues showed that the keratinizing stratified epithelium of the esophagus can also be cultured as organoids in “mini-gut” medium (DeWard et al., 2014). Basal cells in the mouse esophagus represent a heterogeneous population of proliferative cells. When plated as single cells, these give rise to organoids that were morphologically similar to normal esophageal tissue, with small basal-like cells in contact with the extracellular matrix, large flat suprabasal-like cells in the interior, and hardened keratinized material in the center. Expression of specific markers for each of these cell types confirmed the correct layering of the organoid walls. It will be of interest to determine whether basal cells from other squamous epithelia (epidermis, vagina) will also be amenable to organoid culture.

### Applications of Organoid Technology

Both PCS- and aSC-based organoids can be initiated from single cells and cultured long-term and are amenable to essentially all cell-biological and molecular analyses that have been developed for “traditional” cell lines. As such, they provide a new window—between cell lines and in vivo studies—to studying basic gene functions and cellular processes. In addition to this, organoid technology also holds great promise for translational research. Below, I give some examples of its translational applications.

#### Infectious Disease

Since organoids—unlike cell lines—ideally represent all cellular components of a given organ, they are theoretically well suited for infectious disease studies, particularly of pathogens that are restricted to man and are dependent on specialized cell types. In an illustrative application, iPS-derived lung organoids were generated from an otherwise healthy child who suffered life-threatening influenza and carried null alleles in the interferon

regulatory factor 7 gene. These organoids produced less type I interferon and displayed increased influenza virus replication (Ciancanelli et al., 2015). In another example, human stomach organoids, grown from PSCs or aSCs, can be productively infected by *Helicobacter pylori* (Bartfeld et al., 2015; McCracken et al., 2014).

As a striking example, Qian et al. developed a miniaturized spinning bioreactor to generate forebrain-specific organoids from human iPSCs, following the Lancaster/Knoblich protocol. These organoids recapitulate many features of cortical development, including the formation of a distinct human-specific outer radial glia cell layer. Infection of these developing forebrain organoids with Zika virus (ZIKV) resulted in the preferential infection of neural progenitors, resulting in cell death, decreased proliferation, and a reduced neuronal cell-layer volume, thus modeling ZIKV-associated microcephaly. The authors propose this as a versatile experimental for mechanistic studies as well as for testing of potential ZIKV antiviral drugs (Qian et al., 2016).

#### Hereditary Disease

Organoids can be used to study and model organ-specific monogenic hereditary diseases. Knoblich and colleagues identified a patient with a mutation in the CDK5RAP2 and severe microcephaly. The corresponding iPS cells made significant smaller “mini-brains,” containing only occasional neuroepithelial regions with signs of remature neural differentiation, a phenotype that could be rescued by reintroducing the CDK5RAP2 protein (Lancaster et al., 2013).

Cystic fibrosis (CF) is caused by a spectrum of mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) chloride channel that is normally expressed in epithelial cells of many organs. Mirroring the in vivo situation, surface expression of CFTR was absent in iPS-derived lung organoids from CF patients but could be restored by treatment with a (then still experimental) small molecule that corrects some of the common CF-processing mutations (Wong et al., 2012). Dekkers and

colleagues derived intestinal organoids from rectal biopsies of a series of CF patients. Forskolin induces a robust swelling of wild-type organoids due to fluid transport to the organoid lumen. This swelling response is absent in CF organoids yet can be restored for the common, temperature-sensitive CFTR-F508del mutant by culturing at 27°C and also by the addition of experimental CFTR corrector compounds (Dekkers et al., 2013).

Independently, the Verma lab generated iPS cells from CF patients and corrected the mutation by CRISPR/Cas9. The corrected iPS cells were subsequently converted to mature airway epithelial cells demonstrating recovery of normal CFTR function (Firth et al., 2015).

Liver organoids from alpha 1-antitrypsin deficiency patients reproduced the deleterious effects of the mutant protein precipitates in hepatocytes, while the absence of mature biliary cells in liver organoids from an Alagille syndrome patient mirrored the in vivo biliary tree abnormalities (Huch et al., 2015). Liver organoids from dogs deficient in the copper-transporter COMMD1 mimicked the disease by accumulating toxic levels of copper, which could be salvaged by re-expression of wild-type COMMD1 protein (Nantasanti et al., 2015).

### Toxicology

The possibility to grow human organoids representative of the main targets for drug-related toxicity (gut, liver, kidney) opens up theoretical avenues to complement animal-based toxicology with assays performed directly on these vulnerable human tissues. In one such example, Little and colleagues have utilized human kidney organoids to illustrate that cisplatin acts as a nephrotoxicant (Takasato et al., 2015).

### Cancer

Once culturing protocols for human aSC-based organoids were established, we have shown the feasibility of growing organoids from primary colon, prostate, and pancreatic cancers (Boj et al., 2015; Gao et al., 2014; Sato et al., 2011; van de Wetering et al., 2015). These cancer organoids provide the unique opportunity for functional testing (e.g., for drug sensitivity) and for correlating such data with the genetic make-up of individual tumors.

Cancer can also be modeled in organoids derived from wild-type stem cells. Kuo and colleagues probed the metastatic potential of TGFBR2 loss in murine stomach organoids by its shRNA knockdown within *Cdh1*<sup>-/-</sup>; *Tp53*<sup>-/-</sup> organoids. This resulted in invasive phenotypes in vitro and in robust metastasis in vivo (Nadauld et al., 2014). Using similar shRNA-based strategies, the same authors created combinatorial Apc, p53, KrasG12D, and Smad4 mutations in wild-type murine colon organoids and observed progressive transformation to an invasive adenocarcinoma-like histology in vivo, recapitulating the multi-hit model of colorectal cancer (CRC) (Li et al., 2014). In a different approach, Huang et al. established a three-step culture method using human PSCs to generate 3D structures closely resembling human fetal exocrine pancreas (Huang et al., 2015). Expression of mutant *KRAS* and/or *TP53* in these early pancreas organoids induced abnormal ductal architecture and nuclear morphology consistent with neoplastic transformation in culture and in vivo. 17 of 20 primary human pancreas cancers could be propagated under the “third-step” conditions (i.e., EGF and FGF). Thus,

more than one culture condition can be applied to efficiently grow human pancreas cancer organoids (Boj et al., 2015; Huang et al., 2015).

In a marriage between organoid and CRISPR/CAS9 technologies, two independent studies have modeled the “adenoma-carcinoma sequence” by introducing four sequential mutations into human colon organoid stem cells. Xenotransplantation revealed the progressive transformation of the wild-type stem cells into adenocarcinomas (Drost et al., 2015; Matano et al., 2015). These approaches may allow rapid modeling of novel (combinations of) gene mutations, as found in the ongoing genome- and worldwide sequencing efforts on large solid tumor panels.

### Personalized Medicine

In principle, the aSC-based organoid technology allows rapid ex vivo testing of drug responses on the affected tissue of individual patients. As a first example, the colon organoid-based CF test (Dekkers et al., 2013) can be read out in weeks after biopsy. The approach has already been applied for identification and successful treatment of patients with very rare CFTR mutations, who otherwise have no access to the recently introduced CF drugs (Dekkers, 2016). The feasibility of culturing various solid tumors directly from the patient in the form of tumor organoids (see above) holds a similar promise, yet the applicability of such an approach is less clear than in the case of the “single genetic lesion” CF organoids. Tumor organoids grow with unpredictable and often slower kinetics when compared to wild-type organoids, and—like the original tumors—display a heterogeneous genetic make-up. Ongoing trials will reveal the validity and applicability of tumor organoids in the assessment of drug response at the level of the individual patient.

### Regenerative Medicine and Gene Therapy

Proof-of-concept studies have demonstrated the feasibility of expanding organoids from (single) aSCs followed by safe transplantation into animals. This was first done for murine small intestine (Fordham et al., 2013; Yui et al., 2012). When small intestinal organoids are transplanted to colon, they retain original small intestinal features like villus formation and the presence of Paneth cells, indicative of the phenotypic stability of cultured aSC organoids (Fukuda et al., 2014). Of note, it may be advantageous to modify the culture conditions to selectively expand stem cells at the cost of differentiated cells (Wang et al., 2015; Yin et al., 2014) prior to transplantation. We have used CRISPR/Cas9 genome editing to correct the CFTR locus by homologous recombination in cultured intestinal stem cells of CF patients. The corrected allele was fully functional, as demonstrated in clonally expanded organoids (Schwank et al., 2013). This approach can presumably be used for gene correction in any clonally expandable cell population derived from monogenic hereditary disease patients.

### Conclusions

In this Review, I have attempted to describe the state of the art of the explosively developing field of PSC- and aSC-based organoids. The current versions of organoids have clear limitations, e.g., innervation, blood vessels, and immune cells are absent, and as a consequence, disease processes are only partially recapitulated. Yet it is anticipated that the potent self-organizing

properties of organoids may extend beyond their current boundaries and allow the proper incorporation of additional cellular (or microbial) elements. From a basic science perspective, PSC-based organoids will by their very nature play a key role in understanding the developmental biology of organs and will thus complement the long tradition of *in vivo* studies in this field. From the same perspective, aSC-based organoids provide basic insights into the processes that allow aSCs to maintain and repair established tissues. Yet, because of the ease of production and the close resemblance to human organs in health and disease, organoids hold great appeal for translational research and invite an almost immediate application into the clinic.

## ACKNOWLEDGMENTS

Thanks for comments to the text to Melissa Little, Hans Snoeck, Juergen Knoblich, and Esther Verheyen and to Janny van Eldik for secretarial help. Apologies to those scientists whose work could not be cited due to space restrictions. Written as visiting professor at Memorial Sloan Kettering Cancer Center, New York, NY. H.C. is named as inventor on several patents related to Lgr5 stem-cell-based organoid technology.

## REFERENCES

- Antonica, F., Kasprzyk, D.F., Opitz, R., Iacovino, M., Liao, X.H., Dumitrescu, A.M., Refetoff, S., Peremans, K., Manto, M., Kyba, M., and Costagliola, S. (2012). Generation of functional thyroid from embryonic stem cells. *Nature* 491, 66–71.
- Barkauskas, C.E., Crouse, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., and Hogan, B.L. (2013). Type 2 alveolar cells are stem cells in adult lung. *J. Clin. Invest.* 123, 3025–3036.
- Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., and Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* 449, 1003–1007.
- Barker, N., Huch, M., Kujala, P., van de Wetering, M., Snippert, H.J., van Es, J.H., Sato, T., Stange, D.E., Begthel, H., van den Born, M., et al. (2010). *Lgr5*(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units *in vitro*. *Cell Stem Cell* 6, 25–36.
- Bartfeld, S., Bayram, T., van de Wetering, M., Huch, M., Begthel, H., Kujala, P., Vries, R., Peters, P.J., and Clevers, H. (2015). *In vitro* expansion of human gastric epithelial stem cells and their responses to bacterial infection. *Gastroenterology* 148, 126–136.e6.
- Boj, S.F., Hwang, C.I., Baker, L.A., Chio, I.I., Engle, D.D., Corbo, V., Jager, M., Ponz-Sarvisé, M., Tiriác, H., Spector, M.S., et al. (2015). Organoid models of human and mouse ductal pancreatic cancer. *Cell* 160, 324–338.
- Camp, J.G., Badsha, F., Florio, M., Kanton, S., Gerber, T., Wilsch-Bräuninger, M., Lewitus, E., Sykes, A., Hevers, W., Lancaster, M., et al. (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proc. Natl. Acad. Sci. USA* 112, 15672–15677.
- Chen, K.G., Mallon, B.S., McKay, R.D., and Robey, P.G. (2014). Human pluripotent stem cell culture: considerations for maintenance, expansion, and therapeutics. *Cell Stem Cell* 14, 13–26.
- Cherry, A.B., and Daley, G.Q. (2012). Reprogramming cellular identity for regenerative medicine. *Cell* 148, 1110–1122.
- Chua, C.W., Shibata, M., Lei, M., Toivanen, R., Barlow, L.J., Bergren, S.K., Badani, K.K., McKiernan, J.M., Benson, M.C., Hibshoosh, H., et al. (2014). Single luminal epithelial progenitors can generate prostate organoids in culture. *Nat. Cell Biol.* 16, 951–961.
- Ciancanelli, M.J., Huang, S.X., Luthra, P., Garner, H., Itan, Y., Volpi, S., Lafaille, F.G., Trouillet, C., Schmolke, M., Albrecht, R.A., et al. (2015). Infectious dis-
- ease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. *Science* 348, 448–453.
- Clevers, H. (2013). The intestinal crypt, a prototype stem cell compartment. *Cell* 154, 274–284.
- Clevers, H. (2015). STEM CELLS. What is an adult stem cell? *Science* 350, 1319–1320.
- Clevers, H., Loh, K.M., and Nusse, R. (2014). Stem cell signaling. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science* 346, 1248012.
- Dekkers, F.e.a. (2016). Identifying potential clinical responders to CFTR-modulating drugs using rectal cystic fibrosis organoids. *Sci. Transl. Med.*, in press.
- Dekkers, J.F., Wiegerinck, C.L., de Jonge, H.R., Bronsveld, I., Janssens, H.M., de Winter-de Groot, K.M., Brandsma, A.M., de Jong, N.W., Bijvelds, M.J., Scholte, B.J., et al. (2013). A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat. Med.* 19, 939–945.
- Desai, T.J., Brownfield, D.G., and Krasnow, M.A. (2014). Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* 507, 190–194.
- DeWard, A.D., Cramer, J., and Lagasse, E. (2014). Cellular heterogeneity in the mouse esophagus implicates the presence of a nonquiescent epithelial stem cell population. *Cell Rep.* 9, 701–711.
- Dorrell, C., Tarlow, B., Wang, Y., Canaday, P.S., Haft, A., Schug, J., Streeter, P.R., Finegold, M.J., Shenje, L.T., Kaestner, K.H., and Grompe, M. (2014). The organoid-initiating cells in mouse pancreas and liver are phenotypically and functionally similar. *Stem Cell Res. (Amst.)* 13, 275–283.
- Drost, J., van Jaarsveld, R.H., Ponsioen, B., Zimmerlin, C., van Boxtel, R., Buijs, A., Sachs, N., Overmeer, R.M., Offerhaus, G.J., Begthel, H., et al. (2015). Sequential cancer mutations in cultured human intestinal stem cells. *Nature* 521, 43–47.
- Dye, B.R., Hill, D.R., Ferguson, M.A., Tsai, Y.H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., et al. (2015). *In vitro* generation of human pluripotent stem cell derived lung organoids. *eLife* 4, 4.
- Eiraku, M., and Sasai, Y. (2012). Self-formation of layered neural structures in three-dimensional culture of ES cells. *Curr. Opin. Neurobiol.* 22, 768–777.
- Eiraku, M., Watanabe, K., Matsuo-Takasaki, M., Kawada, M., Yonemura, S., Matsumura, M., Wataya, T., Nishiyama, A., Muguruma, K., and Sasai, Y. (2008). Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell* 3, 519–532.
- Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* 472, 51–56.
- Firth, A.L., Menon, T., Parker, G.S., Qualls, S.J., Lewis, B.M., Ke, E., Dargitz, C.T., Wright, R., Khanna, A., Gage, F.H., and Verma, I.M. (2015). Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. *Cell Rep.* 12, 1385–1390.
- Fordham, R.P., Yui, S., Hannan, N.R., Soendergaard, C., Madgwick, A., Schweiger, P.J., Nielsen, O.H., Vallier, L., Pedersen, R.A., Nakamura, T., et al. (2013). Transplantation of expanded fetal intestinal progenitors contributes to colon regeneration after injury. *Cell Stem Cell* 13, 734–744.
- Fukuda, M., Mizutani, T., Mochizuki, W., Matsumoto, T., Nozaki, K., Sakamaki, Y., Ichinose, S., Okada, Y., Tanaka, T., Watanabe, M., and Nakamura, T. (2014). Small intestinal stem cell identity is maintained with functional Paneth cells in heterotopically grafted epithelium onto the colon. *Genes Dev.* 28, 1752–1757.
- Gallico, G.G., 3rd, O'Connor, N.E., Compton, C.C., Kehinde, O., and Green, H. (1984). Permanent coverage of large burn wounds with autologous cultured human epithelium. *N. Engl. J. Med.* 311, 448–451.
- Gao, D., Vela, I., Stoner, A., Iaquinta, P.J., Karthaus, W.R., Gopalan, A., Downing, C., Wanjala, J.N., Undvall, E.A., Arora, V.K., et al. (2014). Organoid cultures derived from patients with advanced prostate cancer. *Cell* 159, 176–187.
- Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 525, 251–255.

- Huang, S.X., Islam, M.N., O'Neill, J., Hu, Z., Yang, Y.G., Chen, Y.W., Mumau, M., Green, M.D., Vunjak-Novakovic, G., Bhattacharya, J., and Snoeck, H.W. (2014). Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nat. Biotechnol.* **32**, 84–91.
- Huang, L., Holtzinger, A., Jagan, I., BeGora, M., Lohse, I., Ngai, N., Nostro, C., Wang, R., Muthuswamy, L.B., Crawford, H.C., et al. (2015). Ductal pancreatic cancer modeling and drug screening using human pluripotent stem cell- and patient-derived tumor organoids. *Nat. Med.* **21**, 1364–1371.
- Huch, M., Bonfanti, P., Boj, S.F., Sato, T., Loomans, C.J., van de Wetering, M., Sojoodi, M., Li, V.S., Schuijers, J., Gracanin, A., et al. (2013a). Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. *EMBO J.* **32**, 2708–2721.
- Huch, M., Dorrell, C., Boj, S.F., van Es, J.H., Li, V.S., van de Wetering, M., Sato, T., Hamer, K., Sasaki, N., Finegold, M.J., et al. (2013b). In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. *Nature* **494**, 247–250.
- Huch, M., Gehart, H., van Boxtel, R., Hamer, K., Blokzijl, F., Verstegen, M.M., Ellis, E., van Wenum, M., Fuchs, S.A., de Ligt, J., et al. (2015). Long-term culture of genome-stable bipotent stem cells from adult human liver. *Cell* **160**, 299–312.
- Jain, R., Barkauskas, C.E., Takeda, N., Bowie, E.J., Aghajanian, H., Wang, Q., Padmanabhan, A., Manderfield, L.J., Gupta, M., Li, D., et al. (2015). Plasticity of Hopx(+) type I alveolar cells to regenerate type II cells in the lung. *Nat. Commun.* **6**, 6727.
- Jung, P., Sato, T., Merlos-Suárez, A., Barriga, F.M., Iglesias, M., Rossell, D., Auer, H., Gallardo, M., Blasco, M.A., Sancho, E., et al. (2011). Isolation and in vitro expansion of human colonic stem cells. *Nat. Med.* **17**, 1225–1227.
- Karthauss, W.R., laquinta, P.J., Drost, J., Gracanin, A., van Boxtel, R., Wongvipat, J., Dowling, C.M., Gao, D., Begthel, H., Sachs, N., et al. (2014). Identification of multipotent luminal progenitor cells in human prostate organoid cultures. *Cell* **159**, 163–175.
- Kessler, M., Hoffmann, K., Brinkmann, V., Thieck, O., Jackisch, S., Toelle, B., Berger, H., Mollenkopf, H.J., Mangler, M., Sehouli, J., et al. (2015). The Notch and Wnt pathways regulate stemness and differentiation in human fallopian tube organoids. *Nat. Commun.* **6**, 8989.
- Korinek, V., Barker, N., Moerer, P., van Donselaar, E., Huls, G., Peters, P.J., and Clevers, H. (1998). Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat. Genet.* **19**, 379–383.
- Kurmann, A.A., Serra, M., Hawkins, F., Rankin, S.A., Mori, M., Astapova, I., Ullas, S., Lin, S., Bilodeau, M., Rossant, J., et al. (2015). Regeneration of thyroid function by transplantation of differentiated pluripotent stem cells. *Cell Stem Cell* **17**, 527–542.
- Lancaster, M.A., and Knoblich, J.A. (2014). Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345**, 1247–1255.
- Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurler, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379.
- Li, X., Nadauld, L., Ootani, A., Corney, D.C., Pai, R.K., Gevaert, O., Cantrell, M.A., Rack, P.G., Neal, J.T., Chan, C.W., et al. (2014). Oncogenic transformation of diverse gastrointestinal tissues in primary organoid culture. *Nat. Med.* **20**, 769–777.
- Lindberg, K., Brown, M.E., Chaves, H.V., Kenyon, K.R., and Rheinwald, J.G. (1993). In vitro propagation of human ocular surface epithelial cells for transplantation. *Invest. Ophthalmol. Vis. Sci.* **34**, 2672–2679.
- Linnemann, J.R., Miura, H., Meixner, L.K., Irmeler, M., Kloos, U.J., Hirschi, B., Bartsch, H.S., Sass, S., Beckers, J., Theis, F.J., et al. (2015). Quantification of regenerative potential in primary human mammary epithelial cells. *Development* **142**, 3239–3251.
- Longmire, T.A., Ikononou, L., Hawkins, F., Christodoulou, C., Cao, Y., Jean, J.C., Kwok, L.W., Mou, H., Rajagopal, J., Shen, S.S., et al. (2012). Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. *Cell Stem Cell* **10**, 398–411.
- Ma, R., Latif, R., and Davies, T.F. (2015). Human embryonic stem cells form functional thyroid follicles. *Thyroid* **25**, 455–461.
- Mae, S., Shono, A., Shiota, F., Yasuno, T., Kajiwara, M., Gotoda-Nishimura, N., Arai, S., Sato-Otubo, A., Toyoda, T., Takahashi, K., et al. (2013). Monitoring and robust induction of nephrogenic intermediate mesoderm from human pluripotent stem cells. *Nat. Commun.* **4**, 1367.
- Maimets, M., Rocchi, C., Bron, R., Pringle, S., Kuipers, J., Giepmans, B.N.G., Vries, R.G.J., Clevers, H., De Haan, G., Van Os, R., et al. (2016). Long-term in vitro expansion of salivary gland stem cells driven by Wnt signals. *Stem Cell Reports* **6**, 1–13.
- Matano, M., Date, S., Shimokawa, M., Takano, A., Fujii, M., Ohta, Y., Watanabe, T., Kanai, T., and Sato, T. (2015). Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* **21**, 256–262.
- McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.H., Mayhew, C.N., Spence, J.R., Zavros, Y., and Wells, J.M. (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* **516**, 400–404.
- Muguruma, K., Nishiyama, A., Ono, Y., Miyawaki, H., Mizuhara, E., Hori, S., Kakizuka, A., Obata, K., Yanagawa, Y., Hirano, T., and Sasai, Y. (2010). Ontogeny-recapitulating generation and tissue integration of ES cell-derived Purkinje cells. *Nat. Neurosci.* **13**, 1171–1180.
- Muguruma, K., Nishiyama, A., Kawakami, H., Hashimoto, K., and Sasai, Y. (2015). Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. *Cell Rep.* **10**, 537–550.
- Nadauld, L.D., Garcia, S., Natsoulis, G., Bell, J.M., Miotke, L., Hopmans, E.S., Xu, H., Pai, R.K., Palm, C., Regan, J.F., et al. (2014). Metastatic tumor evolution and organoid modeling implicate TGFBR2 as a cancer driver in diffuse gastric cancer. *Genome Biol.* **15**, 428.
- Nakano, T., Ando, S., Takata, N., Kawada, M., Muguruma, K., Sekiguchi, K., Saito, K., Yonemura, S., Eiraku, M., and Sasai, Y. (2012). Self-formation of optic cups and storable stratified neural retina from human ESCs. *Cell Stem Cell* **10**, 771–785.
- Nanduri, L.S., Baanstra, M., Faber, H., Rocchi, C., Zwart, E., de Haan, G., van Os, R., and Coppes, R.P. (2014). Purification and ex vivo expansion of fully functional salivary gland stem cells. *Stem Cell Reports* **3**, 957–964.
- Nantasanti, S., Spee, B., Kruitwagen, H.S., Chen, C., Geijsen, N., Oosterhoff, L.A., van Wolferen, M.E., Pelaez, N., Fieten, H., Wubbolts, R.W., et al. (2015). Disease Modeling and Gene Therapy of Copper Storage Disease in Canine Hepatic Organoids. *Stem Cell Reports* **5**, 895–907.
- O'Connor, N.E., Mulliken, J.B., Banks-Schlegel, S., Kehinde, O., and Green, H. (1981). Grafting of burns with cultured epithelium prepared from autologous epidermal cells. *Lancet* **1**, 75–78.
- Ootani, A., Li, X., Sangiorgi, E., Ho, Q.T., Ueno, H., Toda, S., Sugihara, H., Fujimoto, K., Weissman, I.L., Capecchi, M.R., and Kuo, C.J. (2009). Sustained in vitro intestinal epithelial culture within a Wnt-dependent stem cell niche. *Nat. Med.* **15**, 701–706.
- Pellegrini, G., Traverso, C.E., Franzini, A.T., Zingirian, M., Cancedda, R., and De Luca, M. (1997). Long-term restoration of damaged corneal surfaces with autologous cultivated corneal epithelium. *Lancet* **349**, 990–993.
- Plaks, V., Brenot, A., Lawson, D.A., Linnemann, J.R., Van Kappel, E.C., Wong, K.C., de Sauvage, F., Klein, O.D., and Werb, Z. (2013). Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. *Cell Rep.* **3**, 70–78.
- Qian, X., Nguyen, H.N., Song, M.M., Hadiono, C., Ogden, S.C., Hammack, C., Yao, B., Hamersky, G.R., Jacob, F., Zhong, C., et al. (2016). Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell* **165**, 1238–1254.
- Rama, P., Matuska, S., Paganoni, G., Spinelli, A., De Luca, M., and Pellegrini, G. (2010). Limbal stem-cell therapy and long-term corneal regeneration. *N. Engl. J. Med.* **363**, 147–155.

- Ren, W., Lewandowski, B.C., Watson, J., Aihara, E., Iwatsuki, K., Bachmanov, A.A., Margolskee, R.F., and Jiang, P. (2014). Single Lgr5- or Lgr6-expressing taste stem/progenitor cells generate taste bud cells *ex vivo*. *Proc. Natl. Acad. Sci. USA* *111*, 16401–16406.
- Rheinwald, J.G., and Green, H. (1975). Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. *Cell* *6*, 331–343.
- Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. *Nature* *506*, 322–327.
- Rock, J.R., Onaitis, M.W., Rawlins, E.L., Lu, Y., Clark, C.P., Xue, Y., Randell, S.H., and Hogan, B.L. (2009). Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proc. Natl. Acad. Sci. USA* *106*, 12771–12775.
- Sakaguchi, H., Kadoshima, T., Soen, M., Narii, N., Ishida, Y., Ohgushi, M., Takahashi, J., Eiraku, M., and Sasai, Y. (2015). Generation of functional hippocampal neurons from self-organizing human embryonic stem cell-derived dorsomedial telencephalic tissue. *Nat. Commun.* *6*, 8896.
- Sato, T., and Clevers, H. (2013). Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. *Science* *340*, 1190–1194.
- Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., and Clevers, H. (2009). Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* *459*, 262–265.
- Sato, T., Stange, D.E., Ferrante, M., Vries, R.G., Van Es, J.H., Van den Brink, S., Van Houdt, W.J., Pronk, A., Van Gorp, J., Siersema, P.D., and Clevers, H. (2011). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* *141*, 1762–1772.
- Schwank, G., Koo, B.K., Sasselli, V., Dekkers, J.F., Heo, I., Demircan, T., Sasaki, N., Boymans, S., Cuppen, E., van der Ent, C.K., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* *13*, 653–658.
- Sinagoga, K.L., and Wells, J.M. (2015). Generating human intestinal tissues from pluripotent stem cells to study development and disease. *EMBO J.* *34*, 1149–1163.
- Spence, J.R., Mayhew, C.N., Rankin, S.A., Kuhar, M.F., Vallance, J.E., Tolle, K., Hoskins, E.E., Kalinichenko, V.V., Wells, S.I., Zorn, A.M., et al. (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue *in vitro*. *Nature* *470*, 105–109.
- Stange, D.E., Koo, B.K., Huch, M., Sibbel, G., Basak, O., Lyubimova, A., Kujala, P., Bartfeld, S., Koster, J., Geahlen, J.H., et al. (2013). Differentiated Trophoblast cells act as reserve stem cells to generate all lineages of the stomach epithelium. *Cell* *155*, 357–368.
- Tadokoro, T., Wang, Y., Barak, L.S., Bai, Y., Randell, S.H., and Hogan, B.L. (2014). IL-6/STAT3 promotes regeneration of airway ciliated cells from basal stem cells. *Proc. Natl. Acad. Sci. USA* *111*, E3641–E3649.
- Taguchi, A., Kaku, Y., Ohmori, T., Sharmin, S., Ogawa, M., Sasaki, H., and Nishinakamura, R. (2014). Redefining the *in vivo* origin of metanephric nephron progenitors enables generation of complex kidney structures from pluripotent stem cells. *Cell Stem Cell* *14*, 53–67.
- Takasato, M., Er, P.X., Becroft, M., Vanslambrouck, J.M., Stanley, E.G., Elefanti, A.G., and Little, M.H. (2014). Directing human embryonic stem cell differentiation towards a renal lineage generates a self-organizing kidney. *Nat. Cell Biol.* *16*, 118–126.
- Takasato, M., Er, P.X., Chiu, H.S., Maier, B., Baillie, G.J., Ferguson, C., Parton, R.G., Wolvetang, E.J., Roost, M.S., Chuva de Sousa Lopes, S.M., and Little, M.H. (2015). Kidney organoids from human iPSC cells contain multiple lineages and model human nephrogenesis. *Nature* *526*, 564–568.
- Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.R., Ueno, Y., Zheng, Y.W., Koike, N., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* *499*, 481–484.
- Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* *509*, 371–375.
- van de Wetering, M., Francies, H.E., Francis, J.M., Bounova, G., Iorio, F., Pronk, A., van Houdt, W., van Gorp, J., Taylor-Weiner, A., Kester, L., et al. (2015). Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* *161*, 933–945.
- von Furstenberg, R.J., Gulati, A.S., Baxi, A., Doherty, J.M., Stappenbeck, T.S., Gracz, A.D., Magness, S.T., and Henning, S.J. (2011). Sorting mouse jejunal epithelial cells with CD24 yields a population with characteristics of intestinal stem cells. *Am. J. Physiol. Gastrointest. Liver Physiol.* *300*, G409–G417.
- Wang, F., Scoville, D., He, X.C., Mahe, M.M., Box, A., Perry, J.M., Smith, N.R., Lei, N.Y., Davies, P.S., Fuller, M.K., et al. (2013). Isolation and characterization of intestinal stem cells based on surface marker combinations and colony-formation assay. *Gastroenterology* *145*, 383–395.e1–e21.
- Wang, X., Yamamoto, Y., Wilson, L.H., Zhang, T., Howitt, B.E., Farrow, M.A., Kern, F., Ning, G., Hong, Y., Khor, C.C., et al. (2015). Cloning and variation of ground state intestinal stem cells. *Nature* *522*, 173–178.
- Watson, C.L., Mahe, M.M., Múnera, J., Howell, J.C., Sundaram, N., Poling, H.M., Schweitzer, J.I., Vallance, J.E., Mayhew, C.N., Sun, Y., et al. (2014). An *in vivo* model of human small intestine using pluripotent stem cells. *Nat. Med.* *20*, 1310–1314.
- Wong, A.P., Bear, C.E., Chin, S., Pasceri, P., Thompson, T.O., Huan, L.J., Ratjen, F., Ellis, J., and Rossant, J. (2012). Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein. *Nat. Biotechnol.* *30*, 876–882.
- Xia, Y., Nivet, E., Sancho-Martinez, I., Gallegos, T., Suzuki, K., Okamura, D., Wu, M.Z., Dubova, I., Esteban, C.R., Montserrat, N., et al. (2013). Directed differentiation of human pluripotent cells to ureteric bud kidney progenitor-like cells. *Nat. Cell Biol.* *15*, 1507–1515.
- Yin, X., Farin, H.F., van Es, J.H., Clevers, H., Langer, R., and Karp, J.M. (2014). Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny. *Nat. Methods* *11*, 106–112.
- Yui, S., Nakamura, T., Sato, T., Nemoto, Y., Mizutani, T., Zheng, X., Ichinose, S., Nagaishi, T., Okamoto, R., Tsuchiya, K., et al. (2012). Functional engraftment of colon epithelium expanded *in vitro* from a single adult Lgr5+ stem cell. *Nat. Med.* *18*, 618–623.

Introducing  
**STAR★METHODS**

**Empowering methods,  
to empower you.**

STAR Methods promote rigor and robustness with an intuitive, consistent framework that integrates seamlessly into the scientific information flow—making reporting easier for the author and replication easier for the reader.

STAR Methods—a new methods section designed with authors and readers in mind—are now available in *Cell* and *Cell Systems*. Learn more about how we're empowering methods.

Visit [www.cell.com/star-methods](http://www.cell.com/star-methods) today!

**CellPress**

# The Cellular Phase of Alzheimer's Disease

Bart De Strooper<sup>1,2,3,\*</sup> and Eric Karran<sup>2,3,4</sup>

<sup>1</sup>VIB Center for the Biology of Disease, VIB-Leuven, 3000 Leuven, Belgium

<sup>2</sup>Center for Human Genetics, Universitaire Ziekenhuizen and LIND, KU Leuven, 3000 Leuven, Belgium

<sup>3</sup>Institute of Neurology, University College London, Queen Square, WC1N 3BG London, UK

<sup>4</sup>Alzheimer Research UK, Cambridge, CB21 6AD, UK

\*Correspondence: bart.destrooper@cme.vib-kuleuven.be

<http://dx.doi.org/10.1016/j.cell.2015.12.056>

**The amyloid hypothesis for Alzheimer's disease (AD) posits a neuron-centric, linear cascade initiated by A $\beta$  and leading to dementia. This direct causality is incompatible with clinical observations. We review evidence supporting a long, complex cellular phase consisting of feedback and feedforward responses of astrocytes, microglia, and vasculature. The field must incorporate this holistic view and take advantage of advances in single-cell approaches to resolve the critical junctures at which perturbations initially amenable to compensatory feedback transform into irreversible, progressive neurodegeneration.**

## Introduction

Alzheimer's disease (AD) impairs cognition, memory, and language and causes dementia. In 1907, Alois Alzheimer described the extensive distribution of neuronal tangles and amyloid plaques in brain that, together with astrogliosis, neuronal dystrophy, neuronal loss, and vascular alterations, constitute the hallmarks of the disorder.

For over 20 years, the "amyloid cascade hypothesis" has provided the main theoretical construct for AD (Hardy and Selkoe, 2002). It proposes amyloid plaques or their major constituents, the A $\beta$ -peptides, as the direct cause of the progressive neurodegeneration. The concept that a biochemical/biophysical process underlies the complex cognitive alterations has transformed AD research, replacing earlier descriptive studies with a molecular, mechanistic view. This has led to diagnostic and therapeutic trials for a disease believed before to be either incurable or an inevitable consequence of aging. The tremendous increase in knowledge on the molecular biology, pathophysiology, and diagnosis of AD is exciting and holds promise for future prevention and therapies but also starts to erode the assumptions of its main theoretical foundation.

To understand the disease, the well-studied biochemistry needs to be integrated into the complex cellular context of the brain. We examine here how a failing vascular system contributes to disease progression; how inhibitory and excitatory neurons, neuronal networks, microglia, astroglia, and, finally, oligodendrocytes all contribute to a complex cellular phase of the disease evolving over decades, and how initial benign reactions ultimately become chronic, resulting in an irreversible dys-homeostasis of the brain. Emerging approaches could provide us with a more holistic understanding of the spatial, temporal, and cellular aspects of the disease process. We start with a brief critique of the amyloid cascade hypothesis itself.

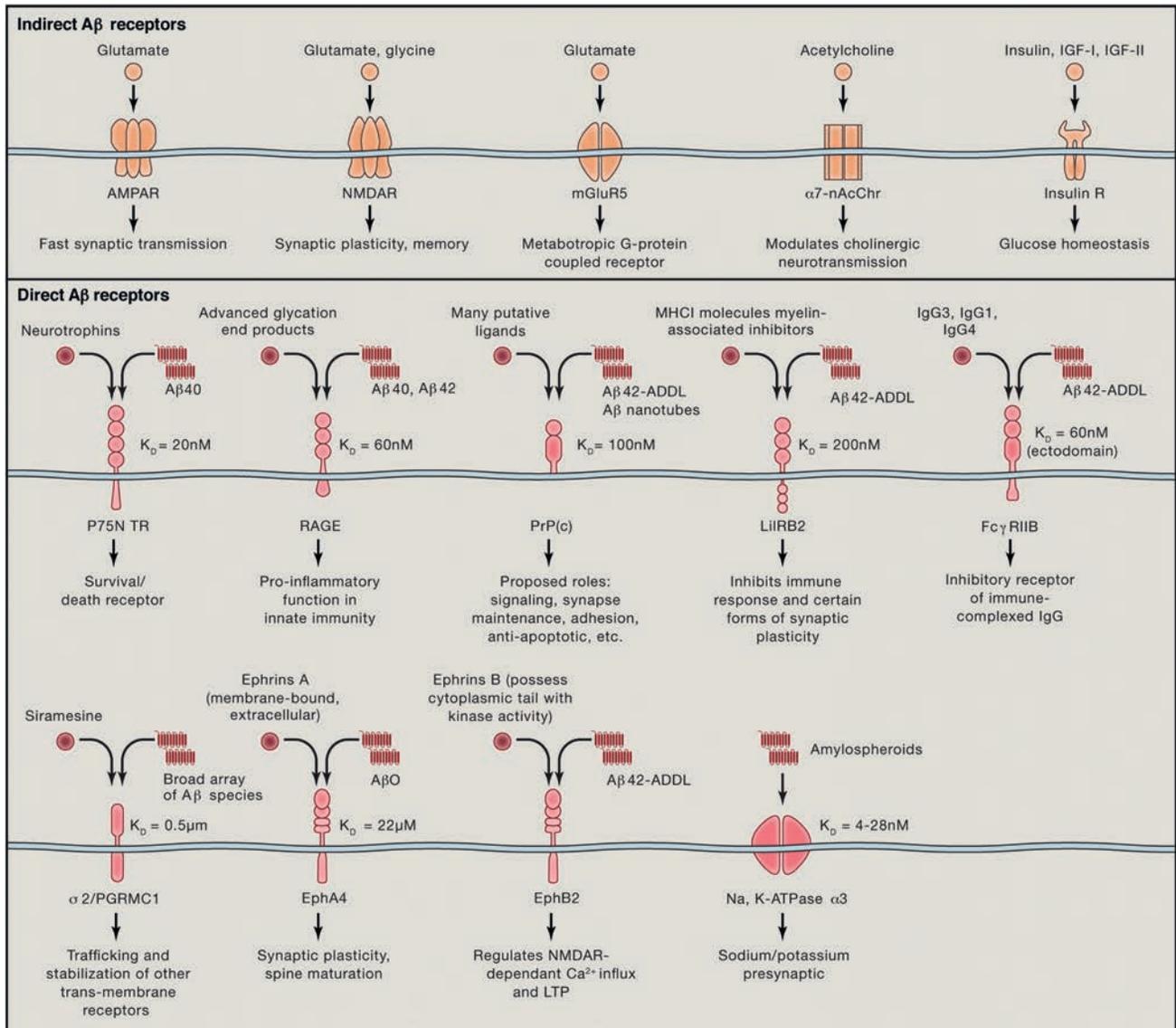
## The Amyloid Hypothesis

The amyloid hypothesis for AD is a neuron-centric, linear, and quantitative model postulating direct cause and consequences

in a cascade initiated by A $\beta$  deposition and leading progressively to Tau pathology, synaptic dysfunction, inflammation, neuronal loss, and, ultimately, dementia.

The linearity of the cascade remains very controversial. For instance, exploration of the direct link between A $\beta$  and neurotoxicity has resulted in a rather confusing literature with at least ten different molecular mechanisms and receptors (see Figure 1) (Benilova et al., 2012). It is debatable whether they all equally contribute to the disease process in humans, as many of the supporting studies used non-physiologically high concentrations of A $\beta$  or transgenic mice significantly overexpressing APP (see Nilsson et al., 2014). However, an alternative explanation for the richness of these findings is that different mechanisms or even the same mechanism exert multiple effects at different stages of disease progression. For instance, the interaction of A $\beta$  with the receptor for advanced glycation end products (RAGE) causes oxidative stress in neurons but enhances inflammatory responses in microglia and is involved in endothelial cells in "reversed" transport of A $\beta$  across the blood-brain barrier (BBB) (Deane et al., 2012). Thus, the neuron-centric view has to be expanded to one that considers the contribution of different cell types, their interactions with each other, and the gradual evolution of the disease.

Preclinical or prodromal AD stages have been defined by clinicians to underscore the concept that AD starts insidiously many decades before full dementia becomes apparent (Carrillo et al., 2013; Dubois et al., 2014). The amyloid cascade hypothesis provides no explanation for the silent incubation period of AD. Even in familial AD (FAD), where the role of A $\beta$  is established because of causative mutations in APP and Presenilin, the amyloid hypothesis is found wanting. Why, for instance, does abnormal A $\beta$  production, already present from conception, manifest its toxicity only in midlife? Clearly the complex compensation mechanisms in the brain that maintain almost normal cognitive performance for decades should be accommodated into any theoretical construct of the disease. In our view, AD only starts when a pathological cellular reaction is initiated.



**Figure 1. Aβ Receptors and Aβ Toxicity Mechanisms**

All membrane receptors implied in Aβ pathogenic mechanisms are indicated. Some receptors are only indirectly affected by Aβ, inferred by physiological alterations of cells exposed to Aβ. The second series are receptors shown to directly bind Aβ (see Benilova et al., 2012).

The quantitative aspects of the hypothesis (the concept that more amyloid plaques, or, in more recent versions, more Aβ<sub>42</sub>, is causing disease) imply that quantitative lowering of Aβ-plaques or Aβ-oligomers should be sufficient to halt progression of AD (critically discussed in Karran et al., 2011). The discrepancy between amyloid load and disease symptoms in sporadic AD (SAD) has, however, been pointed out in the past: “it seems as if the Aβ plaques appear at the wrong time and in the wrong places with respect to the clinical dementia” (Mesulam, 1999). Less appreciated is the fact that the quantitative concept is equally not tenable for FAD. Most Presenilin (PSEN1) mutations do not increase Aβ generation. PSEN1 is part of the γ-secretase complex, and mutations affect mainly the “trimming” function of

these proteases whereby the C-terminal region of long Aβ is progressively cleaved. The mutations might destabilize the three-dimensional structure of the γ-secretases, which might explain the premature release of longer (insufficiently “trimmed”) hydrophobic Aβ<sub>42</sub>, Aβ<sub>43</sub>, and, hypothetically, Aβ<sub>>43</sub>, peptides (Chávez-Gutiérrez et al., 2012; Szaruga et al., 2015; Saito et al., 2011). In this view, even lowering of total Aβ could still be pathogenic if the Aβ produced is longer than Aβ<sub>40</sub>. Such long Aβ, even in low amounts, could provide seeds for further nucleation and amyloidosis. Thus, in contrast to what is assumed in the amyloid hypothesis, the pathological effects of the PSEN1 mutations on Aβ production are qualitative and not quantitative (Kuperstein et al., 2010; Szaruga et al., 2015).

### The Biochemical Phase in AD Causes Proteostatic Stress and Homeostatic Cellular Responses

The “biochemical phase” of AD is characterized by abnormal aggregation of A $\beta$  and potentially abnormal processing of APP; formation of Tau conformers followed by hyper phosphorylation; the seeding and propagation in a prionoid fashion of these abnormally folded proteins (Walker and Jucker, 2015); the generation of oligomeric amyloids (A $\beta$ - and Tau-oligomers); and the generation of amyloid plaques and neuronal tangles. The different conformations of A $\beta$  and Tau exert “proteopathic” or “aggregate” stress on different cells of the brain, i.e., they interact with proteins and membranes and interfere with signaling and other functions. They also bind APP and Tau, which disturbs the normal activities of these proteins in neurotransmission (Palop and Mucke, 2010). The temporal sequence of these events is undefined. Age-related deficiencies of the proteostatic network could cause accumulation of aggregating proteins; conversely, protein aggregates might disturb the complex cellular molecular network that sustains protein folding (Labbadia and Morimoto, 2015). The lysosomal/endolysosomal system and, in particular, autophagy is a key regulator of this proteopathic stage (Nixon, 2013). In AD brains, a striking increase in classic autophagosomes, featuring a double membrane that encapsulates heterogeneous intracellular components, is seen in an early stage of the disorder (Nixon, 2013). Deletion of key macroautophagy components such as ATG5 or ATG7 in neurons in mice causes motor deficits, ubiquitinated neuronal inclusions, and neurodegeneration (Komatsu et al., 2006). In Niemann-Pick’s disease type C, loss of function of the *NPC1* or *NPC2* genes leads to late endosomal/lysosomal accumulation of lipids and defective autophagocytosis (Elrick et al., 2012). Affected neurons display tau inclusions that are indistinguishable from those in AD. Thus, abnormalities in autophagy are linked to tau pathology. The genetic support for a role of autophagy in AD compared to Parkinson disease, for example, remains surprisingly limited, however (Lambert et al., 2013).

None of the initial effects of proteopathy should be considered overwhelming or irreversible, as brain cells can apparently survive this stress for many years. The elicited responses are physiological and can be cell or non-cell autonomous. They can be biochemical, maintaining homeostasis in the proteostatic network (Labbadia and Morimoto, 2015), or functional, involving various synaptic plasticity mechanisms in which astroglia may play crucial roles. Inflammatory responses contribute initially to sustaining homeostasis. It is only when these compensating mechanisms turn into chronic, irreversible, and pathological processes that disease progresses inexorably. The transition of an early reversible to a chronic, irreversible, autonomous cellular response, likely no longer dependent on the initial A $\beta$  and Tau aggregate stressors, represents a critical phase in the disease process.

The separate locations of Tau and amyloid pathology and their natural progression in the brain (Braak and Braak, 1995; Thal et al., 2002) suggest that the two biochemical pathologies progress relatively independently from each other (Mesulam, 1999; Small and Duff, 2008). Tangles are universally found in all aged people and are already present early in life (Braak and Del Tredici, 2011). This primary age-related tauopathy or PART (Crary

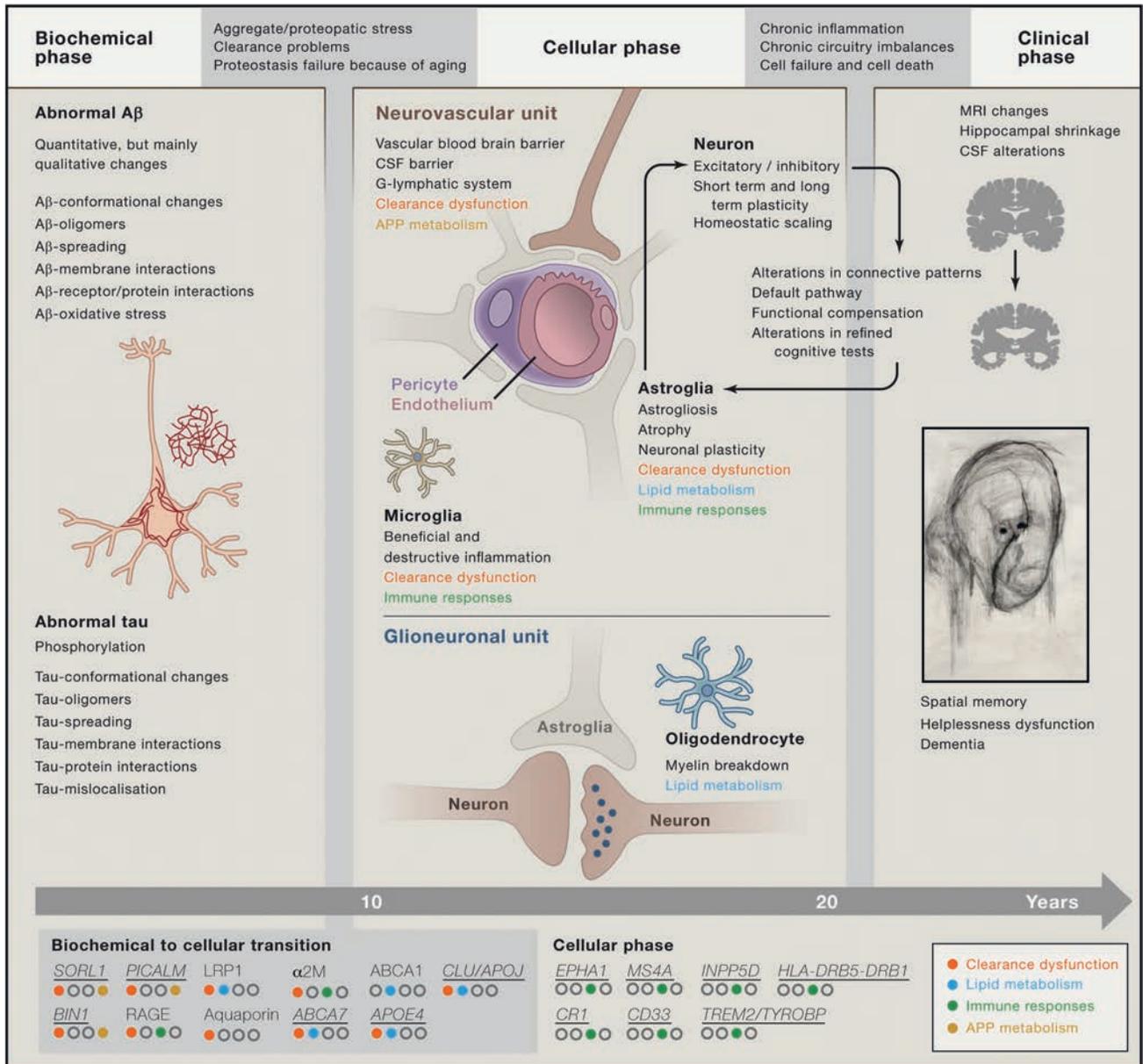
et al., 2014) is mild. The effects of age-associated A $\beta$  accumulation are initially also mild—they are found in many aged people without causing clinical symptoms of brain dysfunction. A $\beta$  stress might enhance the progression of Tau pathology (Khan et al., 2014), but described molecular links are multiple and indirect. Alternatively, A $\beta$ -stress and PART could be considered independent stressors that each cause a slowly evolving, relatively benign, age-related neurodegenerative process with no or only mild symptomatology. Only when they come together and reciprocally augment their deleterious effects is a more aggressive disease manifested with multiple cellular effects. Even in pure FAD, this scenario is conceivable: the abnormal A $\beta$  generated by the mutations becomes only part of a disease process when the “natural” process of progressive accumulation of neurofibrillar tangles has reached a threshold and both pathologies start to interact. The rather stereotypical evolution of the disease is compatible with this view (Karran et al., 2011).

In the current Review, we try to go beyond the discussion—which is nevertheless important—of the biochemistry and toxicity of A $\beta$  and Tau (De Strooper, 2010). We see AD as a complex process consisting of different interacting phases encompassing feedback and feedforward mechanisms between cells, which, after many years, lead to dementia. Most importantly, in this view, A $\beta$  and Tau accumulations in SAD are considered risk factors—i.e., even if their accumulation is a sign of increased proteopathic stress, additional cellular factors are determinative in the evolution toward dementia. Instead of A $\beta$  and Tau, the real causes of sporadic disease are upstream of these proteopathies and are likely manifold, with aging being the major driver. Nevertheless, the proteopathic stress of the biochemical phase results in the cellular phase of AD. The clinical phase of the disease is initiated when the cellular reaction can no longer maintain homeostasis (Figure 2).

### Defective Clearance Mechanisms Are Part of the Initial Cellular Phase

Defective clearance of A $\beta$  (and Tau), considered to underlie the majority of sporadic AD cases, illustrates the complexity of the cellular response. Many proteases can break down A $\beta$  peptides in the biochemical phase (Figure 2) (De Strooper, 2010), but, as shown by injecting labeled A $\beta$  or Tau in the brain parenchyma and following its elimination (Iliff et al., 2014), it is now clear that a large part of normal A $\beta$  turnover depends on bulk flow via the perivascular circulation and the glymphatic system in the brain (Tarasoff-Conway et al., 2015). Genetic risk associated with *APOE4*, *PICALM*, and *APOJ* (or *CLU*) is thought to affect A $\beta$  clearance (Verghese et al., 2013; Zhao et al., 2015).

The cellular component of the clearance system is complex. The BBB consists of tightly connected endothelial cells and astroglia end-feet. Pericytes and smooth muscle cells provide contractile elements. The glia barrier is relatively open because of gap junctions allowing passage of A $\beta$  and Tau. The endothelial barrier is, however, not permeable, and therefore, specialized transport proteins (receptors of the LDL receptor family, e.g., LRP1&2, RAGE, ABC transporter ABCB1 or P-glycoprotein, ABCA1, and A $\beta$  free or bound to proteins such as  $\alpha$ 2-macroglobulin, APOJ, or APOE particles) are responsible for A $\beta$  (and Tau) egress to the circulation. For instance, heterozygote *Picalm*<sup>+/-</sup>



**Figure 2. The Biochemical, Cellular, and Clinical Phases of AD**

The transition of biochemical to cellular and cellular to clinical phases are indicated. The cellular phase is complex and evolves gradually over two decades from reversible physiological reactions upon proteopathic stress to irreversible immune and electrophysiological compensation mechanisms that disturb the normal homeostasis of the brain. The effects are cell and non-cell autonomous and occur in the context of neurovascular or glioneuronal units (see main text). Processes are derived from GWAS studies (clearance dysfunction, lipid metabolism, immune response, protein trafficking) and are indicated in different colors. Genes are in italics when genetic evidence links them to AD. Other genes (LRP1, RAGE and others) are functionally implicated in the disease process. Each gene is associated to a color code indicating the pathways in which they operate. The picture is a late self-portrait from the German painter William Utermohlen, who suffered from AD. Image courtesy of Chris Boicos Fine Arts, Paris.

mice show decreased LRP1-mediated clearance of Aβ via the BBB. Although neurons with downregulated *Picalm* produce less Aβ, accelerated amyloidosis and worsening in behavioral parameters is seen in these mice; this could be rescued via adenoviral expression of PICALM in the endothelium (Zhao et al., 2015). PICALM proteolysis has also been associated with Tau-tangle formation and disturbed endocytosis in SAD (Ando et al., 2013). The multifactorial effects of genes involved

in AD are a recurrent theme in this Review and illustrate the importance of placing the biochemical process into the context of the cellular phase.

In contrast to Aβ, much less is known regarding Tau clearance. Previously, Tau pathology was seen as a cell-autonomous phenomenon, and studies of Tau clearance were predominantly focused on intracellular processes such as autophagy. However, Tau pathology can progress through the brain in a prion-like

fashion (Walker and Jucker, 2015), and Tau is secreted after increased electrical activity (Yamada et al., 2014). Some secreted Tau is cleared by the glymphatic flow (Iliff et al., 2014); thus, it is reasonable to predict that Tau gets cleared from the brain via similar cell-regulated bulk flow pathways as A $\beta$ . However, the role of the vasculature in AD goes far beyond clearance.

### The Neurovascular Unit and the Vascular Hypothesis for AD

The term “neurovascular unit” is used to describe the close anatomic and functional interrelationship between the cellular components of the brain. The vascular hypothesis proposes that initial vascular damage precipitates AD. Hypoperfusion and hypoxia are one aspect of the problem, but the associated breakdown of the BBB also results in accumulation of neurotoxic serum proteins in brain, inflammation, vascular, and synaptic dysfunction that lead secondarily to defects in A $\beta$  and Tau metabolism and clearance, which in turn cause vascular problems such as cerebral amyloid angiopathy (CAA) (Zlokovic, 2011). CAA is caused by widespread deposition of A $\beta$  in the vasculature, resulting in fibrinoid necrosis, microaneurysms, and bleedings. Co-existence of AD and vascular pathology increases the risk for dementia considerably (Attems and Jellinger, 2014).

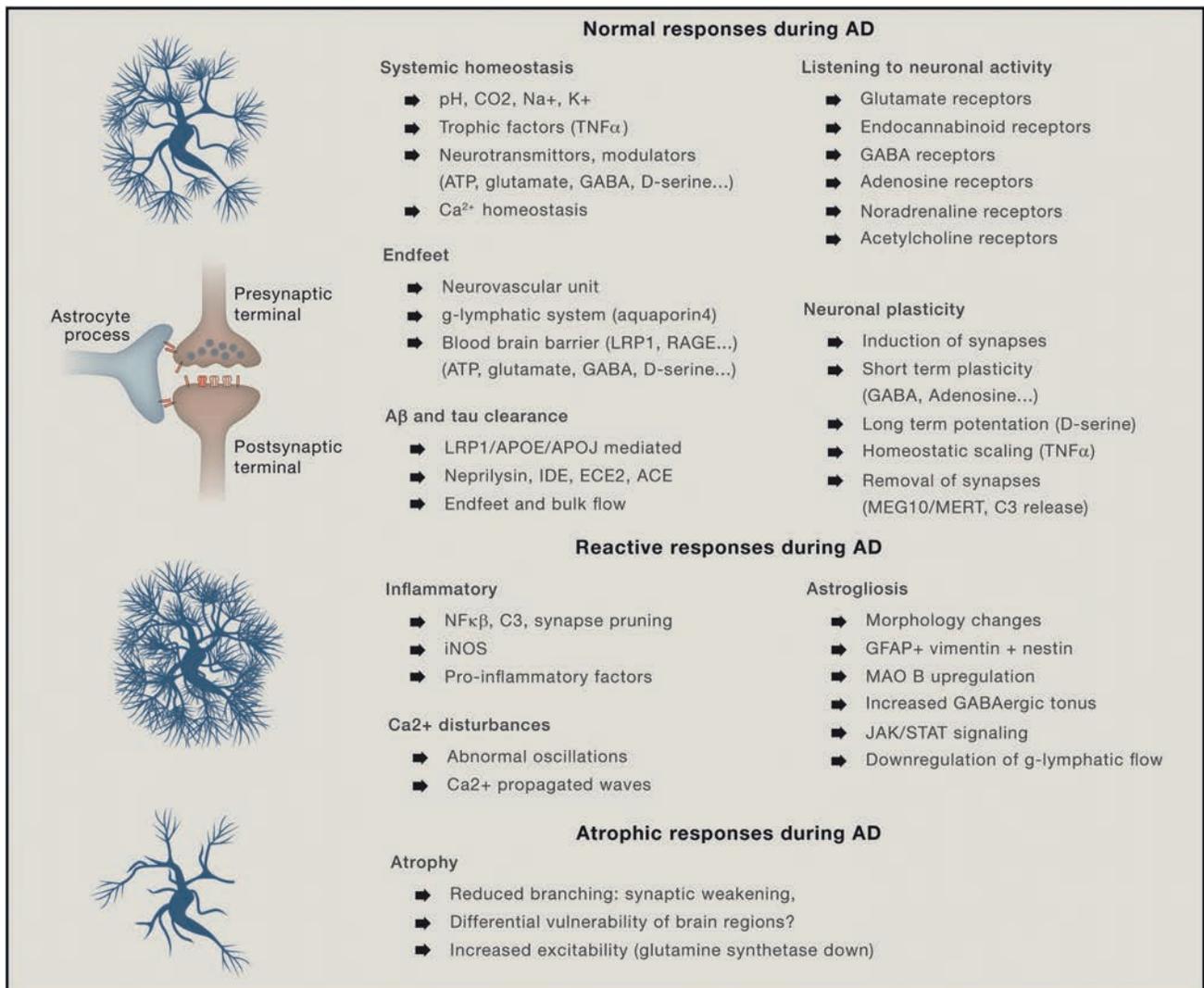
Many functional studies support the central postulate of the vascular hypothesis—i.e., the primacy of vascular lesions in driving the AD phenotype (Zlokovic, 2011). Mice that express the major genetic risk factor for AD, human APOE4 (on a mouse *ApoE*<sup>-/-</sup> background), show a curious degenerative response of endothelial cells and pericytes mediated by cyclophilin A and NF $\kappa$ B that leads to damage of the BBB (Bell et al., 2012), which is also seen in human APOE4 carriers. Other genetic manipulations of receptors, transporters, and cells support the idea that early alterations in the BBB could play a major role in the initiation of AD. For example, AD patients typically display diminished glucose uptake in the brain as measured by 2-[18F]fluoro-2-deoxy-d-glucose positron emission tomography (FDG-PET) (Protas et al., 2013). Interestingly, heterozygous deficiency of *GLUT1* (*Slc2a1*<sup>+/-</sup>) in endothelium accelerated degenerative changes in the APPsw mouse model for AD (Winkler et al., 2015). Degeneration of pericytes or vascular smooth muscle cells under hypoxic stress also leads to neurodegeneration and A $\beta$  clearance problems (Winkler et al., 2014), suggesting that atherosclerosis and hypoperfusion of the brain might be an upstream cause in a subgroup of patients. In models of traumatic brain injury, the clearance flow is slowed down, which might contribute to accumulation of secreted Tau (Iliff et al., 2014). Sleep increases, and aging and/or stiffening of the arterial wall slows down the glymphatic clearance flow (Tarasoff-Conway et al., 2015). While it seems that vascular defects lead to amyloid accumulation, it may also be the case that A $\beta$  accumulation can lead to vascular defects as seen with some clinical APP mutations that manifest as CAA. Thus, A $\beta$  is both cause and consequence of vascular pathology, and this destructive cycle is an important part of the cellular action and reaction that underlies the cellular phase of AD.

### Neurons and Neuronal Circuitry as Part of the Cellular Phase of AD

In the prevailing neuron-centric view, dysfunction of the brain is mediated by decrements in synaptic plasticity: e.g., long-term potentiation and depression, changes in homeostatic scaling, and disruption of neuronal connectivity. These alterations are currently often studied by recording synaptic function in neuronal cell and hippocampal slice models treated with A $\beta$  peptides, either synthetically generated, isolated from “natural” brain material, or expressed in transgenic mouse models. These models are too acute to mimic the progressive cellular disruption that occurs in AD brains over many years.

The complexity of how neurons are progressively affected by A $\beta$  stress has been addressed by a few groups. In a series of landmark publications, Palop and Mucke studied how A $\beta$  stress results in excitatory, (non-convulsive) seizure activity in neuronal networks (Palop and Mucke, 2010). The study nicely illustrates how work in preclinical animal models can translate into novel insights relevant to human pathogenesis: epileptic seizures and neuronal hyperexcitability in AD patients became a recognized part of the condition (Bakker et al., 2012). To explain the unexpected hyper-reactivity, they looked to cellular alterations and found loss of calbindin expression and ectopic expression of the inhibitory neuropeptide Y (NPY) in granule cells and mossy fibers, and GABAergic synaptic sprouting in the molecular layer emanating from NPY- and somatostatin positive interneurons. They also found a reduction in the voltage-gated sodium channel subunit Nav1.1 in parvalbumin cells, leading to defective inhibitory input that could be rectified by restoring channel expression (Verret et al., 2012). Their data suggest that increased A $\beta$  stress in AD leads to a complex pattern of neuronal actions and reactions, causing aberrant excitatory network activity and compensatory inhibitory responses involving learning and memory circuitry (Palop and Mucke, 2010). These fascinating studies are but a first step to address the complex cellular processes in AD. For example, recent work suggests that reactive astroglia in these mice secrete GABA, which paradoxically could contribute to seizure-like activity by suppressing inhibitory input on excitatory circuitry (Jo et al., 2014).

Other mouse models also show that the neuronal reaction to Tau and A $\beta$  proteopathic stress is complex and not just simple cell death. Pyramidal cells in the frontal cortex show phases of hypo- and/or hyperexcitability in Tau and A $\beta$  mouse models (Menkes-Caspi et al., 2015). Oscillatory network activity is altered in many AD models. Finally, the activity of hippocampal place cells that normally show a robust relation to animal location in the environment lose spatial selectivity in AD mice (Cheng and Ji, 2013). Importantly, the recorded alterations are likely due to a mixture of direct and indirect effects of A $\beta$  on neurons and glial cells, synaptic dysfunctions that impact local network activity, and afferent activity in downstream targets and compensatory mechanisms. Disease progression likely entails an initial localized and acute direct or indirect effect of A $\beta$  species on cell and synaptic function, where distinct cell types may have a varying degree of sensitivity. The early local changes in network activity patterns would then progressively affect connected brain structures, followed by the spread of pathology.



**Figure 3. Molecular and Cellular Links between Astroglia and AD**

Astroglia are central in the cellular phase of AD. They interact directly with A $\beta$  and display normal synaptic and metabolic responses in AD. They also display reactive and atrophic responses in AD. Major pathways are schematically represented.

An important caveat is that functional disturbances in mice are typically occurring in the absence of severe tangle formation and neuronal loss that characterizes AD dementia. Some investigators therefore severely doubt the validity of mouse experiments: a more moderate and constructive view is that the alterations in these mice recapitulate early alterations occurring in incipient AD brain (Zahs and Ashe, 2010). In that sense, they are very relevant to the study of the early cellular phase of AD.

Morphological and biochemical studies of human brain material are limited to evolved stages of the disease. However, they too demonstrate the complex cellular phase in AD. Loss of innervation by the entorhinal cortex (EC) of the outer molecular layer of the dentate gyrus is associated with extensive sprouting of cholinergic innervation (Mufson et al., 2015). Loss of input via the Schaffer collaterals is associated initially with increases in length and complexity of the basal dendritic trees of CA1 pyrami-

dal neurons that become reduced only in the very advanced state of AD (Mufson et al., 2015).

### Astrocytes Are Central Players in the Cellular Phase of AD

In contrast to neurons, the astroglia population is strikingly under-investigated in AD-related studies. Each astrocyte deploys many fine processes to contact up to 140,000 synapses in the CA1 region of rat hippocampus (Bushong et al., 2002). With their end feet, they also contact blood vessels, forming the glia limitans of the BBB (Figure 2). The astrocytes occupy anatomically and functionally discrete micro-domains (Bushong et al., 2002), and arrays of astrocyte and neurons are organized around capillaries, providing the basis for higher-order neurovascular units (Allen, 2014).

The metabolic roles of astroglia are well recognized (Figure 3). Importantly, astroglia and oligodendrocytes synthesize all brain

cholesterol and secrete apolipoprotein E and apolipoprotein J that are loaded with cholesterol by ATP-binding cassette transporters—in particular, ABCA1. Synaptogenesis is promoted by glia-derived cholesterol, but there is also complex overlap between the lipid metabolism function of these various genes and the clearance of A $\beta$  over the BBB and blood-CSF barriers. Lipoprotein metabolism is strongly associated with risk for AD as repeatedly suggested in GWAS (Karch and Goate, 2015).

Astroglia play an active role in synapse formation and synaptic strength regulation and in synchronizing and integrating the many synaptic processes under their control. Astroglia possess receptors and cognate signaling machinery for glutamate, GABA, adenosine, noradrenaline, acetylcholine, and endocannabinoids and respond to neurotransmitter release with calcium signaling (Figure 3). They are thus well equipped to monitor neuronal activity. They also actively secrete gliotransmitters like ATP, glutamate, and D-serine that modulate activity of neuronal receptors (Allen, 2014). In fact, astroglia participate in the “tripartite synapse,” providing a third component to the classical “bipartite” flow between pre- and postsynaptic neurons. Astrocytes thus both sense and modulate synaptic output and may exert an integrating role in the overall computing function of the brain (Allen, 2014; Gittis and Brasier, 2015). Astrocytes are also involved in elimination of synapses via phagocytic receptors MEG10 and MERT (Chung et al., 2013) or via guiding microglial cells to synapses that are pruned involving the complement pathway. Recently, it was shown how this might operate in AD: NF $\kappa$ B is upregulated by A $\beta$  (Lian et al., 2015) and induces C3 release from the astrocytes. C3 binding to the G-protein-linked receptor C3aR on neurons leads to alterations of dendritic morphology and network dysfunction. A C3aR antagonist improves performance of AD mice in the Morris water maze (Lian et al., 2015).

Astrocytes are intimately involved in A $\beta$  catabolism, and A $\beta$  peptides affect their metabolic phenotypes. Astroglia are activated, express inflammatory markers early (Heneka et al., 2005), and display abnormal synchronous Ca<sup>2+</sup> transients over long distances upon A $\beta$  exposure (Kuchibhotla et al., 2009). Elevated Ca<sup>2+</sup> increases 1,4,5-triphosphate (IP<sub>3</sub>) signaling and N-cadherin expression, which underlies astrogliosis. Likely several other signaling pathways are involved as well, and astroglial functionality should be considered when investigating mechanisms of A $\beta$  toxicity in brain.

Local ablation of astrocytes in specific domains of the mouse brain causes neurodegeneration. This is partially because of increased excitotoxicity, as glutamate is no longer efficiently removed. However, the number of astrocytes is not markedly reduced in aging or AD brains (Pelvig et al., 2003) nor in various AD mice (Olabarria et al., 2010). The contribution of astrocytes to AD neurodegeneration is thus clearly more complex than simple “astrodegeneration” and encompasses a whole spectrum of beneficial and harmful reactions (Figure 3).

### Reactive, Degrading, and Atrophic Astrocytes in AD

The “reactive” astrogliosis in AD was noticed by Dr. Alzheimer himself. Reactive astrogliosis is a complex and dynamic response to brain damage (Burda and Sofroniew, 2014) characterized by astrocytic hypertrophy, proliferation, and augmented

expression of intermediate filaments (glial fibrillar acidic protein [GFAP], vimentin, and nestin). It is a physiological response and should, a priori, be considered a protective reaction; however, in AD, it could contribute directly to defective clearance because astrogliosis results in downregulation of Aquaporin 4 interfering with normal glymphatic flow. Astrogliosis is a relatively early event seen before the appearance of A $\beta$  deposits in an AD mouse model (Heneka et al., 2005). PET studies measuring monoamine oxidase B (MaoB) activity in astrocytes confirm that astrogliosis is an early event in patients, with the largest signals seen in prodromal AD (Carter et al., 2012). The relationship between amyloidosis, tau pathology, and astrogliosis remains enigmatic.

Experiments in various mouse models suggest that the astroglial reaction does not lead to astroglial scars as seen in injury models—i.e., the hypertrophy reflects morphological thickening of processes and soma, and the astroglia sustain support for the synapses in their unit. Indeed, attenuation of astrocyte activation (by deleting *Gfap* and *vimentin* genes) accelerates plaque pathogenesis in APP/PS1 mice (Kraft et al., 2013). Moreover, a marked increase in dystrophic neurites was seen, indicating the protective nature of the astrogliosis reaction in the mice. Another group, however, found the opposite effect in similar models (Kamphuis et al., 2015), and whether astrogliosis is beneficial or harmful remains an open question. Indeed, it might be that the answer depends upon the stage of the cellular phase of AD and the brain region affected (Olabarria et al., 2010). The paradoxes in the different studies will probably only be resolved when the progressive and divergent cellular responses in AD are systematically and comprehensively delineated.

Atrophy of astroglia, characterized by reduced branching, is also seen at 1 month in the entorhinal cortex and at 3 months in the prefrontal cortex of triple transgenic AD mice (Kulijewicz-Nawrot et al., 2012). This could lead to defective metabolic support and shrinking of their territorial domain, leading to synaptic dysfunction. Atrophy of astroglia was independently observed in another AD mouse (Beauquis et al., 2014). Studies linking morphological changes to functional alterations are limited. Jo et al. (2014) provide insight in how MaoB (which is upregulated in AD patients [Carter et al., 2012]) increases GABA in reactive, GFAP-positive astroglia in the dentate gyrus of APP/PS1 mice. Release of GABA causes a tonic inhibition of perforant path/dentate granule cell synapses. Two other studies link the G<sub>s</sub>-protein-coupled adenosine receptor A<sub>2A</sub> expressed on astroglia to memory defects in Tau- and APP-mice (Orr et al., 2015). The A<sub>2A</sub>-receptor is increased in AD brain, and these new studies demonstrate that specific antagonists can restore memory (and possibly decrease Tau-pathology) in models for AD.

### Microglia and Inflammation: Beneficial or AD Promoting Effects?

Microglia, the phagocytic cells of the CNS, comprise between 5% and 12% of the cell population in the adult mouse brain (Lawson et al., 1990) and are a recognized part of the cellular phase of AD. Seminal observations by McGeer et al. (1988) demonstrated their intimate association with senile plaques. Since then, their role and, more generally, the role of inflammatory processes (which also involves astrocytes and monocytes)

in the pathogenesis of AD have been intensively investigated [Heneka et al., 2015]). The role of microglia in AD is likely broader than inflammation alone. These cells constitute a dynamic population that samples the local brain environment continuously with highly ramified processes and interacts intimately with neurons (Li et al., 2012). Microglia phagocytose supernumerary neurons during normal development and can directly induce neuronal apoptosis (Salter and Beggs, 2014). More subtle actions, such as synaptic pruning, are also part of their repertoire.

Highly replicable data show complement activation, increases in inflammatory cytokines in the CSF, increased levels of reactive oxygen species, and so on, in AD. Epidemiological studies suggest, in addition, a potential inverse correlation between the use of non-steroidal anti-inflammatory drugs and AD prevalence (Heneka et al., 2015), although interventional studies with anti-inflammatory drugs have all been negative to date. Recent genetic data unequivocally place inflammatory processes as being important in AD. A series of GWASs of increasing size revealed 22 susceptibility loci (Lambert et al., 2013) of which a large group clearly could be ascribed to pathways within innate immune system regulation (Figure 2): *complement receptor 1 (CR1)*, *clusterin*, *CD33*, the *MS4A6-MS4A4* cluster, *ABCA7*, *CD2AP*, *EPHA1*, *HLA-DRB5-DRB1*, *INPP5D*, and *MEF2C* (Karch and Goate, 2015). However, for most of these associations, it is not possible to ascribe changes in biological function to the identified gene or, indeed, whether the polymorphisms are in disequilibrium with other genes that harbor the relevant biological function.

A critically important finding from two groups revealed that rare mutations to the triggering receptor expressed on myeloid cells 2 (*TREM2*) gene significantly increased the risk of late onset AD by about 3- to 5-fold (Guerreiro et al., 2013; Jonsson et al., 2013)—an increase in risk on a par with that of *APOE4*. *TREM2* mutations can also increase the risk of other neurodegenerative diseases, suggesting some commonality in the underlying pathological processes. At the same time, a very large study of late onset AD patients versus controls using whole-genome gene-expression profiling (Zhang et al., 2013) demonstrated that a module grouping innate immunity/microglia related genes correlated best with clinical disease. TYROBP ranked highest as the module's potential regulator. TYROBP, otherwise known as DAP12, encodes the signaling partner for *TREM2*. Thus, using radically different approaches, microglial function and *TREM2* signaling are found to play a major role in SAD.

*TREM2* is a type I transmembrane protein that is expressed on the surface of osteoclasts, immature dendritic cells, macrophages, and, in the CNS, on microglia. *TREM2* ranks as one of highest expressed receptors on microglia, where its expression is >300-fold that of astrocytes (Hickman and El Khoury, 2014). Homozygous, loss-of-function mutations in *TREM2* and its adaptor molecule DAP12 (TYROBP) cause Nasu Hakola disease that is characterized by dementia, white matter loss, and cystic bone lesions. *TREM2* binds to a range of poorly defined ligands such as phospholipids, bacterial products, and cell debris, and receptor binding mediates microglial phagocytosis and promotes an anti-inflammatory cytokine profile. The greatest risk for AD is associated with the R47H variant, which causes loss of function by preventing normal folding of the protein (Klein-

berger et al., 2014). An obvious hypothesis is that *TREM2* is triggering the phagocytosis of A $\beta$  plaques (Kleinberger et al., 2014). Another microglia risk gene, *CD33*, apparently also inhibits uptake of A $\beta$  but via a gain-of-function mechanism, as it is upregulated in SAD (Griciuc et al., 2013).

To assess the effects of *Trem2* in vivo, *APP/PS1* transgenic mice were crossed with *Trem2*<sup>-/-</sup> mice (Ulrich et al., 2014; Wang et al., 2015). The results are divergent. In one study (Ulrich et al., 2014), no effects on A $\beta$  plaque load, cytokine levels, or microglial activation were recorded. The only reported difference was a reduction in the numbers of microglia that clustered around plaques. In the second study, which used a different model—i.e., the 5XFAD mice—*Trem2* knockout resulted in a gene dosage increase in brain A $\beta$  plaques, accompanied by significantly reduced expression of microglial markers of activation and inflammatory cytokines. Thus, this work suggests that *Trem2* plays an important role in the microglial inflammatory response: in the absence of *Trem2*, this response is diminished.

Apart from their phagocytic capacities, microglia express a range of cell-surface receptors that activate innate immune responses. This allows microglia to mount protective responses to damage-associated molecular patterns, such as A $\beta$  aggregates. In particular, a number of studies have implicated the CD36 scavenger receptor and the Toll-like Receptor 4 (TLR-4) in mediating inflammatory cytokine release in response to aggregated A $\beta$ . Data on the role of microglia and inflammatory mediators on the deposition and clearance of A $\beta$  plaques in APP transgenic mice are again mixed, with some researchers finding amelioration and others exacerbation of plaque deposition. Recently, two groups looked at the effects of Interleukin 10 (IL10) on amyloid pathology in APP transgenic mice, with very concordant results. Interleukin 10 is a cytokine that inhibits inflammatory cytokine signaling. Increasing its expression in two APP transgenic mouse models using adeno-associated virus mediated brain expression (Chakrabarty et al., 2015) caused an increase of deposited A $\beta$ <sub>42</sub> and plaque burden. The authors postulate that IL10-mediated increases in astrocyte ApoE production led, in turn, to an elevation of ApoE binding to A $\beta$  plaque that reduced microglial phagocytosis—although this appears to be at variance with data showing binding of ApoE to *Trem2* (Atagi et al., 2015). In the second study (Guillot-Sestier et al., 2015), the investigators crossed *APP/PS1* transgenic mice with *IL10*<sup>-/-</sup> knockout mice. In a similar fashion, but with an opposite polarity, the *IL10* deficiency reduced insoluble, deposited A $\beta$ . Another study investigated the effects of removing microglia from the brain altogether (Grathwohl et al., 2010). Using different gangciclovir dosing regimens, it was possible to remove nearly 100% of the microglia, albeit for a relatively short duration of up to 4 weeks. In no case was the deposition or clearance of amyloid plaques altered.

Collectively, these examples illustrate the conundrum of inflammatory processes: in some situations they can be beneficial, in others, damaging. A more integrated approach seems necessary to investigate the progressive inflammatory and other responses, and a more comprehensive analysis of the role of the different cells (microglia, but also astroglia and infiltrating monocytes) over the different stages of the cellular phase of AD is needed. Indeed, what is beneficial in the beginning could

become destructive later. Finally, inflammation must be placed into the wider context of aging. The AD brain carries the hallmarks of an inflammatory process, but it is also apparent that age itself predisposes to inflammation—a concept that has been coined as “inflammaging.” In a fascinating study, an analysis of gene expression profiles of immune- and inflammation-related genes was conducted across a range of ages of normal and AD brains (Cribbs et al., 2012). This study demonstrated that the increase in expression of immune/inflammation-related genes with age far exceeded that seen in AD versus normal brains. The conclusion from this work is that, with age, the brain becomes predisposed to inflammatory processes, which might explain why age is the greatest risk factor for AD.

### Oligodendrocytes: A Silent Majority

The roles of oligodendrocytes in AD are even less studied than those of astrocytes, although evidence supports their role in the cellular phase of the disease (Ettle et al., 2015). Oligodendrocytes constitute ~75% of the neuroglia cells in the neocortex and are therefore the largest group of non-neuronal cells in the brain (Pelvig et al., 2008). They produce myelin and facilitate neuronal transmission via saltatory action potential conductance over the nodes of Ranvier. They also provide buffer capacity for potassium, metabolic, trophic, and mechanic support to axons. They are considered highly vulnerable cells and show a dramatic (–27%) decrease in number in the aging brain (Pelvig et al., 2008). This cellular loss is reflected in age-associated myelin breakdown as seen in MRI scans of patients from age 50 on (Bartzokis, 2011). Oligodendrocyte progenitor cells, characterized by the expression of the proteoglycan nerve-glia antigen NG2, provide up to 5% of the cells in the mature brain. They are the main proliferating cells and can differentiate into oligodendrocytes, astrocytes, and possibly neurons, although this pluripotency remains an issue of debate (Crawford et al., 2014).

Braak suggested a link between vulnerability of neurons and their myelination stage because the spreading of neurofibrillary tangles recapitulates the pattern of myelination in reverse order (Braak and Braak, 1996). Later differentiating oligodendrocytes ensheath more axons with small axon diameters as seen in late myelinating frontal and temporal lobes (Bartzokis, 2011), and it might be that these make these areas more vulnerable to all kinds of stresses. Bartzokis goes even as far as to propose that the recently evolved extensive myelination of the human brain explains increased vulnerability to neurodegenerative and neuropsychiatric disorders (Bartzokis, 2011). Age-related myelin breakdown as assessed by MRI might indeed result in a progressive disconnection of neural networks and presage AD (Bartzokis, 2011). One of the more compelling arguments to increase focus on oligodendrocytes and possible demyelination is the strongly accentuated age-related breakdown of myelin in AD and in *APOE4* patients (Bartzokis, 2011; Scheltens et al., 1992). At the cellular level, A $\beta$  deposition is linked to focal demyelination. In *APP/PS1* mice, an early demyelination phenotype is seen that is apparently “repaired” by proliferation of oligodendrocyte progenitor cells at 6–8 months (Behrendt et al., 2013). Such proliferation is not observed in AD brain, suggesting that this putative “protective” reaction does not occur in patients (Behrendt et al., 2013). Oligodendrocyte progenitor cells appear

to phagocytose A $\beta$  while A $\beta$  triggers deficits in oligodendrocytes in different mice models (Desai et al., 2010). Other molecular links between AD and oligodendrocytes include their central role in lipid metabolism, discussed above, and the role of BACE1 in neuregulin processing.

### Systems Biology Approaches to Study the Cellular Responses in AD

The complexity emerging from the current analysis of the literature stresses the importance of generating a new theoretical framework that encompasses the many parallel processes that go astray in the brains of AD patients. Typically, systems biology addresses such questions, but the “big data” approaches, excepting the genetic studies, have not yet yielded spectacular new insights into AD pathogenesis. Available studies characterize global alterations of mRNA (Bennett et al., 2014; Bossers et al., 2010; Matarin et al., 2015; Miller et al., 2010; Zhang et al., 2013) or miRNA expression (Lau et al., 2013), thus aggregating all changes in neurons, interneurons, astroglia, oligodendrocytes, and microglia of the analyzed brain block with a consequent loss in resolution. Typically, small expression changes of thousands of genes are recorded. Systems biology uses sophisticated statistical methods to correct for the multiple testing of hypotheses and involves a series of assumptions to increase the power of analysis—for instance, that complex diseases disturb complex biological pathways and that the genes in these pathways are co-regulated (Santiago and Potashkin, 2014). Other assumptions are that key regulators of biological networks can be discovered by identifying regulatory patterns of co-expressed genes or by finding gene expression correlation patterns between different samples (differential co-expression network analysis or DCA) and that proteins in pathways interact directly in a protein complex or in a small protein interaction network (protein-protein interaction or PPI analysis). In AD, real experimental validation of the conclusions of such analyses is mostly lacking. A notable exception is the study on transcription/neuron-restrictive silencer factor (REST/NRF) (Lu et al., 2014). Data mining of transcription profiles from prefrontal cortex of aging humans predicted that REST/NRF was activated in aging brains but less so in AD (Lu et al., 2014). An impressive series of experiments in human brain and in mouse and *C. elegans* genetic models demonstrated that REST expression is neuroprotective, represses genes that promote cell death, increases oxidative stress resistance by de-repressing FOXO1a and SOD1 expression, and maintains neuronal viability in the aging cortex (Lu et al., 2014), providing crucial insight in protective mechanisms that, when malfunctioning, contribute to the risk for neurodegenerative disease.

As stated, other available systems biology studies in the field are more descriptive. Altered regulation of transcription and cell proliferation, changes in energy pathways or mitochondria, upregulation of complement activation, inflammation, lipid metabolism, or downregulation of synaptic transmission are expected (Matarin et al., 2015; Zhang et al., 2013). Further analyses try to identify molecular players that act as drivers or are critical in the overall processes identified. One study found downregulation of tumor suppressors and focused also on changes in oligodendrocyte growth factors. Another study found significant changes in a

human-specific module (group of highly co-expressed genes) that closely correlated with AD progression. The authors speculate that the hubs in this network are important parts of the AD process (Miller et al., 2010). However, none of those hubs has been further validated, and one (FLJ12151) has even been withdrawn as a gene by NCBI. Overall, this study found major links with oligodendrocytes, astroglia, and microglia, re-enforcing the major theme of the current review. As previously mentioned, TYROBP, was identified as a regulator of an immune- and microglia-specific module in a study in 549 brains from late-onset patients and non-demented subjects (Zhang et al., 2013). Although only upregulated by 18%, the authors boldly conclude that TYROBP may be a therapeutic target in prevention of neuronal damage in LOAD, supported by the genetic data found in other studies (vide supra).

There are many reasons why systems biology approaches are currently not yielding novel insights in our field. Further massive data gathering is of questionable use unless systematic, quality controlled, in-depth phenotypic information is associated with the samples (Bennett et al., 2014). However, in addition, massively collecting “omic” data from blocks of brain material probably lacks the granularity required to dissect key aspects of the disease process. The brain is the most complex structure of the human body with an incredible variety of cell types. The lack of cellular resolution makes it very likely that dramatic changes in gene expression in particular cell types or groups of cells will be diluted or lost by the absence of effects in other cell groups. The interpretation of data from brain blocks might be very misleading with regard to the cell type affected in the disorder.

### Single-Cell Populations and Resolution over Different Stages of Disease

A recent systems biology study in AD mice strikingly illustrated the above point by looking separately at astrocytes and microglia (Orre et al., 2014). The immune pathways were much more affected in astrocytes than in microglia, and interestingly, while the recorded alterations in astrocytes were remarkably similar to the alterations recorded in human AD brain, the microglial changes showed much less similarity (Orre et al., 2014). Individual changes in expression of genes were much more pronounced than in the “whole brain block” studies. Tyrobp expression increased almost 10-fold (compared to the overall 18% increase reported in human AD [Zhang et al., 2013]), and Trem2 increased more than 9-fold in the astrocytes while both were conspicuously absent in the list of microglia associated gene changes (Orre et al., 2014). While Trem2 expression is much lower in astrocytes than in microglia (Hickman and El Khoury, 2014), these dramatic and consistent changes in gene expression provide high confidence that they are real and will translate to biologically significant effects. The study clearly illustrates how important it is to provide cellular resolution when performing whole-genome analyses of the brain.

Temporal resolution is also important. In particular, studies on human tissue tend to pool the data without taking into account disease duration or evolution. A study following global gene (Bossers et al., 2010) and miRNA (Lau et al., 2013) expression in the prefrontal cortex over different Braak stages of AD demon-

strated major changes between Braak stage II and III, which is just before or at the onset of plaque and tangle pathology. Unexpectedly, and in contrast with other genome-wide studies, increases in synaptic plasticity genes came up in this early phase, while downregulation was only seen in the late stage of AD (Bossers et al., 2010). The miRNA study demonstrates complex patterns of change in 41 miRNA with a 2- to 3-fold decrease in miRNA132 expression standing out in this analysis (Lau et al., 2013). Unfortunately, the dataset encompasses only 49 brains at the moment, and therefore, statistical power is limited. Expanding this dataset and providing in addition cellular resolution would be extremely helpful for the field.

Real temporal resolution can only come from animal studies where brain samples can be taken from cohorts of mice at different stages of the disease. One such study compared gene expression alterations over time in A $\beta$ - and Tau-models of the disease (Matarin et al., 2015). Mainly immune-related changes were recorded, with an immune module strongly associated to amyloid pathology, and genes related to synaptic function most altered in the Tau model. Cellular resolution, however, was lacking.

### The Future: Single-Cell Biology and Cellular Network Analysis in the Study of AD

Breakthroughs in the study of the biochemical phase of AD are still needed, and an excellent example was the recent elucidation of the atomic structure of  $\gamma$ -secretases, opening a complete new perspective on drug discovery for this important AD target (Bai et al., 2015; De Strooper, 2014). However, it is unlikely that molecular genetics or molecular cell biology will continue to propel our understanding of the disorder in as productive a way as it has over the last 20 years. The major challenge for AD research is now to understand the complex cellular reaction underlying the long prodromal phase of AD. Cellular and temporal resolution will be crucial to map systematically the progressive cellular alterations and to generate a comprehensive cellular theory of the disorder. AD is indeed not a biochemical or molecular problem but a physiological one of disrupted cellular connectivity. The disease can therefore only be fully understood in the context of the complex cellular interactions that maintain homeostasis in the brain—i.e., at the level of the neurovascular and glioneuronal units discussed above.

Spectacular progress in “single-cell biology” (Sandberg, 2014) over the last 3 years makes it possible to map alterations in gene expression in many different cell types in parallel. The strong advantages of a single-cell approach are (1) agnosticism with respect to cell type (allowing also the discovery of new cell types) and (2) that information on the variation in responses over a subpopulation of cells can be evaluated—e.g., do all astroglia, interneurons and so on, respond in the same way and at the same time to A $\beta$  stress? Two recent publications provide proof of concept that genomic expression data from hundreds of individual cells from mouse hippocampus (Zeisel et al., 2015) and human cortex (Darmanis et al., 2015) can be resolved successfully. In addition, spatial transcriptomics is rapidly evolving, and spatial resolution of genome-wide gene expression in cells is becoming possible. A limitation of current approaches is that changes are recorded only at the transcriptional level. However,

cells react by changes in gene expression, and pathway analysis will help to indicate processes affected in the different cells.

A framework could be generated by measuring all changes in the different cell types of the brain over the different Braak stages in a few relevant brain areas, which would constitute a valuable resource for the field and propel AD research into the complex biology of the 21<sup>st</sup> century. Further experimentation and validation of hypotheses at the protein and in vivo level will still use classical reductionist approaches and yield individual insights, but being able to fit such pieces into the time, cell, and network dependent atlas that describes the evolution of Alzheimer's disease will be transformative for the field. Novel ideas will emanate from such an insight and an integrated conceptual framework will provide a scientifically coherent basis for targeted therapeutics that address different elements of the disease in a stage-dependent manner.

#### ACKNOWLEDGMENTS

B.D.S. is a consultant for Janssen Pharmaceuticals and receives research funding from this company. E.K. is an employee of ABBVIE.

#### REFERENCES

- Allen, N.J. (2014). Astrocyte regulation of synaptic behavior. *Annu. Rev. Cell Dev. Biol.* 30, 439–463.
- Ando, K., Brion, J.-P., Stygelbout, V., Suain, V., Authélet, M., Dedecker, R., Chanut, A., Lacor, P., Lavaur, J., Sazdovitch, V., et al. (2013). Clathrin adaptor CALM/PICALM is associated with neurofibrillary tangles and is cleaved in Alzheimer's brains. *Acta Neuropathol.* 125, 861–878.
- Atagi, Y., Liu, C.-C., Painter, M.M., Chen, X.-F., Verbeeck, C., Zheng, H., Li, X., Rademakers, R., Kang, S.S., Xu, H., et al. (2015). Apolipoprotein E is a Ligand for Triggering Receptor Expressed on Myeloid Cells 2 (TREM2). *J. Biol. Chem.* 290, 26043–26050.
- Attems, J., and Jellinger, K.A. (2014). The overlap between vascular disease and Alzheimer's disease—lessons from pathology. *BMC Med.* 12, 206.
- Bai, X.C., Yan, C., Yang, G., Lu, P., Ma, D., Sun, L., Zhou, R., Scheres, S.H.W., and Shi, Y. (2015). An atomic structure of human  $\gamma$ -secretase. *Nature* 525, 212–217.
- Bakker, A., Krauss, G.L., Albert, M.S., Speck, C.L., Jones, L.R., Stark, C.E., Yassa, M.A., Bassett, S.S., Shelton, A.L., and Gallagher, M. (2012). Reduction of hippocampal hyperactivity improves cognition in amnesic mild cognitive impairment. *Neuron* 74, 467–474.
- Bartzokis, G. (2011). Alzheimer's disease as homeostatic responses to age-related myelin breakdown. *Neurobiol. Aging* 32, 1341–1371.
- Beauquis, J., Vinuesa, A., Pomilio, C., Pavia, P., Galván, V., and Saravia, F. (2014). Neuronal and glial alterations, increased anxiety, and cognitive impairment before hippocampal amyloid deposition in PDAPP mice, model of Alzheimer's disease. *Hippocampus* 24, 257–269.
- Behrendt, G., Baer, K., Buffo, A., Curtis, M.A., Faull, R.L., Rees, M.I., Götz, M., and Dimou, L. (2013). Dynamic changes in myelin aberrations and oligodendrocyte generation in chronic amyloidosis in mice and men. *Glia* 61, 273–286.
- Bell, R.D., Winkler, E.A., Singh, I., Sagare, A.P., Deane, R., Wu, Z., Holtzman, D.M., Betsholtz, C., Armulik, A., Sallstrom, J., et al. (2012). Apolipoprotein E controls cerebrovascular integrity via cyclophilin A. *Nature* 485, 512–516.
- Benilova, I., Karran, E., and De Strooper, B. (2012). The toxic A $\beta$  oligomer and Alzheimer's disease: an emperor in need of clothes. *Nat. Neurosci.* 15, 349–357.
- Bennett, D.A., Yu, L., and De Jager, P.L. (2014). Building a pipeline to discover and validate novel therapeutic targets and lead compounds for Alzheimer's disease. *Biochem. Pharmacol.* 88, 617–630.
- Bossers, K., Wirz, K.T.S., Meerhoff, G.F., Essing, A.H.W., van Dongen, J.W., Houba, P., Kruse, C.G., Verhaagen, J., and Swaab, D.F. (2010). Concerted changes in transcripts in the prefrontal cortex precede neuropathology in Alzheimer's disease. *Brain* 133, 3699–3723.
- Braak, H., and Braak, E. (1995). Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol. Aging* 16, 271–278, discussion 278–284.
- Braak, H., and Braak, E. (1996). Development of Alzheimer-related neurofibrillary changes in the neocortex inversely recapitulates cortical myelogenesis. *Acta Neuropathol.* 92, 197–201.
- Braak, H., and Del Tredici, K. (2011). The pathological process underlying Alzheimer's disease in individuals under thirty. *Acta Neuropathol.* 121, 171–181.
- Burda, J.E., and Sofroniew, M.V. (2014). Reactive gliosis and the multicellular response to CNS damage and disease. *Neuron* 81, 229–248.
- Bushong, E.A., Martone, M.E., Jones, Y.Z., and Ellisman, M.H. (2002). Protoplasmic astrocytes in CA1 stratum radiatum occupy separate anatomical domains. *J. Neurosci.* 22, 183–192.
- Carrillo, M.C., Dean, R.A., Nicolas, F., Miller, D.S., Berman, R., Khachaturian, Z., Bain, L.J., Schindler, R., and Knopman, D.; Alzheimer's Association Research Roundtable (2013). Revisiting the framework of the National Institute on Aging-Alzheimer's Association diagnostic criteria. *Alzheimers Dement.* 9, 594–601.
- Carter, S.F., Schöll, M., Almkvist, O., Wall, A., Engler, H., Långström, B., and Nordberg, A. (2012). Evidence for astrocytosis in prodromal Alzheimer disease provided by 11C-deuterium-L-deprenyl: a multitracer PET paradigm combining 11C-Pittsburgh compound B and 18F-FDG. *J. Nucl. Med.* 53, 37–46.
- Chakrabarty, P., Li, A., Ceballos-Diaz, C., Eddy, J.A., Funk, C.C., Moore, B., DiNunno, N., Rosario, A.M., Cruz, P.E., Verbeeck, C., et al. (2015). IL-10 alters immunoproteostasis in APP mice, increasing plaque burden and worsening cognitive behavior. *Neuron* 85, 519–533.
- Chávez-Gutiérrez, L., Bammens, L., Benilova, I., Vandersteen, A., Benurwar, M., Borgers, M., Lismont, S., Zhou, L., Van Cleynebreugel, S., Esselmann, H., et al. (2012). The mechanism of  $\gamma$ -Secretase dysfunction in familial Alzheimer disease. *EMBO J.* 31, 2261–2274.
- Cheng, J., and Ji, D. (2013). Rigid firing sequences undermine spatial memory codes in a neurodegenerative mouse model. *eLife* 2, e00647.
- Chung, W.-S., Clarke, L.E., Wang, G.X., Stafford, B.K., Sher, A., Chakrabarty, C., Joung, J., Foo, L.C., Thompson, A., Chen, C., et al. (2013). Astrocytes mediate synapse elimination through MEGF10 and MERTK pathways. *Nature* 504, 394–400.
- Crary, J.F., Trojanowski, J.Q., Schneider, J.A., Abisambra, J.F., Abner, E.L., Alafuzoff, I., Arnold, S.E., Attems, J., Beach, T.G., Bigio, E.H., et al. (2014). Primary age-related tauopathy (PART): a common pathology associated with human aging. *Acta Neuropathol.* 128, 755–766.
- Crawford, A.H., Stockley, J.H., Tripathi, R.B., Richardson, W.D., and Franklin, R.J.M. (2014). Oligodendrocyte progenitors: adult stem cells of the central nervous system? *Exp. Neurol.* 260, 50–55.
- Cribbs, D.H., Berchtold, N.C., Perreau, V., Coleman, P.D., Rogers, J., Tenner, A.J., and Cotman, C.W. (2012). Extensive innate immune gene activation accompanies brain aging, increasing vulnerability to cognitive decline and neurodegeneration: a microarray study. *J. Neuroinflammation* 9, 179.
- Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. USA* 112, 7285–7290.
- De Strooper, B. (2010). Proteases and proteolysis in Alzheimer disease: a multifactorial view on the disease process. *Physiol. Rev.* 90, 465–494.
- De Strooper, B. (2014). Lessons from a failed  $\gamma$ -secretase Alzheimer trial. *Cell* 159, 721–726.
- Deane, R., Singh, I., Sagare, A.P., Bell, R.D., Ross, N.T., LaRue, B., Love, R., Perry, S., Paquette, N., Deane, R.J., et al. (2012). A multimodal RAGE-specific inhibitor reduces amyloid  $\beta$ -mediated brain disorder in a mouse model of Alzheimer disease. *J. Clin. Invest.* 122, 1377–1392.

- Desai, M.K., Mastrangelo, M.A., Ryan, D.A., Sudol, K.L., Narrow, W.C., and Bowers, W.J. (2010). Early oligodendrocyte/myelin pathology in Alzheimer's disease mice constitutes a novel therapeutic target. *Am. J. Pathol.* *177*, 1422–1435.
- Dubois, B., Feldman, H.H., Jacova, C., Hampel, H., Molinuevo, J.L., Blennow, K., DeKosky, S.T., Gauthier, S., Selkoe, D., Bateman, R., et al. (2014). Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol.* *13*, 614–629.
- Erick, M.J., Yu, T., Chung, C., and Lieberman, A.P. (2012). Impaired proteolysis underlies autophagic dysfunction in Niemann-Pick type C disease. *Hum. Mol. Genet.* *21*, 4876–4887.
- Ettle, B., Schlachetzki, J.C.M., and Winkler, J. (2015). Oligodendroglia and Myelin in Neurodegenerative Diseases: More Than Just Bystanders? *Mol. Neurobiol.* Published online May 13, 2015. <http://dx.doi.org/10.1007/s12035-015-9205-3>.
- Gittis, A.H., and Brasier, D.J. (2015). NEUROSCIENCE. Astrocytes tell neurons when to listen up. *Science* *349*, 690–691.
- Grathwohl, S.A., Kälin, R.E., Bolmont, T., Prokop, S., Winkelmann, G., Kaeser, S., Odenthal, J., Radde, R., Eldh, T., Gandy, S., et al. (2010). Formation and maintenance of Alzheimer's disease beta-amyloid plaques in the absence of microglia. *Nat. Neurosci.* *12*, 1361–1363.
- Griciuc, A., Serrano-Pozo, A., Parrado, A.R., Lesinski, A.N., Asselin, C.N., Mullin, K., Hooli, B., Choi, S.H., Hyman, B.T., and Tanzi, R.E. (2013). Alzheimer's disease risk gene CD33 inhibits microglial uptake of amyloid beta. *Neuron* *78*, 631–643.
- Guerreiro, R., Wojtas, A., Bras, J., Carrasquillo, M., Rogava, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J.S.K., Younkin, S., et al.; Alzheimer Genetic Analysis Group (2013). TREM2 variants in Alzheimer's disease. *N. Engl. J. Med.* *368*, 117–127.
- Guillot-Sestier, M.-V., Doty, K.R., Gate, D., Rodriguez, J., Jr., Leung, B.P., Rezaei-Zadeh, K., and Town, T. (2015). IL10 deficiency rebalances innate immunity to mitigate Alzheimer-like pathology. *Neuron* *85*, 534–548.
- Hardy, J., and Selkoe, D.J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* *297*, 353–356.
- Heneka, M.T., Sastre, M., Dumitrescu-Ozimek, L., Dewachter, I., Walter, J., Klockgether, T., and Van Leuven, F. (2005). Focal glial activation coincides with increased BACE1 activation and precedes amyloid plaque deposition in APP[V717I] transgenic mice. *J. Neuroinflammation* *2*, 22.
- Heneka, M.T., Carson, M.J., El Khoury, J., Landreth, G.E., Brosseron, F., Feinstein, D.L., Jacobs, A.H., Wyss-Coray, T., Vitorica, J., Ransohoff, R.M., et al. (2015). Neuroinflammation in Alzheimer's disease. *Lancet Neurol.* *14*, 388–405.
- Hickman, S.E., and El Khoury, J. (2014). TREM2 and the neuroimmunology of Alzheimer's disease. *Biochem. Pharmacol.* *88*, 495–498.
- Illiff, J.J., Chen, M.J., Plog, B.A., Zeppenfeld, D.M., Soltero, M., Yang, L., Singh, I., Deane, R., and Nedergaard, M. (2014). Impairment of glymphatic pathway function promotes tau pathology after traumatic brain injury. *J. Neurosci.* *34*, 16180–16193.
- Jo, S., Yarishkin, O., Hwang, Y.J., Chun, Y.E., Park, M., Woo, D.H., Bae, J.Y., Kim, T., Lee, J., Chun, H., et al. (2014). GABA from reactive astrocytes impairs memory in mouse models of Alzheimer's disease. *Nat. Med.* *20*, 886–896.
- Jonsson, T., Stefansson, H., Steinberg, S., Jonsdottir, I., Jonsson, P.V., Snaedal, J., Bjornsson, S., Huttenlocher, J., Levey, A.I., Lah, J.J., et al. (2013). Variant of TREM2 associated with the risk of Alzheimer's disease. *N. Engl. J. Med.* *368*, 107–116.
- Kamphuis, W., Kooijman, L., Orre, M., Stassen, O., Pekny, M., and Hol, E.M. (2015). GFAP and vimentin deficiency alters gene expression in astrocytes and microglia in wild-type mice and changes the transcriptional response of reactive glia in mouse model for Alzheimer's disease. *Glia* *63*, 1036–1056.
- Karch, C.M., and Goate, A.M. (2015). Alzheimer's disease risk genes and mechanisms of disease pathogenesis. *Biol. Psychiatry* *77*, 43–51.
- Karran, E., Mercken, M., and De Strooper, B. (2011). The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nat. Rev. Drug Discov.* *10*, 698–712.
- Khan, U.A., Liu, L., Provenzano, F.A., Berman, D.E., Profaci, C.P., Sloan, R., Mayeux, R., Duff, K.E., and Small, S.A. (2014). Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nat. Neurosci.* *17*, 304–311.
- Kleinberger, G., Yamanishi, Y., Suárez-Calvet, M., Czirr, E., Lohmann, E., Cuyvers, E., Struyfs, H., Pettkus, N., Wenninger-Weinzierl, A., Mazaheri, F., et al. (2014). TREM2 mutations implicated in neurodegeneration impair cell surface transport and phagocytosis. *Sci. Transl. Med.* *6*, 243ra86.
- Komatsu, M., Waguri, S., Chiba, T., Murata, S., Iwata, J., Tanida, I., Ueno, T., Koike, M., Uchiyama, Y., Kominami, E., and Tanaka, K. (2006). Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature* *441*, 880–884.
- Kraft, A.W., Hu, X., Yoon, H., Yan, P., Xiao, Q., Wang, Y., Gil, S.C., Brown, J., Wilhelmsson, U., Restivo, J.L., et al. (2013). Attenuating astrocyte activation accelerates plaque pathogenesis in APP/PS1 mice. *FASEB J.* *27*, 187–198.
- Kuchibhotla, K.V., Lattarulo, C.R., Hyman, B.T., and Bacskai, B.J. (2009). Synchronous hyperactivity and intercellular calcium waves in astrocytes in Alzheimer mice. *Science* *323*, 1211–1215.
- Kulijewicz-Nawrot, M., Verkhratsky, A., Chvátal, A., Syková, E., and Rodríguez, J.J. (2012). Astrocytic cytoskeletal atrophy in the medial prefrontal cortex of a triple transgenic mouse model of Alzheimer's disease. *J. Anat.* *221*, 252–262.
- Kuperstein, I., Broersen, K., Benilova, I., Rozenski, J., Jonckheere, W., Debupeap, M., Vandersteen, A., Segers-Nolten, I., Van Der Werf, K., Subramaniam, V., et al. (2010). Neurotoxicity of Alzheimer's disease A $\beta$  peptides is induced by small changes in the A $\beta$ 42 to A $\beta$ 40 ratio. *EMBO J.* *29*, 3408–3420.
- Labbadia, J., and Morimoto, R.I. (2015). The biology of proteostasis in aging and disease. *Annu. Rev. Biochem.* *84*, 435–464.
- Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B., et al.; European Alzheimer's Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* *45*, 1452–1458.
- Lau, P., Bossers, K., Janky, R., Salta, E., Frigerio, C.S., Barbash, S., Rothman, R., Sierksma, A.S.R., Thathiah, A., Greenberg, D., et al. (2013). Alteration of the microRNA network during the progression of Alzheimer's disease. *EMBO Mol. Med.* *5*, 1613–1634.
- Lawson, L.J., Perry, V.H., Dri, P., and Gordon, S. (1990). Heterogeneity in the distribution and morphology of microglia in the normal adult mouse brain. *Neuroscience* *39*, 151–170.
- Li, Y., Du, X.F., Liu, C.S., Wen, Z.L., and Du, J.L. (2012). Reciprocal regulation between resting microglial dynamics and neuronal activity in vivo. *Dev. Cell* *23*, 1189–1202.
- Lian, H., Yang, L., Cole, A., Sun, L., Chiang, A.C., Fowler, S.W., Shim, D.J., Rodriguez-Rivera, J., Tagliavolterra, G., Jankowsky, J.L., et al. (2015). NF $\kappa$ B-activated astroglial release of complement C3 compromises neuronal morphology and function associated with Alzheimer's disease. *Neuron* *85*, 101–115.
- Lu, T., Aron, L., Zullo, J., Pan, Y., Kim, H., Chen, Y., Yang, T.-H., Kim, H.-M., Drake, D., Liu, X.S., et al. (2014). REST and stress resistance in ageing and Alzheimer's disease. *Nature* *507*, 448–454.
- Matarin, M., Salih, D.A., Yasvoina, M., Cummings, D.M., Guelfi, S., Liu, W., Nahaboo Solim, M.A., Moens, T.G., Paublete, R.M., Ali, S.S., et al. (2015). A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. *Cell Rep.* *10*, 633–644.
- McGeer, P.L., Itagaki, S., Tago, H., and McGeer, E.G. (1988). Occurrence of HLA-DR reactive microglia in Alzheimer's disease. *Ann. N Y Acad. Sci.* *540*, 319–323.

- Menkes-Caspi, N., Yamin, H.G., Kellner, V., Spires-Jones, T.L., Cohen, D., and Stern, E.A. (2015). Pathological tau disrupts ongoing network activity. *Neuron* 85, 959–966.
- Mesulam, M.-M. (1999). Neuroplasticity failure in Alzheimer's disease: bridging the gap between plaques and tangles. *Neuron* 24, 521–529.
- Miller, J.A., Horvath, S., and Geschwind, D.H. (2010). Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proc. Natl. Acad. Sci. USA* 107, 12698–12703.
- Mufson, E.J., Mahady, L., Waters, D., Counts, S.E., Perez, S.E., DeKosky, S.T., Ginsberg, S.D., Ikonomic, M.D., Scheff, S.W., and Binder, L.I. (2015). Hippocampal plasticity during the progression of Alzheimer's disease. *Neuroscience* 309, 51–67.
- Nilsson, P., Saito, T., and Saido, T.C. (2014). New mouse model of Alzheimer's. *ACS Chem. Neurosci.* 5, 499–502.
- Nixon, R.A. (2013). The role of autophagy in neurodegenerative disease. *Nat. Med.* 19, 983–997.
- Olabarria, M., Noristani, H.N., Verkhratsky, A., and Rodríguez, J.J. (2010). Concomitant astroglial atrophy and astrogliosis in a triple transgenic animal model of Alzheimer's disease. *Glia* 58, 831–838.
- Orr, A.G., Hsiao, E.C., Wang, M.M., Ho, K., Kim, D.H., Wang, X., Guo, W., Kang, J., Yu, G.-Q., Adame, A., et al. (2015). Astrocytic adenosine receptor A2A and Gs-coupled signaling regulate memory. *Nat. Neurosci.* 18, 423–434.
- Orre, M., Kamphuis, W., Osborn, L.M., Jansen, A.H.P., Kooijman, L., Bossers, K., and Hol, E.M. (2014). Isolation of glia from Alzheimer's mice reveals inflammation and dysfunction. *Neurobiol. Aging* 35, 2746–2760.
- Palop, J.J., and Mucke, L. (2010). Amyloid- $\beta$ -induced neuronal dysfunction in Alzheimer's disease: from synapses toward neural networks. *Nat. Neurosci.* 13, 812–818.
- Pelvig, D.P., Pakkenberg, H., Regeur, L., Oster, S., and Pakkenberg, B. (2003). Neocortical glial cell numbers in Alzheimer's disease. A stereological study. *Dement. Geriatr. Cogn. Disord.* 16, 212–219.
- Pelvig, D.P., Pakkenberg, H., Stark, A.K., and Pakkenberg, B. (2008). Neocortical glial cell numbers in human brains. *Neurobiol. Aging* 29, 1754–1762.
- Protas, H.D., Chen, K., Langbaum, J.B., Fleisher, A.S., Alexander, G.E., Lee, W., Bandy, D., de Leon, M.J., Mosconi, L., Buckley, S., et al. (2013). Posterior cingulate glucose metabolism, hippocampal glucose metabolism, and hippocampal volume in cognitively normal, late-middle-aged persons at 3 levels of genetic risk for Alzheimer disease. *JAMA Neurol.* 70, 320–325.
- Saito, T., Suemoto, T., Brouwers, N., Slegers, K., Funamoto, S., Mihira, N., Matsuba, Y., Yamada, K., Nilsson, P., Takano, J., et al. (2011). Potent amyloidogenicity and pathogenicity of A $\beta$ 43. *Nat. Neurosci.* 14, 1023–1032.
- Salter, M.W., and Beggs, S. (2014). Sublime microglia: expanding roles for the guardians of the CNS. *Cell* 158, 15–24.
- Sandberg, R. (2014). Entering the era of single-cell transcriptomics in biology and medicine. *Nat. Methods* 11, 22–24.
- Santiago, J.A., and Potashkin, J.A. (2014). A network approach to clinical intervention in neurodegenerative diseases. *Trends Mol. Med.* 20, 694–703.
- Scheltens, P., Barkhof, F., Valk, J., Algra, P.R., van der Hoop, R.G., Nauta, J., and Wolters, E.C. (1992). White matter lesions on magnetic resonance imaging in clinically diagnosed Alzheimer's disease. Evidence for heterogeneity. *Brain* 115, 735–748.
- Small, S.A., and Duff, K. (2008). Linking A $\beta$  and tau in late-onset Alzheimer's disease: a dual pathway hypothesis. *Neuron* 60, 534–542.
- Szaruga, M., Veugelen, S., Benurwar, M., Lismont, S., Sepulveda-Falla, D., Lleo, A., Ryan, N.S., Lashley, T., Fox, N.C., Murayama, S., et al. (2015). Qualitative changes in human  $\gamma$ -secretase underlie familial Alzheimer's disease. *J. Exp. Med.* 212, 2003–2013.
- Tarasoff-Conway, J.M., Carare, R.O., Osorio, R.S., Glodzik, L., Butler, T., Fieremans, E., Axel, L., Rusinek, H., Nicholson, C., Zlokovic, B.V., et al. (2015). Clearance systems in the brain—implications for Alzheimer disease. *Nat. Rev. Neurol.* 11, 457–470.
- Thal, D.R., Rüb, U., Orantes, M., and Braak, H. (2002). Phases of A $\beta$  deposition in the human brain and its relevance for the development of AD. *Neurology* 58, 1791–1800.
- Ulrich, J.D., Finn, M.B., Wang, Y., Shen, A., Mahan, T.E., Jiang, H., Stewart, F.R., Piccio, L., Colonna, M., and Holtzman, D.M. (2014). Altered microglial response to A $\beta$  plaques in APPPS1-21 mice heterozygous for TREM2. *Mol. Neurodegener.* 9, 20.
- Vergheze, P.B., Castellano, J.M., Garai, K., Wang, Y., Jiang, H., Shah, A., Bu, G., Frieden, C., and Holtzman, D.M. (2013). ApoE influences amyloid- $\beta$  (A $\beta$ ) clearance despite minimal apoE/A $\beta$  association in physiological conditions. *Proc. Natl. Acad. Sci. USA* 110, E1807–E1816.
- Verret, L., Mann, E.O., Hang, G.B., Barth, A.M.I., Cobos, I., Ho, K., Devidze, N., Masliah, E., Kreitzer, A.C., Mody, I., et al. (2012). Inhibitory interneuron deficit links altered network activity and cognitive dysfunction in Alzheimer model. *Cell* 149, 708–721.
- Walker, L.C., and Jucker, M. (2015). Neurodegenerative diseases: expanding the prion concept. *Annu. Rev. Neurosci.* 38, 87–103.
- Wang, Y., Cella, M., Mallinson, K., Ulrich, J.D., Young, K.L., Robinette, M.L., Gilfillan, S., Krishnan, G.M., Sudhakar, S., Zinselmeyer, B.H., et al. (2015). TREM2 lipid sensing sustains the microglial response in an Alzheimer's disease model. *Cell* 160, 1061–1071.
- Winkler, E.A., Sagare, A.P., and Zlokovic, B.V. (2014). The pericyte: a forgotten cell type with important implications for Alzheimer's disease? *Brain Pathol.* 24, 371–386.
- Winkler, E.A., Nishida, Y., Sagare, A.P., Rege, S.V., Bell, R.D., Perlmutter, D., Sengillo, J.D., Hillman, S., Kong, P., Nelson, A.R., et al. (2015). GLUT1 reductions exacerbate Alzheimer's disease vasculo-neuronal dysfunction and degeneration. *Nat. Neurosci.* 18, 521–530.
- Yamada, K., Holth, J.K., Liao, F., Stewart, F.R., Mahan, T.E., Jiang, H., Cirrito, J.R., Patel, T.K., Hochgräfe, K., Mandelkow, E.-M., and Holtzman, D.M. (2014). Neuronal activity regulates extracellular tau in vivo. *J. Exp. Med.* 211, 387–393.
- Zahs, K.R., and Ashe, K.H. (2010). 'Too much good news' - are Alzheimer mouse models trying to tell us how to prevent, not cure, Alzheimer's disease? *Trends Neurosci.* 33, 381–389.
- Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Jureus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347, 1138–1142.
- Zhang, B., Gaiteri, C., Bodea, L.-G., Wang, Z., McElwee, J., Podtelezhnikov, A.A., Zhang, C., Xie, T., Tran, L., Dobrin, R., et al. (2013). Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* 153, 707–720.
- Zhao, Z., Sagare, A.P., Ma, Q., Halliday, M.R., Kong, P., Kisler, K., Winkler, E.A., Ramanathan, A., Kanekiyo, T., Bu, G., et al. (2015). Central role for PICALM in amyloid- $\beta$  blood-brain barrier transcytosis and clearance. *Nat. Neurosci.* 18, 978–987.
- Zlokovic, B.V. (2011). Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders. *Nat. Rev. Neurosci.* 12, 723–738.



# OUR NETWORK IS YOUR NETWORK

## With Cell Press Webinars, our network is your network!

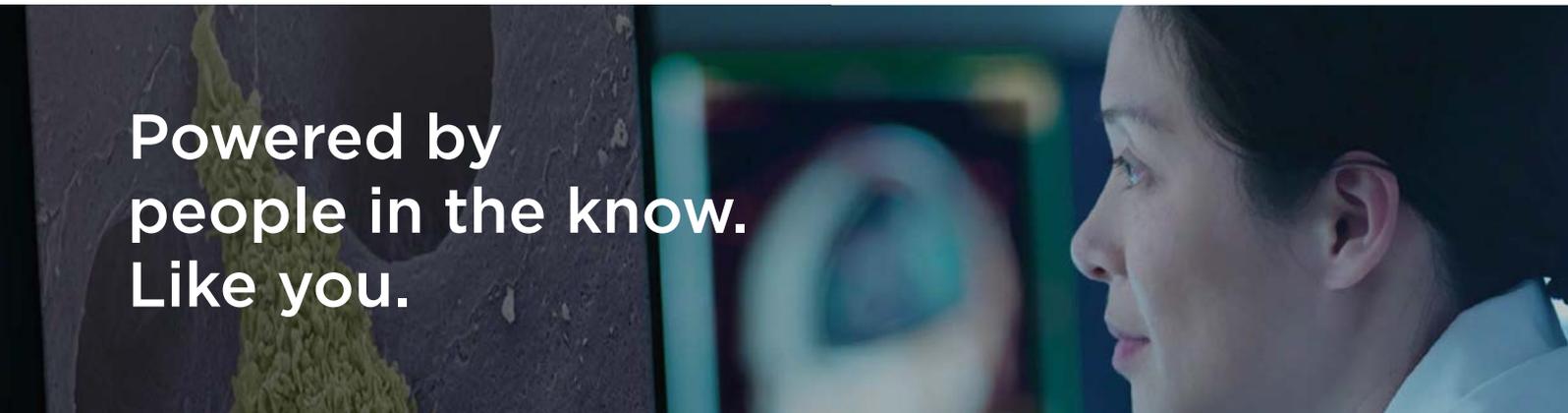
Cell Press Webinars give you access to hot topics in emerging research and the application of new technology.

Watch essential, need-to-know webcasts via live streaming or on demand from the comfort and convenience of your office, lab, or home.

Need-to-know topics, editorially curated

World-class presenters, experts in their field

Moderated by Cell Press editors



Powered by  
people in the know.  
Like you.

Tap into our network today!  
Visit [www.cell.com/webinars](http://www.cell.com/webinars)

**CellPress**  
Webinars

# Mitochondrial Dynamics Controls T Cell Fate through Metabolic Programming

Michael D. Buck,<sup>1,2</sup> David O'Sullivan,<sup>1</sup> Ramon I. Klein Geltink,<sup>1</sup> Jonathan D. Curtis,<sup>1</sup> Chih-Hao Chang,<sup>2</sup> David E. Sanin,<sup>1</sup> Jing Qiu,<sup>1,2</sup> Oliver Kretz,<sup>3,4</sup> Daniel Braas,<sup>5</sup> Gerritje J.W. van der Windt,<sup>6</sup> Qiongyu Chen,<sup>2</sup> Stanley Ching-Cheng Huang,<sup>2</sup> Christina M. O'Neill,<sup>2</sup> Brian T. Edelson,<sup>2</sup> Edward J. Pearce,<sup>1,7</sup> Hiromi Sesaki,<sup>8</sup> Tobias B. Huber,<sup>3,9</sup> Angelika S. Rambold,<sup>10,11</sup> and Erika L. Pearce<sup>1,\*</sup>

<sup>1</sup>Department of Immunometabolism, Max Planck Institute of Immunobiology and Epigenetics, 79108 Freiburg, Germany

<sup>2</sup>Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA

<sup>3</sup>Renal Division, University Medical Center Freiburg, 79106 Freiburg, Germany

<sup>4</sup>Department of Neuroanatomy, University of Freiburg, 79104 Freiburg, Germany

<sup>5</sup>University of California Los Angeles Metabolomics Center, Los Angeles, CA 90095, USA

<sup>6</sup>Academic Medical Center, 1105 AZ Amsterdam, the Netherlands

<sup>7</sup>Faculty of Biology, University of Freiburg, 79104 Freiburg, Germany

<sup>8</sup>Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

<sup>9</sup>BIOS Centre for Biological Signaling Studies, 79104 Freiburg, Germany

<sup>10</sup>Center for Chronic Immunodeficiency, University Medical Center Freiburg and University of Freiburg, 79106 Freiburg, Germany

<sup>11</sup>Department of Developmental Immunology, Max Planck Institute of Immunobiology and Epigenetics, 79108 Freiburg, Germany

\*Correspondence: pearce@ie-freiburg.mpg.de

<http://dx.doi.org/10.1016/j.cell.2016.05.035>

## SUMMARY

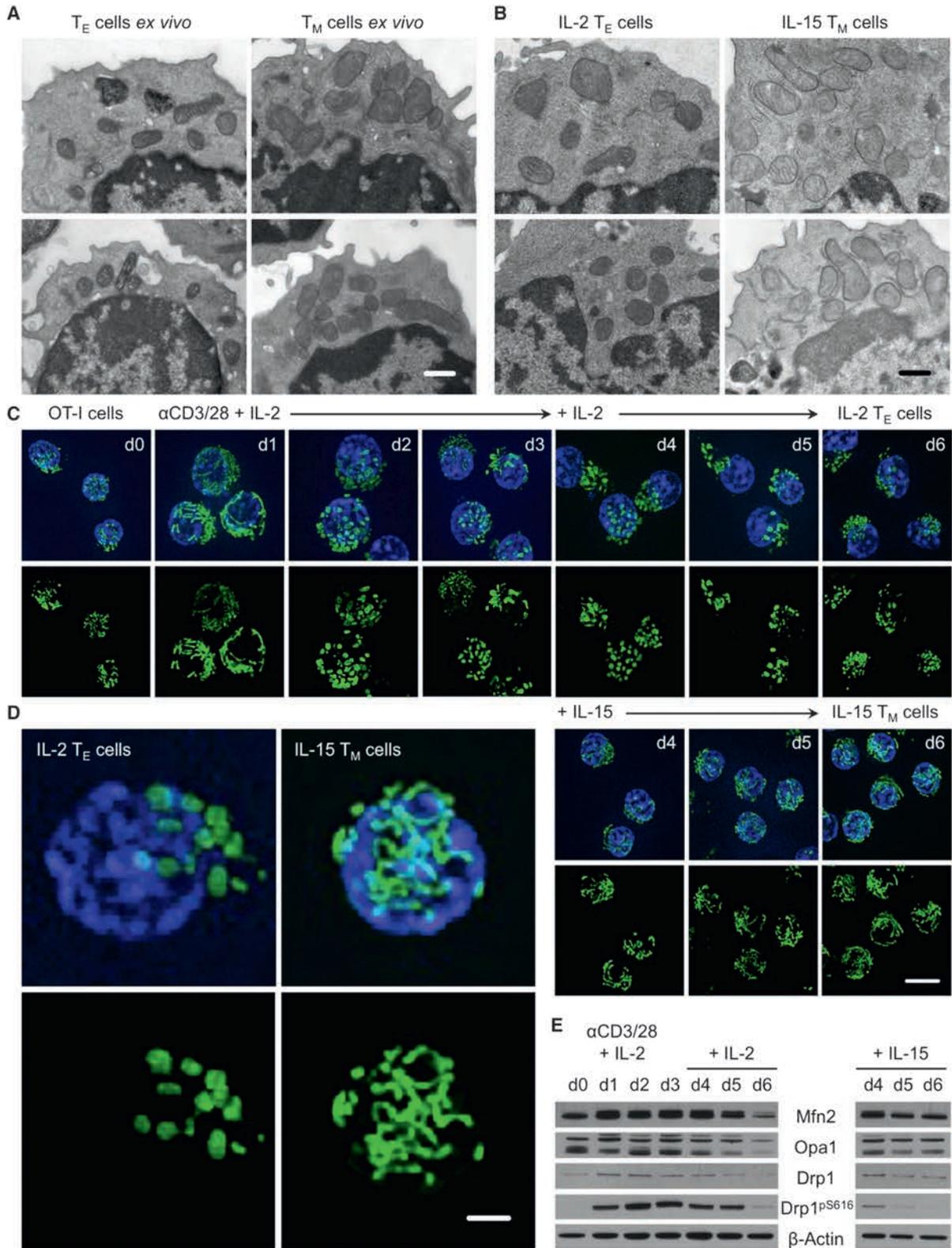
Activated effector T ( $T_E$ ) cells augment anabolic pathways of metabolism, such as aerobic glycolysis, while memory T ( $T_M$ ) cells engage catabolic pathways, like fatty acid oxidation (FAO). However, signals that drive these differences remain unclear. Mitochondria are metabolic organelles that actively transform their ultrastructure. Therefore, we questioned whether mitochondrial dynamics controls T cell metabolism. We show that  $T_E$  cells have punctate mitochondria, while  $T_M$  cells maintain fused networks. The fusion protein Opa1 is required for  $T_M$ , but not  $T_E$  cells after infection, and enforcing fusion in  $T_E$  cells imposes  $T_M$  cell characteristics and enhances antitumor function. Our data suggest that, by altering cristae morphology, fusion in  $T_M$  cells configures electron transport chain (ETC) complex associations favoring oxidative phosphorylation (OXPHOS) and FAO, while fission in  $T_E$  cells leads to cristae expansion, reducing ETC efficiency and promoting aerobic glycolysis. Thus, mitochondrial remodeling is a signaling mechanism that instructs T cell metabolic programming.

## INTRODUCTION

T cells mediate protective immunity against pathogens and cancer and possess the unique ability to proliferate at an unparalleled rate in an adult organism. In this regard, one naive T ( $T_N$ ) cell can clonally expand into millions of “armed” effector T ( $T_E$ ) cells in just a few days (Williams and Bevan, 2007). Concomitant with T cell activation is the engagement of aerobic glycolysis and

elevated OXPHOS (Chang et al., 2013; Sena et al., 2013), the former of which is characteristic of the Warburg effect shared by tumor cells and unicellular organisms (Vander Heiden et al., 2009). Once antigen is cleared, most  $T_E$  cells die, but a subset of long-lived memory T ( $T_M$ ) cells persists with enhanced mitochondrial capacity marked by a reliance on FAO to fuel OXPHOS, which equips them to rapidly respond should infection or cancer recur (Pearce et al., 2013). These extensive changes in phenotype and function of T cells go along with a dynamic metabolic range (MacIver et al., 2013; Buck et al., 2015). Failure to engage specific metabolic programs impairs the function and differentiation of T cells. As such, T cells represent an amenable system to study changes in cell metabolism that occur as part of normal development and not as a result of transformation. Establishing the precise reasons why and how these and other cells emphasize one particular metabolic pathway over another remains a challenge.

Mitochondria are essential hubs of metabolic activity, antiviral responses, and cell death that constantly remodel their structure via nuclear-encoded GTPases (Nunnari and Suomalainen, 2012). Mitochondrial fission generates discrete and fragmented mitochondria that can increase reactive oxygen species (ROS) production (Yu et al., 2006), facilitate mitophagy (Frank et al., 2012; Toyama et al., 2016), accelerate cell proliferation (Taguchi et al., 2007), and mediate apoptosis (Youle and Karbowski, 2005). Dynamin-related protein 1 (Drp1) is a cytosolic protein that translocates to the outer mitochondrial membrane (OMM) upon phosphorylation to scission mitochondria. Fusion of mitochondria into linear or tubular networks limits deleterious mutations in mtDNA (Santel et al., 2003), induces supercomplexes of the ETC maximizing OXPHOS activity (Cogliati et al., 2013; Mishra et al., 2014), and enhances ER interactions important for  $Ca^{2+}$  flux (de Brito and Scorrano, 2008). In addition, mitochondria elongate as a survival mechanism in response to nutrient starvation and stress, linking fusion to cell longevity



(legend on next page)

(Gomes et al., 2011; Rambold et al., 2011; Friedman and Nunnari, 2014). OMM fusion is mediated by mitofusin 1 and 2 (Mfn1, Mfn2), while inner membrane fusion is controlled by optic atrophy 1 (Opa1). Total deletion in any of these proteins is embryonically lethal and mutations in the genes that encode them underlie the cause of several human diseases (Chan, 2012; Archer, 2014).

Mitochondrial membrane remodeling is acutely responsive to changes in cell metabolism (Mishra and Chan, 2016; Wai and Langer, 2016), but whether it instructs metabolic pathway utilization has been inferred but not extensively studied. In general, deletion of any of the dynamics machinery perturbs OXPHOS and glycolytic rates at baseline (Liesa and Shirihai, 2013). Tissue-specific deletion of Mfn2 in muscles of mice disrupts glucose homeostasis (Sebastián et al., 2012), and Drp1 ablation in the liver results in reduced adiposity and elevated whole-body energy expenditure, protecting mice from diet-induced obesity (Wang et al., 2015). A recent study has also suggested a link between Drp1-mediated fission and its effect on glycolysis during cell transformation (Serasinghe et al., 2015). The central question of whether fission/fusion and associated changes in cristae morphology actively control the adoption of distinct metabolic programs and therefore regulates T cell responses, however, remains unanswered.

## RESULTS

### Unlike T<sub>E</sub> Cells, T<sub>M</sub> Cells Maintain a Fused Mitochondrial Network

We reported that T<sub>M</sub> cells have more mitochondrial mass than T<sub>E</sub> or T<sub>N</sub> cells and suggested that mitochondria in these T cell subsets are morphologically distinct (van der Windt et al., 2012, 2013). These observations prompted us to assess mitochondrial structure in T cells. We infected mice with *Listeria monocytogenes* expressing ovalbumin (OVA) (LmOVA) and isolated T<sub>E</sub> and T<sub>M</sub> cells for ultrastructure analysis by electron microscopy (EM). We found that T<sub>E</sub> cells had small, distinct mitochondria dispersed in the cytoplasm, while T<sub>M</sub> cells had densely packed, somewhat tubular, mitochondria (Figure 1A). To investigate these morphological differences, we differentially cultured activated OVA-specific T cell receptor (TCR) transgenic OT-I cells in interleukin-2 (IL-2) and IL-15 to generate IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells (Figure S1) (Carrio et al., 2004). These culture conditions approximate T cell responses in vivo and allow us to generate large numbers of cells amenable to further experimentation in vitro (O'Sullivan et al., 2014). We found that IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells possessed similar mitochondrial ultrastructure as their ex vivo isolated counterparts (Figure 1B). Using confocal

microscopy, we observed that while a day after activation the mitochondria appeared fused, from days 2–6 the IL-2 T<sub>E</sub> cells exhibited predominantly punctate mitochondria (Figure 1C). In contrast, once cells were exposed to IL-15, a cytokine that supports T<sub>M</sub> cell formation (Schluns et al., 2002), the mitochondria formed elongated tubules. Magnified images from these experiments emphasized the marked differences in mitochondrial morphology between IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells (Figure 1D). We also examined the protein expression of several critical regulators of mitochondrial dynamics. We found that by day 6, fusion mediators Mfn2 and Opa1 were lower in T<sub>E</sub> cells compared to T<sub>M</sub> cells, while fission factor Drp1 was phosphorylated at its activating site Ser616 in T<sub>E</sub> cells (Figure 1E) (Marsboom et al., 2012). Together, these data suggest that mitochondria in T<sub>E</sub> cells actively undergo fission, while in T<sub>M</sub> cells, these organelles exist in a fused state.

### Mitochondrial Inner Membrane Fusion Protein Opa1 Is Necessary for T<sub>M</sub> Cell Generation

To test whether mitochondrial fusion was important for T<sub>M</sub> cell development, we crossed Mfn1, Mfn2, and Opa1 floxed mice to OT-I CD4 Cre transgenic mice to conditionally delete these proteins in T cells. Peripheral T cell frequencies in these mice were grossly normal (data not shown). We differentially cultured Mfn1<sup>-/-</sup>, Mfn2<sup>-/-</sup>, and Opa1<sup>-/-</sup> OT-I T cells in IL-2 and IL-15 and found that only Opa1<sup>-/-</sup> T cells had a survival defect when cultured in IL-15 (Figure 2A), but not in IL-2. We measured the efficiency of gene deletion by mRNA and/or protein analyses (Figures S2A–S2C). While Mfn1 and 2 were efficiently deleted, we found residual expression of Opa1 particularly in IL-15 T<sub>M</sub> cells, suggesting that cells that retained some expression of Opa1 in IL-15 cultures had a survival advantage (Figure 2A). Assessment of mitochondrial ultrastructure revealed that the cristae were altered and disorganized in the absence of Opa1 in agreement with published results for other cells (Figure 2B) (Zhang et al., 2011; Cogliati et al., 2013). Consistent with their survival defect, Opa1<sup>-/-</sup> IL-15 T<sub>M</sub> cells had decreased OXPHOS activity, as measured by O<sub>2</sub> consumption rate (OCR, an indicator of OXPHOS) to extracellular acidification rate (ECAR, an indicator of aerobic glycolysis) ratio, and spare respiratory capacity (SRC), compared to normal cells (Figure 2C). SRC is the extra mitochondrial capacity available in a cell to produce energy under conditions of increased work or stress and is thought to be important for long-term cell survival and function (measured as OXPHOS activity above basal after uncoupling with FCCP) (Nicholls, 2009; van der Windt et al., 2012). To determine whether Opa1 is required for T<sub>M</sub> cell development in vivo, we adoptively transferred Opa1<sup>-/-</sup> OT-I T<sub>N</sub> cells into congenic recipients, infected

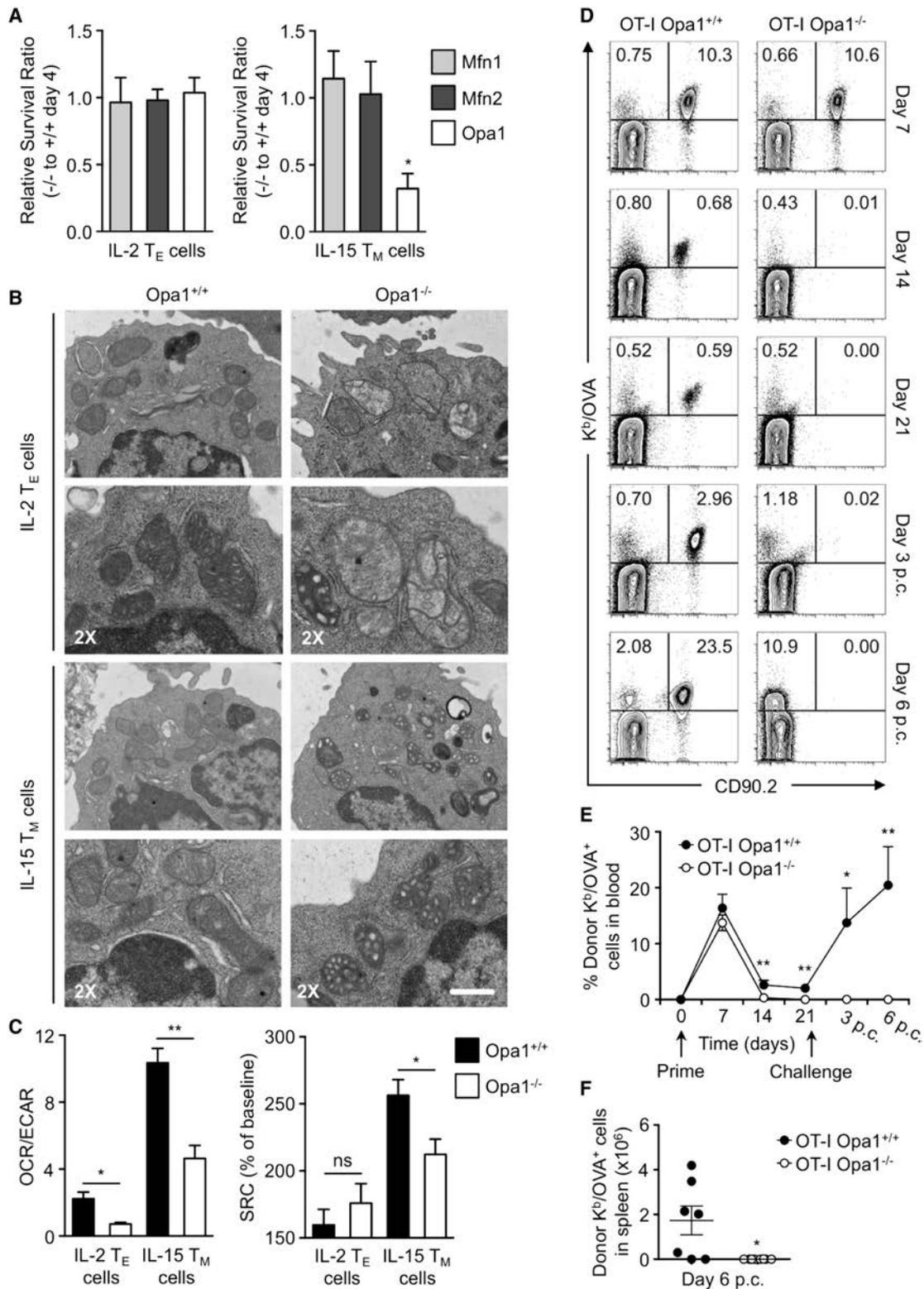
#### Figure 1. Effector and Memory T Cells Possess Distinct Mitochondrial Morphologies

(A and B) Effector (T<sub>E</sub>, CD44<sup>hi</sup> CD62L<sup>lo</sup>, 7 days post infection) and memory T (T<sub>M</sub>, CD44<sup>hi</sup> CD62L<sup>hi</sup>, 21 days post infection) cells (A) sorted from C57BL/6 mice infected i.p. with 10<sup>7</sup> CFU LmOVA and IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells (B) generated from differential culture of OT-I cells activated with OVA peptide and IL-2 using IL-2 or IL-15 analyzed by EM. Scale bar, 0.5 μm. Representative of two experiments.

(C and D) Mitochondrial morphology in live OT-I PhAM cells before and after αCD3/CD28 activation and differential cytokine culture by spinning disk confocal microscopy. Mitochondria are green (GFP) and nuclei are blue (Hoechst). Scale bar, 5 μm (C); 1 μm (D). Representative of two experiments.

(E) Immunoblot analysis of cell protein extracts from (C), probed for Mfn2, Opa1, Drp1, phosphorylated Drp1 at Ser616 (Drp1<sup>PS616</sup>), and β-actin. Representative of two experiments.

See also Figure S1.



**Figure 2. Memory T Cell Development and Survival, Unlike Effectors, Requires Mitochondrial Fusion**

(A) Relative in vitro survival ratios of Mfn1, Mfn2, or Opa1-deficient (CD4 Cre<sup>+</sup>, Cre<sup>-/-</sup>) to wild-type (CD4 Cre<sup>-</sup>, Cre<sup>+/+</sup>) OT-I IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells (\*p = 0.0465). Data normalized from two to three independent experiments shown as mean ± SEM.

(legend continued on next page)

these mice with LmOVA, and assessed  $T_M$  cell formation in the weeks after infection. Control and  $Opa1^{-/-}$  OT-I T cells mounted normal  $T_E$  cell responses (day 7) to infection, while  $Opa1^{-/-}$  OT-I  $T_M$  cell formation (days 14–21) was drastically impaired (Figure 2D). Consistent with diminished  $T_M$  cell development, a significantly higher proportion of short-lived effector to memory precursor effector cells were present within the  $Opa1^{-/-}$  OT-I donor cell population 7 days after infection (Figure S2D) (Kaeche et al., 2003). In addition, at day 10 post infection, a time point when  $T_E$  cells contract while  $T_M$  cells emerge,  $Opa1^{-/-}$  T cells isolated ex vivo had decreased SRC compared to control cells (Figure S2E), correlating with their decreased survival. To assess whether  $Opa1^{-/-}$   $T_M$  cells existed in too low an abundance to be discerned by flow cytometry, we challenged these mice with a second infection. We observed no recall response (day 3 and 6 p.c.) from  $Opa1^{-/-}$  T cells when assessing frequency (Figure 2E) or absolute numbers (Figure 2F), while there was considerable expansion of control donor cells. These data illustrate that  $Opa1$  is required for  $T_M$ , but not  $T_E$  cell generation.

### Mitochondrial Fusion Imposes a $T_M$ Cell Phenotype, Even in the Presence of Activating Signals

Genetic loss of function of  $Opa1$  revealed that this protein is critical for  $T_M$  cell formation. Given the fused phenotype of mitochondria in  $T_M$  cells, we hypothesized that  $Opa1$ -mediated mitochondrial fusion supports the metabolism needed for  $T_M$  cell development. We used a gain of function approach to enhance mitochondrial fusion. Culturing T cells with the “fusion promoter” M1 (Wang et al., 2012) and the “fission inhibitor” Mdivi-1 (Cassidy-Stone et al., 2008) (Figure 3A) induced mitochondrial fusion in IL-2  $T_E$  cells, rendering them morphologically similar to IL-15  $T_M$  cells (Figure 3B). Treatment with these drugs enhanced other  $T_M$  cell properties in activated IL-2  $T_E$  cells, including increased mitochondrial mass (Figure 3C), OXPHOS and SRC (Figure 3D), CD62L expression (Figure 3E), and robust metabolic activity, as indicated by bioenergetic profiling in response to secondary stimulation with PMA + ionomycin, followed by addition of oligomycin (ATP synthase inhibitor), FCCP, and rotenone with antimycin A (ETC complex I and III inhibitors), all drugs that stress the mitochondria (Figures 3F and S3A). However, we did not observe increased mtDNA in these cells (Figure S3B). We found that ECAR and the OCR/ECAR ratio increased after drug treatment (Figure S3C), indicating elevated metabolic activity overall, with a predominant increase in OXPHOS over glycolysis. While we observed these changes in mitochondrial activity, we did not measure any significant differences in mitochondrial membrane potential or ROS after drug treatment (Figure S3D). The expression of other activation markers were also not substantially

affected, although a small decrease in KLRG1 and increase in CD25 was measured (Figure S3E). Additionally, we performed a genetic gain of function experiment and transduced activated IL-2  $T_E$  cells with retrovirus expressing Mfn1, Mfn2, or  $Opa1$ . Similar to enforcement of fusion pharmacologically, we found that cells transduced with  $Opa1$  had more mitochondria (Figure 3G) and OXPHOS (Figure 3H), than empty vector control or Mfn-transduced T cells, as well as increased overall metabolic activity, with a predominant increase in OXPHOS over glycolysis (Figure S3F).  $T_M$  cell-associated markers such as CCR7 and CD127 were increased on transduced cells, as well as  $T_E$  cell proteins, such as PD-1 (Figure S3G). Overexpression of each target gene over the control was confirmed by mRNA expression (Figure S3H). Together, our results show that mitochondrial fusion confers  $T_M$  cell phenotypes on activated  $T_E$  cells even in culture conditions that program  $T_E$  cell differentiation.

### T Cell Mitochondrial Fusion Improves Adoptive Cellular Immunotherapy against Tumors

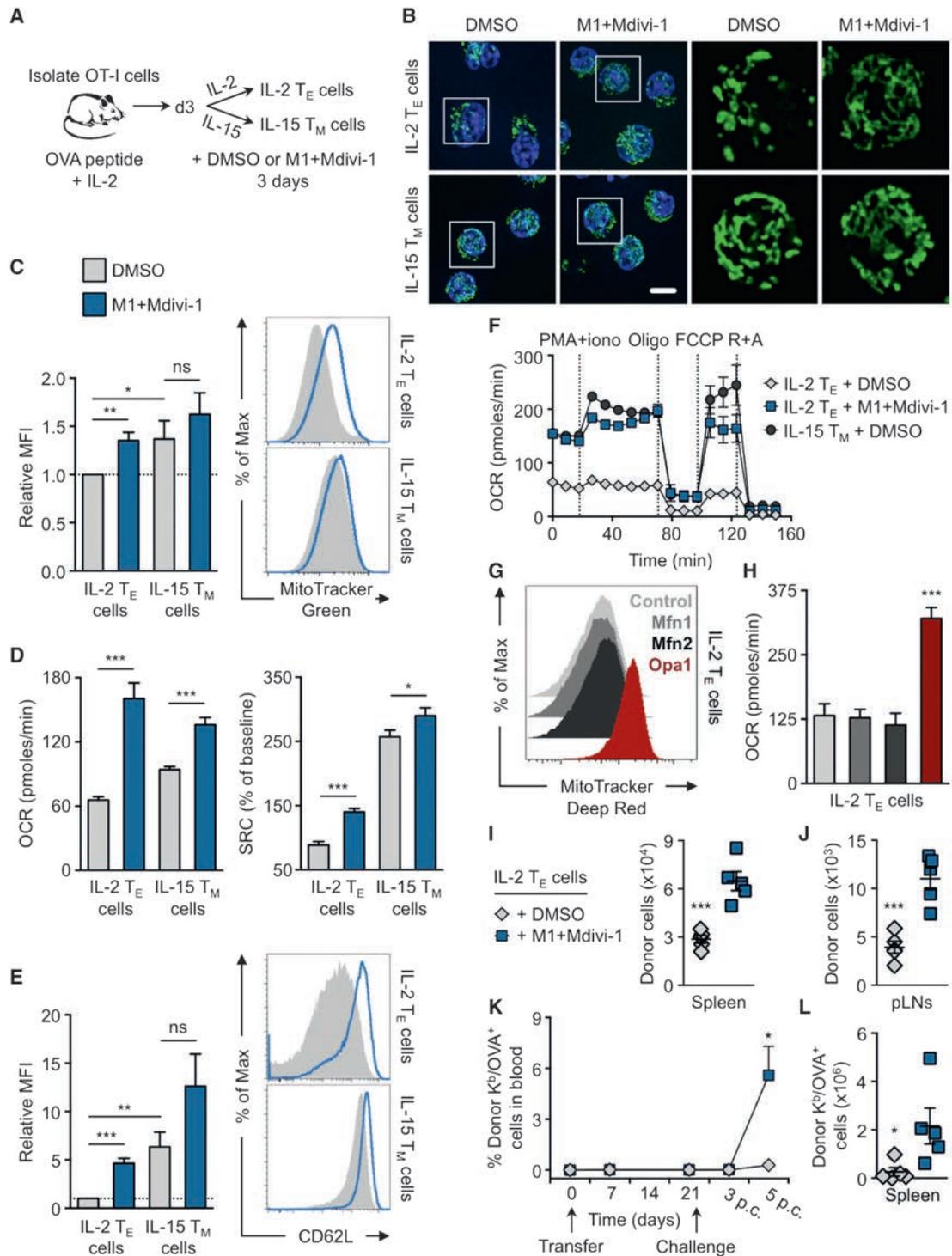
A consideration when designing adoptive cellular immunotherapy (ACI) is to improve T cell fitness during ex vivo culture so that when T cells are re-introduced into a patient they are able to function efficiently and persist over time (Restifo et al., 2012; Maus et al., 2014; O’Sullivan and Pearce, 2015). Our data showed that fusion-promoting drugs created metabolically fit T cells. We predicted that enforcing fusion would also enhance the longevity of IL-2  $T_E$  cells in vivo. To test this, we adoptively transferred control and M1 + Mdivi-1-treated OT-I T cells into congenic mice and tracked donor cell survival. We found significantly more drug-treated T cells in the spleen (Figure 3I) and lymph nodes (Figure 3J) 2 days after transfer. To determine if the persistence of these cells would be maintained better long term than controls, we infected mice with LmOVA >3 weeks later and measured T cell responses against the bacteria. We found that drug-treated cells selectively expanded in response to infection (Figure 3K) and could be recovered in significantly greater numbers in the spleen 6 days post challenge (Figure 3L).

Next, we assessed whether these drugs could be used to promote T cell function in an ACI model. We injected EL4-OVA tumor cells into mice. Then either 5 or 12 days later we adoptively transferred IL-2  $T_E$  cells that had been treated with DMSO or M1 + Mdivi-1. In both settings, mice that received “fusion-promoted” T cells were able to control tumor growth significantly better than mice that received control-treated cells (Figures 4A and 4B). The cytolytic ability (Figure S4A) and proliferation (Figure S4B) of the fusion enforced IL-2  $T_E$  cells were similar to controls, however, they expressed significantly higher levels of IFN- $\gamma$  and TNF- $\alpha$  when restimulated with PMA + ionomycin

(B and C) Mitochondrial morphology of OT-I  $Opa1$  wild-type and  $Opa1^{-/-}$  IL-2  $T_E$  and IL-15  $T_M$  cells analyzed by (B) EM (scale bar, 0.5  $\mu$ m, represents one experiment) and (C) Seahorse EFA. Bar graphs represent (left) ratios of  $O_2$  consumption rates (OCR, indicator of OXPHOS) to extracellular acidification rates (ECAR, indicator of aerobic glycolysis) at baseline and (right) spare respiratory capacity (SRC) (% max OCR after FCCP injection of baseline OCR) of indicated cells (\* $p$  < 0.03, \*\* $p$  = 0.0079). Data from three experiments are shown as mean  $\pm$  SEM.

(D–F)  $10^4$  OT-I  $Opa1^{+/+}$  or  $Opa1^{-/-}$  T cells were transferred i.v. into C57BL/6 CD90.1 mice infected i.v. with  $10^7$  CFU LmOVA. Blood analyzed by flow cytometry at indicated times post infection. After 21 days, mice were challenged i.v. with  $5 \times 10^7$  CFU LmOVA and blood analyzed post challenge (p.c.). (D) % Donor  $K^b/OVA^+$  CD90.2 $^+$  cells shown in representative flow plots and (E) line graph with mean  $\pm$  SEM (\* $p$  = 0.0238, \*\* $p$  < 0.005). (F) Number of donor  $K^b/OVA^+$  cells from spleens of infected mice shown with mean  $\pm$  SEM (\* $p$  = 0.0126). Representative of two experiments (n = 9–11/group).

See also Figure S2.



**Figure 3. Enhancing Mitochondrial Fusion Promotes the Generation of Memory-like T Cells**

(A–F and I–L) OVA peptide and IL-2-activated OT-I cells differentiated into IL-2 T<sub>E</sub> or IL-15 T<sub>M</sub> cells for 3 days in the presence of DMSO or 20 μM fusion promoter M1 and 10 μM fission inhibitor Mdivi-1 (M1 + Mdivi-1) as shown (A) pictorially. (B) Representative spinning disk confocal images from three experiments of live cells from OT-I PHAM mice. Mitochondria are green (GFP) and nuclei are blue (Hoechst). Scale bar, 5 μm. (C) Cells stained with MitoTracker Green and analyzed by flow cytometry. Relative MFI (left) from 6 experiments (\*p = 0.0394, \*\*p = 0.0019) with representative histograms (right). (D) Baseline OCR and SRC from three to four experiments (\*p = 0.0485, \*\*\*p < 0.0001), and (E) CD62L expression analyzed by flow cytometry of indicated cells. Relative MFI (left) from seven experiments (\*p = 0.0325, \*\*p = 0.0019, \*\*\*p < 0.0001) with representative histograms (right). (F) OCR of indicated cells at baseline and in response to PMA and ionomycin stimulation (PMA + iono), oligomycin (Oligo), FCCP, and rotenone plus antimycin A (R + A). Represents two experiments. (C–F) Shown as mean ± SEM. (I–L) A total

(legend continued on next page)

in vitro (Figure S4C). We also exposed activated human T cells to M1 + Mdivi-1 treatment in vitro and found that they had visibly more fused mitochondria (Figure 4C) and exhibited the bioenergetic profile (Figure 4D) and surface marker expression (Figure 4E) characteristic of  $T_M$  cells, compared to control-treated cells. Parameters such as mitochondrial mass (Figure 4E) and other surface markers (Figure S4D) were not significantly altered. These data suggest that promoting fusion in T cells may be a translatable treatment for enhancing human therapy.

### Mitochondrial Fusion Promotes $T_M$ Cell Metabolism, but Opa1 Is Not Required for FAO

Our data showed that Opa1 was necessary for  $T_M$  cell formation, but the question of how Opa1 acted to support  $T_M$  cells remained. We hypothesized that mitochondrial fusion, via Opa1 function, was needed for FAO, as the engagement of this pathway is required for  $T_M$  cell development and survival (Pearce et al., 2009; van der Windt et al., 2012, 2013). This hypothesis was based on our observations that these two processes seemed to be linked in  $T_M$  cells and also on a recent report that mitochondrial fusion is important for efficient FAO via lipid droplet trafficking under starvation conditions (Rambold et al., 2015). We treated IL-2  $T_E$  and IL-15  $T_M$  cells with M1 + Mdivi-1 or vehicle and then measured OCR in response to etomoxir, a specific inhibitor of mitochondrial long chain FAO (Deberardinis et al., 2006), and mitochondrial inhibitors. We found that the increased OCR and SRC evident in these cells after M1 + Mdivi-1 treatment was due to enhanced FAO (Figures 5A and S5A). IL-2  $T_E$  cells transduced with Opa1 also exhibited elevated OCR that decreased in the presence of etomoxir compared to controls (Figure 5B). Bone marrow-derived macrophages (BM-Macs) cultured with M1 + Mdivi-1 increased OCR and SRC to levels similar to M2 polarized macrophages, which engage FAO much like  $T_M$  cells do (Figure S5B) (Huang et al., 2014). Importantly, M1 + Mdivi-1 treatment did not increase OCR (Figure 5C) or affect ECAR (Figure S5C) in Opa1<sup>-/-</sup> IL-2  $T_E$  cells compared to controls, suggesting a requirement for Opa1 in augmenting OCR and FAO. However, in contrast to what we expected, when we assessed bioenergetics of Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> IL-2  $T_E$  (Figure 5D) and ex vivo isolated  $T_E$  cells (Figure 5E), we found that both are equally responsive to etomoxir. Our results show that while Opa1 can promote FAO in T cells, it is not compulsory for engagement of this metabolic pathway.

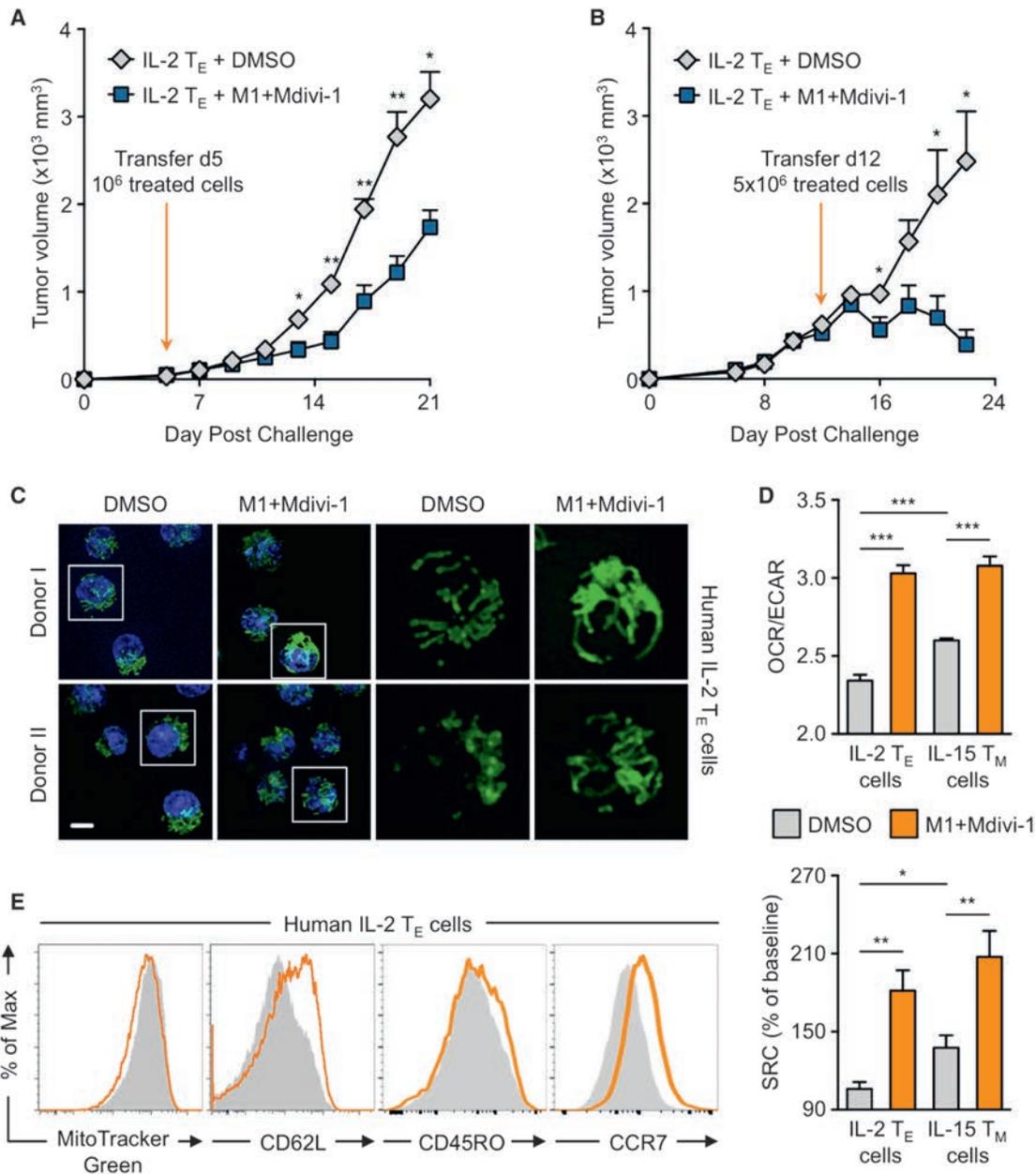
### Mitochondrial Cristae Remodeling Signals Metabolic Adaptations in $T_M$ and $T_E$ Cells

Although Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> IL-2  $T_E$  cells could equally engage FAO (Figure 5D), we observed that ECAR was significantly increased in Opa1<sup>-/-</sup> cells both in vitro and ex vivo (Fig-

ure 6A). Furthermore, unlike controls, we observed no additional drop of OCR in Opa1<sup>-/-</sup> IL-2  $T_E$  cells after the addition of oligomycin (Figure 6B), suggesting that in the absence of Opa1, only FAO supports OXPHOS, and that substrates, such as glucose-derived pyruvate, are not used for mitochondrial ATP production in this setting. We cultured Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> IL-2  $T_E$  cells with <sup>13</sup>C-labeled glucose and traced <sup>13</sup>C into TCA cycle metabolites. While the percent of <sup>13</sup>C-labeled pyruvate was higher in Opa1<sup>-/-</sup> T cells, the frequency of <sup>13</sup>C-labeled TCA cycle intermediates was significantly reduced in Opa1<sup>-/-</sup> T cells compared to controls (Figures 6C and S6A), a result that is supported by their higher ECAR (Figure 6A). These data suggested that without mitochondrial fusion, pyruvate is preferentially secreted as lactate, rather than oxidized in the mitochondria. Therefore, we questioned whether FAO is a “default” pathway for mitochondria in a resting or fused state (i.e., Opa1 sufficiency), and the induction of aerobic glycolysis is a major downstream effect of fission (i.e., Opa1 deficiency). If this were the case, then a balance between fission and fusion, modulated by proteins such as Opa1, could act as a primary signal to dictate the metabolic phenotype of T cells. In support of this idea, T cells from polyclonal T cell-conditional deleted Opa1 mice had higher ECAR and an increased proportion of CD8 T cells with an activated effector phenotype in the basal state based on surface marker expression (Figure S6B).

Opa1 is critical for inner mitochondrial membrane fusion, but also for other processes like cristae remodeling (Frezza et al., 2006; Cogliati et al., 2013). We observed major changes in cristae morphology in Opa1<sup>-/-</sup> T cells (Figure 2B). Given the importance of Opa1 in  $T_M$  cell development (Figure 2), we further assessed cristae morphology in  $T_E$  and  $T_M$  cells isolated ex vivo after LmOVA infection (Figure 6D), as well as IL-2  $T_E$  and IL-15  $T_M$  cells (Figure 6E), and found that  $T_E$  cells had many cristae with slightly wider and loosely organized intermembrane space than  $T_M$  cells. Opa1 overexpression induces cristae tightening and close association of ETC complexes in the inner mitochondrial membrane (Cogliati et al., 2013; Civiletto et al., 2015). Therefore, we surmised that in the absence of Opa1, cristae disorganization leads to dissociation of ETC complexes and subsequently less efficient ETC activity in T cells (Figure 2C). We assessed OCR after oligomycin in relation to OCR after rotenone and antimycin A treatment (i.e., proton leak), which indicates the coupling efficiency of OXPHOS to mitochondrial ATP production. Consistent with decreased OXPHOS efficiency, we observed elevated proton leak in Opa1<sup>-/-</sup> T cells compared to controls (Figure 6F). This was also true for ex vivo isolated  $T_E$  cells compared to  $T_M$  cells (Figure 6G), as well as IL-2  $T_E$  and IL-15  $T_M$  cells (Figure 6H). Together, these data suggest that there are cristae differences between  $T_E$  and  $T_M$  cells, which may contribute to their distinct metabolic phenotypes.

of 1–2 × 10<sup>6</sup> IL-2  $T_E$  cells cultured with DMSO (gray diamonds) or M1 + Mdivi-1 (blue squares) were transferred into congenic C57BL/6 recipient mice. Cell counts of donor cells recovered 2 days later from the (I) spleen (\*\*\*p = 0.005) and (J) peripheral lymph nodes (pLNs, \*\*\*p = 0.0006). Dots are individual mice. (K) Blood from recipient mice analyzed for % donor K<sup>b</sup>/OVA<sup>+</sup> cells post transfer and challenged with 10<sup>7</sup> CFU LmOVA by flow cytometry (\*p = 0.0150, n = 5/group). (L) Donor K<sup>b</sup>/OVA<sup>+</sup> cells recovered from recipient spleens 6 days p.c. (\*p = 0.0383). Dots are individual mice. (I–L) Represents two experiments shown with mean ± SEM. (G and H) OT-I cells were transduced with either empty (Control), Mfn1, Mfn2, or Opa1 expression vectors, sorted, and cultured to generate IL-2  $T_E$  cells. (G) Representative histograms of MitoTracker Deep Red staining from four experiments and (H) basal OCR from two experiments of transduced cells. See also Figure S3.



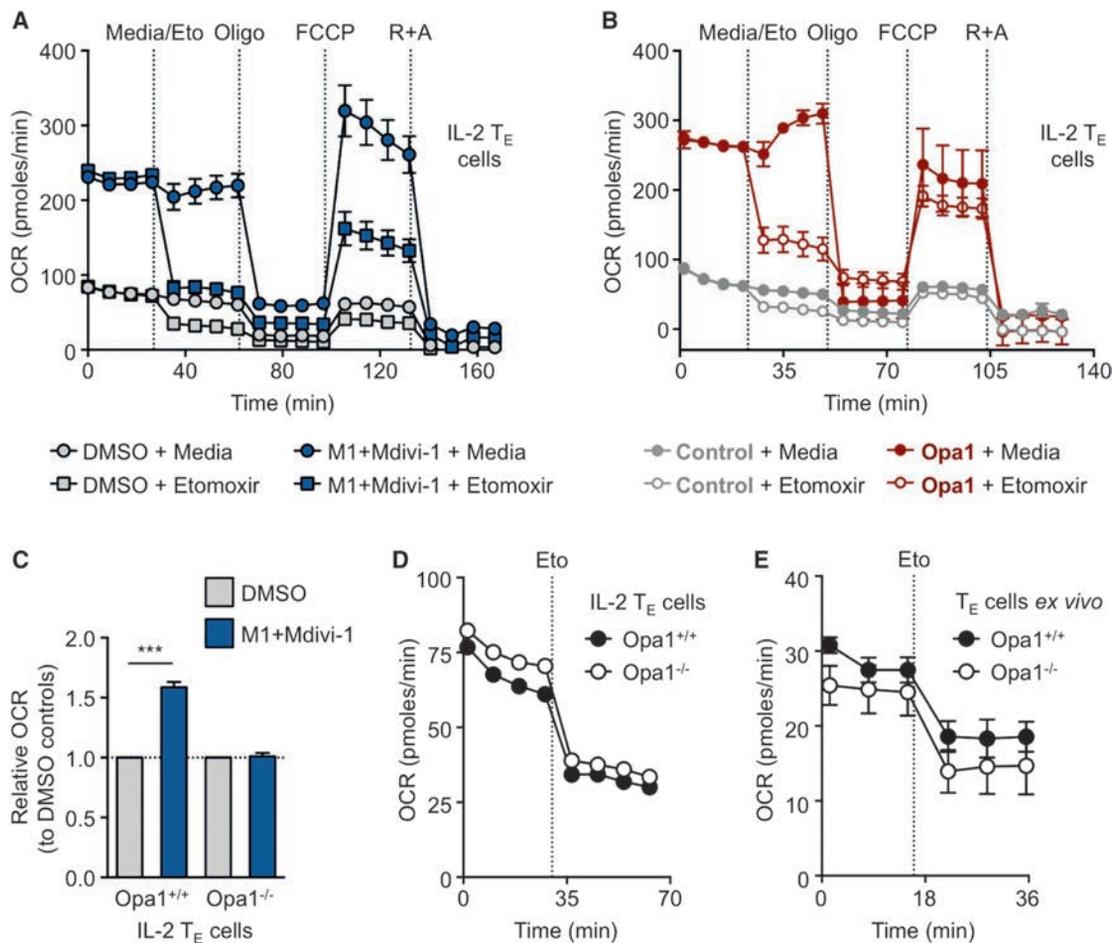
**Figure 4. Mitochondrial Fusion Improves Adoptive Cellular Immunotherapy against Tumors**

(A and B) C57BL/6 mice inoculated s.c. with 10<sup>6</sup> EL4-OVA cells. (A) After 5 or (B) 12 days, 10<sup>6</sup> or 5 × 10<sup>6</sup> OT-I IL-2 T<sub>E</sub> cells cultured with DMSO or M1 + Mdivi-1 were transferred i.v. into recipients and tumor growth assessed. Represents two experiments shown as mean ± SEM (n = 5/group, \*p < 0.05, \*\*p < 0.005).

(C–E) Human CD8<sup>+</sup> PBMCs activated with αCD3/CD28 + IL-2 to generate IL-2 T<sub>E</sub> cells. (C) Confocal images of indicated cells where mitochondria are green (MitoTracker) and nuclei are blue (Hoechst). Representative images from two of four biological donors. Scale bar, 5 μm. (D) OCR/ECAR ratios and SRC of indicated cells from 2 separate donors shown as mean ± SEM (\*\*p = 0.0303, \*\*p < 0.005, \*\*\*p < 0.0001). (E) MitoTracker Green staining and CD62L, CD45RO, and CCR7 expression analyzed by flow cytometry shown with representative histograms from four to six biological replicates. See also Figure S4.

We reasoned that fusion renders tightly configured cristae, which results in closely associated ETC complexes and efficient OXPHOS (Patten et al., 2014), producing conditions that favor the entrance of pyruvate into the TCA cycle. NADH generated from the TCA cycle is able to easily donate electrons to complex I, which are passed efficiently along the ETC. Our data suggest

that this predominantly occurs in T<sub>M</sub> cells. However, if electron transport across the ETC became less efficient, caused by physical separation of the individual complexes due to cristae remodeling via mitochondrial fission, then electrons could linger in the complexes and imbalance redox reactions. NADH levels would build, slowing forward momentum of the TCA cycle. To restore



**Figure 5. Fusion Promotes Memory T Cell Metabolism, but Opa1 Is Not Required for FAO**

OCR measured at baseline and in response to media, etomoxir (Eto) and other drugs as indicated of (A) IL-2 T<sub>E</sub> cells cultured in DMSO or M1 + Mdivi-1, (B) control or Opa1 transduced IL-2 T<sub>E</sub> cells, (C) Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> IL-2 T<sub>E</sub> cells cultured in DMSO or M1 + Mdivi-1, (D) or without drugs, and (E) ex vivo donor OT-I Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> day 7 T<sub>E</sub> cells derived from LmOVA infection. Representative of two experiments shown as mean ± SEM (\*\*p < 0.0001). See also Figure S5.

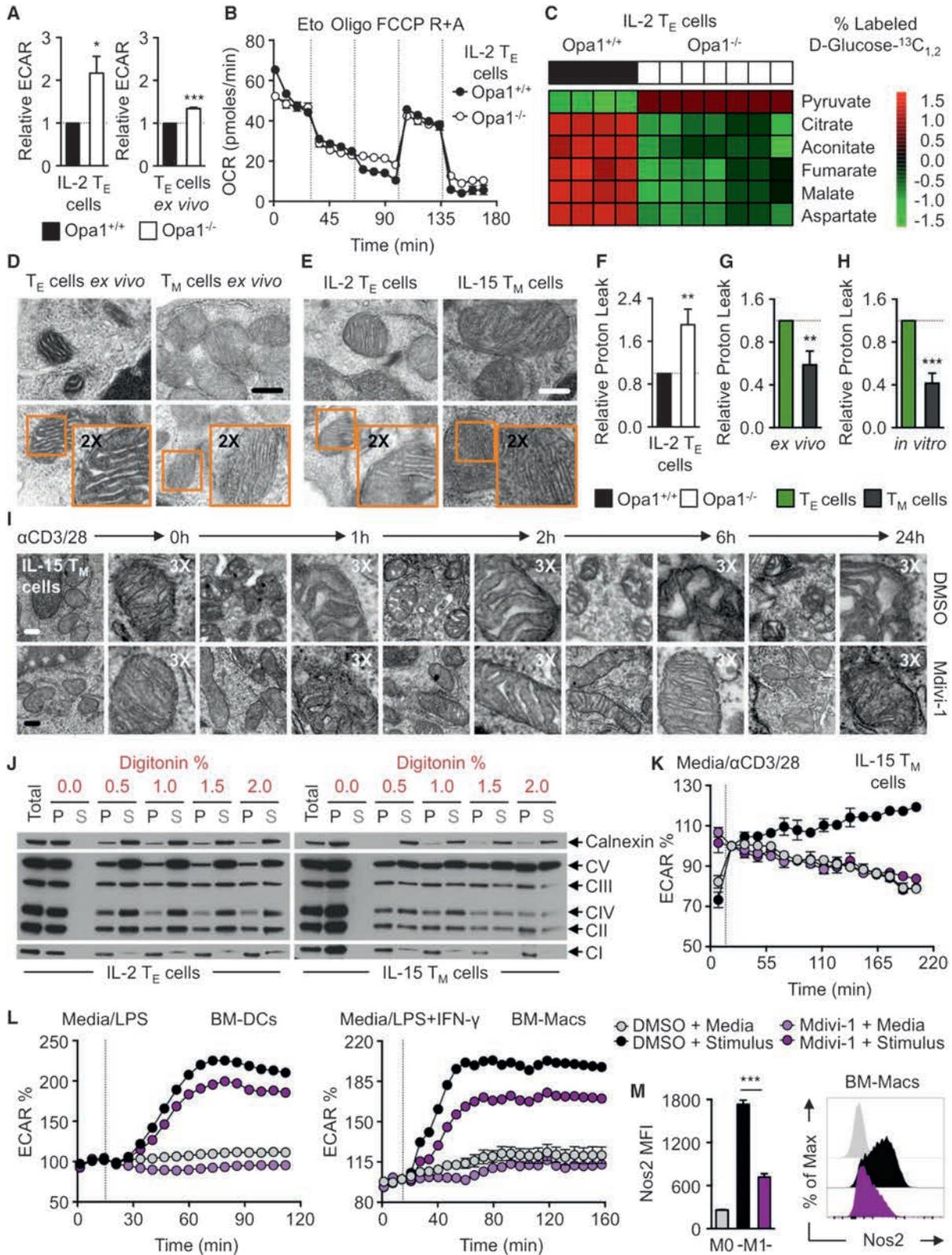
redox balance, cells could augment glycolysis and shunt pyruvate as excreted lactate (i.e., aerobic glycolysis), regenerating NAD<sup>+</sup> from cytosolic NADH. We speculate that this occurs in T<sub>E</sub> cells. Correlating with this idea, we previously reported that T<sub>E</sub> and T<sub>M</sub> cells have different ratios of NAD<sup>+</sup>/NADH (i.e., redox balance) with T<sub>M</sub> cells maintaining higher NAD<sup>+</sup>/NADH than T<sub>E</sub> cells. We also showed that NADH levels dramatically rise in T<sub>M</sub> cells compared to T<sub>E</sub> cells when exposed to rotenone and antimycin A, indicating that T<sub>M</sub> cells consume more NADH for the purpose of donating electrons to the ETC (van der Windt et al., 2012). Together, our data suggest that fission and fusion regulate cristae remodeling, which alter ETC efficiency and redox balance, ultimately controlling metabolic adaptations in T cells.

To examine this idea further, we assessed cristae morphology in T<sub>E</sub> and T<sub>M</sub> cells by EM after TCR stimulation. We hypothesized that if cristae remodeling induces aerobic glycolysis, changes in cristae structure could be visualized after T cell activation. T<sub>M</sub> cells rapidly augment aerobic glycolysis when restimulated (van der Windt et al., 2013). We activated IL-15 T<sub>M</sub> cells with

αCD3/CD28 beads (Figure 6I), or with PMA + ionomycin (Figure S6C), in the presence or absence of Mdivi-1 to modulate activity of mitochondrial fission protein Drp1 (Cassidy-Stone et al., 2008). We observed dramatic changes to cristae morphology by EM, with the intermembrane space widening over time in controls compared to drug-treated cells. These data are consistent with the hypothesis that fission-induced mitochondrial cristae remodeling supports metabolic reprogramming in T cells.

#### T<sub>M</sub> Cells Maintain Tight Cristae with Closely Associated ETC Complexes

Our data suggested that unlike T<sub>E</sub> cells, T<sub>M</sub> cells have tight cristae with closely associated ETC complexes. To investigate this biochemically, we treated native lysates of IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells with increasing concentrations of digitonin to disrupt cell membranes (including mitochondrial). The crude membrane-bound fraction was separated from solubilized proteins by centrifugation. Both pellet and soluble supernatants were loaded on a denaturing reducing gel and then probed for various



(legend on next page)

mitochondrial proteins by western blot. We found that mitochondria in IL-2 T<sub>E</sub> cells were susceptible to digitonin disruption, indicated by the fact that ETC complex proteins became less detectable in the pellet and enriched in the soluble fraction in 0.5% detergent (Figures 6J and S6D). This was in contrast to IL-15 T<sub>M</sub> cells, where ETC proteins did not solubilize to the same extent as those in IL-2 T<sub>E</sub> cells into the supernatant, even when 2% digitonin was used. To investigate whether this phenomenon was unique to the mitochondrial compartment, we also probed for the ER integral protein calnexin and found that it solubilized similarly in 0.5% digitonin in both cell types. Overall, these data suggest that there is more exposed mitochondrial membrane between proteins in IL-2 T<sub>E</sub> than IL-15 T<sub>M</sub> cells and is consistent with the idea that T<sub>M</sub> cells have tight cristae that yield efficient ETC activity, while T<sub>E</sub> cells have looser cristae with less efficient ETC activity, ultimately supporting their distinct metabolic phenotypes.

### Mitochondrial Fission in Activated Immune Cells Facilitates Aerobic Glycolysis

Our data support a model wherein cristae remodeling, through fission and fusion events, is a mechanism to regulate efficient OXPHOS and FAO in T<sub>M</sub> cells, as well as the induction of aerobic glycolysis in T<sub>E</sub> cells. To more directly test this idea, we assessed ECAR of IL-15 T<sub>M</sub> cells stimulated with  $\alpha$ CD3/28 beads in the presence or absence of Mdivi-1. We found that when mitochondrial fission protein Drp1 was inhibited with Mdivi-1, T cell activation did not robustly increase aerobic glycolysis when compared to controls (Figure 6K), which correlated with our EM data (Figure 6I). Because fission is associated with cell division (Taguchi et al., 2007), we tested our idea in a non-proliferating cell that substantially augments aerobic glycolysis upon stimulation (Krawczyk et al., 2010). We stimulated bone marrow-derived dendritic cells (BM-DCs) and macrophages (BM-Macs) with lipopolysaccharide (LPS) with or without interferon (IFN)- $\gamma$  in the presence or absence of Mdivi-1 and measured ECAR. Aerobic glycolysis was curtailed in BM-DCs and BM-Macs after stimulation when Drp1 was inhibited (Figure 6L). The blunted ECAR in Mdivi-1-treated cells correlated with decreased nitric oxide synthase 2 (Nos2) protein expression in the BM-Macs (Figure 6M), indicating that their activation was also repressed. These data

show that cristae remodeling and/or fission acts as a signal to drive the induction of aerobic glycolysis and subsequent cell activation via Drp1.

### DISCUSSION

Although T<sub>M</sub> cells rely on FAO for development and survival, precisely why T<sub>M</sub> cells utilize FAO and the signals that drive the induction of aerobic glycolysis in T<sub>E</sub> cells remain unclear. Our data suggest that manipulating the structure of a single organelle can have profound consequences that impact metabolism and ultimately cell differentiation. We found that Opa1 regulated tight cristae organization in T<sub>M</sub> cells, which facilitated efficient ETC activity and favorable redox balance, allowing continued entrance of pyruvate into mitochondria. We originally hypothesized that Opa1 would be required for FAO. However, we found that Opa1<sup>-/-</sup> IL-2 T<sub>E</sub> and ex vivo T<sub>E</sub> cells generated during infection utilized FAO to the same level as controls. While this was true for T<sub>E</sub> cells, this may not be the case for T<sub>M</sub> cells, whose survival is severely impaired in vitro and in vivo when Opa1-deficient. It is possible that Opa1<sup>-/-</sup> T cells are unable to form T<sub>M</sub> cells because they cannot efficiently engage FAO under the metabolic constraints imposed during T<sub>M</sub> cell development. Previous studies point to the existence of a “futile” cycle of fatty acid synthesis (FAS) and FAO in T<sub>M</sub> cells (O’Sullivan et al., 2014; Cui et al., 2015) whereby carbon derived from glucose is used to build fat that is subsequently burned by mitochondria. T<sub>M</sub> cells have a lower overall metabolic rate than T<sub>E</sub> cells, and tightly configured cristae may be important to ensure that any pyruvate generated will efficiently feed into the TCA cycle not only for reducing equivalents, but also for deriving citrate for FAS. Without tight cristae and efficient ETC activity, electrons may loiter in the complexes causing more ROS that could be damaging, but also provide signals that drive cell activation (Sena et al., 2013).

We did not observe a defect in T<sub>M</sub> cell survival in Mfn1<sup>-/-</sup> or Mfn2<sup>-/-</sup> T cells, but this does not exclude the possibility that OMM fusion or additional activities ascribed to each protein are not important. Mfn1 and Mfn2 form homotypic and heterotypic interactions, suggesting that in the absence of one, the other can compensate (Chen et al., 2003). Our results show that unlike Opa1<sup>-/-</sup> T cells, in vitro cultured Mfn1<sup>-/-</sup> or Mfn2<sup>-/-</sup>

### Figure 6. Mitochondrial Cristae Remodeling Signals Metabolic Pathway Engagement

- (A) Basal ECAR of OT-I Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> IL-2 T<sub>E</sub> cells (left) and day 7 T<sub>E</sub> cells isolated ex vivo after adoptive transfer from LmOVA infection (right). Data combined from two to three experiments (\*p = 0.0412, \*\*\*p < 0.0001).
- (B) OCR at baseline and with indicated drugs, representative of two experiments shown as mean  $\pm$  SEM.
- (C) D-Glucose-<sup>13</sup>C<sub>1,2</sub> trace analysis of OT-I Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> IL-2 T<sub>E</sub> cells. Each lane represents separate mice with a technical replicate.
- (D and E) EM analysis of mitochondrial cristae from (D) T<sub>E</sub> and T<sub>M</sub> cells isolated after LmOVA infection and (E) in vitro cultured IL-2 T<sub>E</sub> and T<sub>M</sub> cells. Representative of two experiments. Scale bar, 0.25  $\mu$ m.
- (F–H) Relative proton leak ( $\Delta$ OCR after oligomycin and subsequent injection of rotenone plus antimycin A) of (F) Opa1<sup>+/+</sup> and Opa1<sup>-/-</sup> IL-2 T<sub>E</sub>, (G) infection-elicited T<sub>E</sub> and T<sub>M</sub>, and (H) IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells. Combined from two to four experiments shown as mean  $\pm$  SEM (p\*\* < 0.005, \*\*\*p < 0.0001).
- (I) EM analysis of IL-15 T<sub>M</sub> cell-mitochondrial cristae before and after  $\alpha$ CD3/CD28 bead stimulation over hours. Scale bar, 0.2  $\mu$ m. Represents one experiment.
- (J) Immunoblot analysis of calnexin and ETC complexes (CI-NDUFB8, CII-SDHB, CIII-UQCRC2, CIV-MTCO1, CV-ATP5A). Equivalent numbers of IL-2 T<sub>E</sub> and IL-15 T<sub>M</sub> cells lysed in native lysis buffer followed by digitonin solubilization of intracellular membranes with pellet (P) and solubilized supernatant (S) fractions resolved on a denaturing gel, representative of two experiments.
- (K–M) IL-15 T<sub>M</sub> cell (K), BM-DCs (L), and BM-Macs (M) % ECAR measured at baseline and after media,  $\alpha$ CD3/CD28 bead, LPS, or LPS+IFN- $\gamma$  injection as indicated. Data baselined prior to or right after injection with stimuli. (M) BM-Macs stained for intracellular Nos2 protein by flow cytometry with MFI values (left) and representative histogram (right). Shown as mean  $\pm$  SEM and represents two to three experiments (\*\*\*p < 0.0001).
- See also Figure S6.

T cells do not have a survival defect when differentiated in IL-15, even though, like  $Opa1^{-/-}$  T cells, they are more glycolytic and OXPHOS-impaired compared to controls (data not shown). Our imaging data showed that  $T_M$  cells maintained fused mitochondrial networks, suggesting that OMM fusion also has a compulsory role in  $T_M$  cell development. However, unlike  $Opa1$ , retroviral expression of  $Mfn1$  and  $Mfn2$  did not confer a  $T_M$  cell phenotype in  $T_E$  cells. In this setting, an increase in OMM fusion without a concomitant increase in inner membrane fusion may still yield loose cristae morphology and a redox state that by default results in sustained excretion of lactate.

The question of what initial signals drive T cell mitochondrial remodeling still remains. In the case of  $T_M$  cell development, withdrawal of activating signals and growth factors may induce fusion, consistent with previous reports that starvation induces mitochondrial hyperfusion (Rambold et al., 2015), an effect we also observe in  $T_E$  cells after IL-2 withdrawal (data not shown). However, pro-survival signals from cytokines like IL-15 or IL-7 are needed to sustain  $T_M$  cell viability and metabolically remodel these cells for FAS and FAO via increased CPT1a (van der Windt et al., 2012) and aquaporin 9 expression (Cui et al., 2015). Factors such as these may enforce fusion and would be consistent with our observations that mitochondria in activated T cells subsequently cultured in IL-15 fuse over time. Another possibility is that during metabolic stress,  $Opa1$  is activated via sirtuin 3 (SIRT3) (Samant et al., 2014). Sirtuins are post-translational modifiers that are activated by  $NAD^+$ , directly tying their activity to cell metabolism (Houtkooper et al., 2012; Wang and Green, 2012). We previously published that there is more available  $NAD^+$  in  $T_M$  cells (van der Windt et al., 2012), which may correlate with this scenario.

In  $T_E$  cells a day after TCR stimulation, we saw activation of Drp1 via its phosphorylation site Ser616 prior to seeing a fission phenotype. TCR signals instigate  $Ca^{2+}$  flux that promotes the phosphatase activity of calcineurin (Smith-Garvin et al., 2009), which in turn dephosphorylates Drp1 at Ser637, leading to its activation (Cereghetti et al., 2008). Initial Drp1 activation could facilitate some level of fission and cristae remodeling, tipping off aerobic glycolysis via the initial shunting of pyruvate to lactate. Although inhibition of Drp1 blocks activation-induced ECAR, our preliminary data do not show overt mitochondrial fragmentation in the early hours after TLR stimulation of DC or macrophages (data not shown). This however, does not exclude the possibility that Drp1 is mediating subtle changes to mitochondrial structure that are not discernable by confocal microscopy. Drp1 also affects cristae structure by altering the fluidity of the mitochondrial membrane (Benard et al., 2007). While Drp1 has been implicated in mitochondrial positioning at the immune synapse (Baixauli et al., 2011), lymphocyte chemotaxis (Campello et al., 2006), and ROS production during T cell activation (Röth et al., 2014), our data suggest that in addition to these processes, cristae remodeling via fission underlies the programming of cells to aerobic glycolysis.

We show that mitochondria in IL-2  $T_E$  cells are more susceptible to digitonin disruption than IL-15  $T_M$  cells, suggesting a more exposed membrane with less densely packed protein complexes. This relatively enhanced permeability, however, does not mean that their mitochondria are damaged or unable

to function. In fact, although  $T_E$  cells have less efficient OXPHOS in terms of coupling to ATP synthesis,  $T_E$  cells are very metabolically active with high OCR and ECAR (Chang et al., 2013; Sena et al., 2013). Our experiments involving pharmacological enforcement of mitochondrial fusion promoted OCR and SRC (and ECAR, albeit to a lesser extent) in IL-2  $T_E$  cells. The drug-modified cells maintained full  $T_E$  cell function with no effect on cytolytic ability or proliferation, but possessed enhanced cytokine expression. Fusion and/or cristae tightening boosted oxidative capacity in  $T_E$  cells, endowing them with longevity and persistence, while their higher aerobic glycolysis supported increased cytokine production, which may explain their superior antitumor function.

Our data suggest a model where morphological changes in mitochondria are a primary signal that shapes metabolic reprogramming during cellular quiescence or activation. We speculate that fission-associated expansion of cristae as a result of TCR stimulation physically separates ETC complexes, decreasing ETC efficiency. With delayed movement of electrons down the ETC, NADH levels rise, slowing forward momentum of the TCA cycle and cause an initial drop in ATP. To correct redox balance, cells will export pyruvate to lactate to regenerate  $NAD^+$  in the cytosol, which can enter the mitochondria through various shuttles to restore redox balance (Dawson, 1979) and increase flux through glycolysis to restore ATP levels, all contributing to the Warburg effect in activated T cells. When cristae are tightly configured, the ETC works efficiently and maintains entrance of pyruvate into the mitochondria with a favorable redox balance. In this case, cristae morphology as a result of fusion directs  $T_M$  cell formation and retains these cells in a quiescent state. Thus, mitochondrial dynamics controls the balance between metabolic pathway engagement and T cell fate.

## EXPERIMENTAL PROCEDURES

See the Supplemental Information for details.

### Mice and Immunizations

Mice were purchased from The Jackson Laboratory. LmOVA deleted for *actA* was used for primary and secondary immunizations. For tumor studies, EL4 lymphoma cells expressing OVA (EL4-OVA) were injected subcutaneously (s.c.) into the flank of mice.

### Flow Cytometry and Imaging

Antibody and H2-K<sup>b</sup>OVA<sub>257-264</sub> (K<sup>b</sup>OVA) MHC-peptide tetramer staining were performed as described (Chang et al., 2015). For electron microscopy (EM), cells were fixed in 2% paraformaldehyde, 2.5% glutaraldehyde in 100 mM sodium cacodylate containing 0.05% malachite green.

### Metabolism Assays

OCR and ECAR were measured using a 96-well XF or XFe extracellular flux analyzer (EFA) (Seahorse Bioscience). For fission inhibition studies, cells were plated in assay media with 10  $\mu$ M Mdivi-1 or vehicle control (DMSO) followed by assay media or  $\alpha$ CD3/CD28 bead or 20 ng/ml LPS  $\pm$  50 ng/ml IFN- $\gamma$  injection.

### Glucose Tracing

Cells were activated in glucose-free media (prepped with dialyzed FBS) supplemented with 11 mM glucose. After 3 days, cells were washed and cultured overnight in media replaced with 11 mM D-[1,2-<sup>13</sup>C] glucose. Metabolites were extracted with MeOH and analyzed by mass spectrometry (MS).

### Adoptive Transfers

For *in vivo* T<sub>M</sub> cell studies,  $\leq 10^4$  OT-I<sup>+</sup> CD8<sup>+</sup> donor cells/mouse were transferred intravenously (i.v.) into congenic C57BL/6 mice. For *in vivo* survival experiments,  $1\text{--}2 \times 10^6$  day 6 IL-2 T<sub>E</sub>-treated cells/mouse were injected i.v. into congenic C57BL/6 mice. For tumor studies,  $1\text{--}5 \times 10^6$  day 6 IL-2 T<sub>E</sub>-treated cells/mouse were injected i.v. into previously EL4-OVA-inoculated mice.

### Retroviral Transduction

Activated OT-I splenocytes were transduced with control (empty vector) or Mfn1, Mfn2, Opa1 expressing retrovirus by centrifugation. GFP or human CD8 marked retroviral expression.

### Statistical Analysis

Comparisons for two groups were calculated using unpaired two-tailed Student's *t* tests and one-way ANOVA followed by Bonferroni's multiple comparison tests for more than two groups. Comparisons over time were calculated using two-way ANOVA followed by Bonferroni's multiple comparison tests.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.05.035>.

### AUTHOR CONTRIBUTIONS

M.D.B., D.O., R.I.K.G., D.E.S., D.B., B.T.E., E.J.P., H.S., T.B.H., A.S.R., and E.L.P. designed the research and analyzed data. M.D.B., D.O., R.I.K.G., J.D.C., C.-H.C., D.E.S., J.Q., O.K., D.B., G.v.d.W., S.C.-C.H., and C.M.O. performed experiments. M.D.B. and E.L.P. wrote the manuscript.

### ACKNOWLEDGMENTS

We thank David Chan (Caltech) for Mfn1/2 floxed mice, Tom Graeber (DMMP, David Geffen School of Medicine, UCLA Metabolomics Center), Wandy Beatty (Molecular Microbiology Imaging Facility, Washington University School of Medicine), Erica Lanteme, Dorjan Brinja, Barbara Joch, Hani Suleiman, Elena Tonc, Tara Bradstreet, and Elizabeth Schwarzkopf for technical assistance. This work was funded by the NIH (CA181125 and AI091965 to E.L.P. and AI110481 to E.J.P.), the Burroughs Wellcome Fund (Investigator in the Pathogenesis of Infectious Disease Award to E.L.P. and Career Award for Medical Scientists to B.T.E.), the BMBF (to A.S.R.), the Max Planck Society, and the NSF Graduate Research Fellowship DGE-1143954 (to M.D.B.).

Received: December 24, 2015

Revised: April 3, 2016

Accepted: May 6, 2016

Published: June 9, 2016

### REFERENCES

- Archer, S.L. (2014). Mitochondrial fission and fusion in human diseases. *N. Engl. J. Med.* *370*, 1074.
- Baixauli, F., Martín-Córfecres, N.B., Morlino, G., Carrasco, Y.R., Calabia-Linares, C., Veiga, E., Serrador, J.M., and Sánchez-Madrid, F. (2011). The mitochondrial fission factor dynamin-related protein 1 modulates T-cell receptor signalling at the immune synapse. *EMBO J.* *30*, 1238–1250.
- Benard, G., Bellance, N., James, D., Parrone, P., Fernandez, H., Letellier, T., and Rossignol, R. (2007). Mitochondrial bioenergetics and structural network organization. *J. Cell Sci.* *120*, 838–848.
- Buck, M.D., O'Sullivan, D., and Pearce, E.L. (2015). T cell metabolism drives immunity. *J. Exp. Med.* *212*, 1345–1360.
- Campello, S., Lacalle, R.A., Bettella, M., Mañes, S., Scorrano, L., and Viola, A. (2006). Orchestration of lymphocyte chemotaxis by mitochondrial dynamics. *J. Exp. Med.* *203*, 2879–2886.
- Carrio, R., Bathe, O.F., and Malek, T.R. (2004). Initial antigen encounter programs CD8<sup>+</sup> T cells competent to develop into memory cells that are activated in an antigen-free, IL-7- and IL-15-rich environment. *J. Immunol.* *172*, 7315–7323.
- Cassidy-Stone, A., Chipuk, J.E., Ingeman, E., Song, C., Yoo, C., Kuwana, T., Kurth, M.J., Shaw, J.T., Hinshaw, J.E., Green, D.R., and Nunnari, J. (2008). Chemical inhibition of the mitochondrial division dynamin reveals its role in Bax/Bak-dependent mitochondrial outer membrane permeabilization. *Dev. Cell* *14*, 193–204.
- Cereghetti, G.M., Stangherlin, A., Martins de Brito, O., Chang, C.R., Blackstone, C., Bernardi, P., and Scorrano, L. (2008). Dephosphorylation by calcineurin regulates translocation of Drp1 to mitochondria. *Proc. Natl. Acad. Sci. USA* *105*, 15803–15808.
- Chan, D.C. (2012). Fusion and fission: interlinked processes critical for mitochondrial health. *Annu. Rev. Genet.* *46*, 265–287.
- Chang, C.H., Curtis, J.D., Maggi, L.B., Jr., Faubert, B., Villarino, A.V., O'Sullivan, D., Huang, S.C., van der Windt, G.J., Blagih, J., Qiu, J., et al. (2013). Post-transcriptional control of T cell effector function by aerobic glycolysis. *Cell* *153*, 1239–1251.
- Chang, C.H., Qiu, J., O'Sullivan, D., Buck, M.D., Noguchi, T., Curtis, J.D., Chen, Q., Gindin, M., Gubin, M.M., van der Windt, G.J., et al. (2015). Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell* *162*, 1229–1241.
- Chen, H., Detmer, S.A., Ewald, A.J., Griffin, E.E., Fraser, S.E., and Chan, D.C. (2003). Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. *J. Cell Biol.* *160*, 189–200.
- Civiletto, G., Varanita, T., Cerutti, R., Gorletta, T., Barbaro, S., Marchet, S., Lamperti, C., Viscomi, C., Scorrano, L., and Zeviani, M. (2015). Opa1 overexpression ameliorates the phenotype of two mitochondrial disease mouse models. *Cell Metab.* *21*, 845–854.
- Cogliati, S., Frezza, C., Soriano, M.E., Varanita, T., Quintana-Cabrera, R., Corrado, M., Cipolat, S., Costa, V., Casarin, A., Gomes, L.C., et al. (2013). Mitochondrial cristae shape determines respiratory chain supercomplexes assembly and respiratory efficiency. *Cell* *155*, 160–171.
- Cui, G., Staron, M.M., Gray, S.M., Ho, P.C., Amezcua, R.A., Wu, J., and Kaech, S.M. (2015). IL-7-induced glycerol transport and TAG synthesis promotes memory CD8<sup>+</sup> T cell longevity. *Cell* *161*, 750–761.
- Dawson, A.G. (1979). Oxidation of cytosolic NADH formed during aerobic metabolism in mammalian cells. *Trends Biochem. Sci.* *4*, 171–176.
- de Brito, O.M., and Scorrano, L. (2008). Mitofusin 2 tethers endoplasmic reticulum to mitochondria. *Nature* *456*, 605–610.
- Deberardinis, R.J., Lum, J.J., and Thompson, C.B. (2006). Phosphatidylinositol 3-kinase-dependent modulation of carnitine palmitoyltransferase 1A expression regulates lipid metabolism during hematopoietic cell growth. *J. Biol. Chem.* *281*, 37372–37380.
- Frank, M., Duvezin-Caubet, S., Koob, S., Occhipinti, A., Jagasia, R., Petcherski, A., Ruonala, M.O., Priault, M., Salin, B., and Reichert, A.S. (2012). Mitophagy is triggered by mild oxidative stress in a mitochondrial fission dependent manner. *Biochim. Biophys. Acta* *1823*, 2297–2310.
- Frezza, C., Cipolat, S., Martins de Brito, O., Micaroni, M., Beznoussenko, G.V., Rudka, T., Bartoli, D., Polishuck, R.S., Danial, N.N., De Strooper, B., and Scorrano, L. (2006). OPA1 controls apoptotic cristae remodeling independently from mitochondrial fusion. *Cell* *126*, 177–189.
- Friedman, J.R., and Nunnari, J. (2014). Mitochondrial form and function. *Nature* *505*, 335–343.
- Gomes, L.C., Di Benedetto, G., and Scorrano, L. (2011). During autophagy mitochondria elongate, are spared from degradation and sustain cell viability. *Nat. Cell Biol.* *13*, 589–598.
- Houtkooper, R.H., Pirinen, E., and Auwerx, J. (2012). Sirtuins as regulators of metabolism and healthspan. *Nat. Rev. Mol. Cell Biol.* *13*, 225–238.
- Huang, S.C., Everts, B., Ivanova, Y., O'Sullivan, D., Nascimento, M., Smith, A.M., Beatty, W., Love-Gregory, L., Lam, W.Y., O'Neill, C.M., et al. (2014).

- Cell-intrinsic lysosomal lipolysis is essential for alternative activation of macrophages. *Nat. Immunol.* **15**, 846–855.
- Kaech, S.M., Tan, J.T., Wherry, E.J., Konieczny, B.T., Surh, C.D., and Ahmed, R. (2003). Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nat. Immunol.* **4**, 1191–1198.
- Krawczyk, C.M., Holowka, T., Sun, J., Blagih, J., Amiel, E., DeBerardinis, R.J., Cross, J.R., Jung, E., Thompson, C.B., Jones, R.G., and Pearce, E.J. (2010). Toll-like receptor-induced changes in glycolytic metabolism regulate dendritic cell activation. *Blood* **115**, 4742–4749.
- Liesa, M., and Shirihai, O.S. (2013). Mitochondrial dynamics in the regulation of nutrient utilization and energy expenditure. *Cell Metab.* **17**, 491–506.
- MacIver, N.J., Michalek, R.D., and Rathmell, J.C. (2013). Metabolic regulation of T lymphocytes. *Annu. Rev. Immunol.* **31**, 259–283.
- Marsboom, G., Toth, P.T., Ryan, J.J., Hong, Z., Wu, X., Fang, Y.H., Thenappan, T., Piao, L., Zhang, H.J., Pogoriler, J., et al. (2012). Dynamin-related protein 1-mediated mitochondrial mitotic fission permits hyperproliferation of vascular smooth muscle cells and offers a novel therapeutic target in pulmonary hypertension. *Circ. Res.* **110**, 1484–1497.
- Maus, M.V., Fraietta, J.A., Levine, B.L., Kalos, M., Zhao, Y., and June, C.H. (2014). Adoptive immunotherapy for cancer or viruses. *Annu. Rev. Immunol.* **32**, 189–225.
- Mishra, P., and Chan, D.C. (2016). Metabolic regulation of mitochondrial dynamics. *J. Cell Biol.* **212**, 379–387.
- Mishra, P., Carelli, V., Manfredi, G., and Chan, D.C. (2014). Proteolytic cleavage of Opa1 stimulates mitochondrial inner membrane fusion and couples fusion to oxidative phosphorylation. *Cell Metab.* **19**, 630–641.
- Nicholls, D.G. (2009). Spare respiratory capacity, oxidative stress and excitotoxicity. *Biochem. Soc. Trans.* **37**, 1385–1388.
- Nunnari, J., and Suomalainen, A. (2012). Mitochondria: in sickness and in health. *Cell* **148**, 1145–1159.
- O'Sullivan, D., and Pearce, E.L. (2015). Targeting T cell metabolism for therapy. *Trends Immunol.* **36**, 71–80.
- O'Sullivan, D., van der Windt, G.J., Huang, S.C., Curtis, J.D., Chang, C.H., Buck, M.D., Qiu, J., Smith, A.M., Lam, W.Y., DiPlato, L.M., et al. (2014). Memory CD8(+) T cells use cell-intrinsic lipolysis to support the metabolic programming necessary for development. *Immunity* **41**, 75–88.
- Patten, D.A., Wong, J., Khacho, M., Soubannier, V., Mailloux, R.J., Pilon-Larose, K., MacLaurin, J.G., Park, D.S., McBride, H.M., Trinkle-Mulcahy, L., et al. (2014). OPA1-dependent cristae modulation is essential for cellular adaptation to metabolic demand. *EMBO J.* **33**, 2676–2691.
- Pearce, E.L., Walsh, M.C., Cejas, P.J., Harms, G.M., Shen, H., Wang, L.S., Jones, R.G., and Choi, Y. (2009). Enhancing CD8 T-cell memory by modulating fatty acid metabolism. *Nature* **460**, 103–107.
- Pearce, E.L., Poffenberger, M.C., Chang, C.H., and Jones, R.G. (2013). Fueling immunity: insights into metabolism and lymphocyte function. *Science* **342**, 1242454.
- Rambold, A.S., Kostecky, B., Elia, N., and Lippincott-Schwartz, J. (2011). Tubular network formation protects mitochondria from autophagosomal degradation during nutrient starvation. *Proc. Natl. Acad. Sci. USA* **108**, 10190–10195.
- Rambold, A.S., Cohen, S., and Lippincott-Schwartz, J. (2015). Fatty acid trafficking in starved cells: regulation by lipid droplet lipolysis, autophagy, and mitochondrial fusion dynamics. *Dev. Cell* **32**, 678–692.
- Restifo, N.P., Dudley, M.E., and Rosenberg, S.A. (2012). Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat. Rev. Immunol.* **12**, 269–281.
- Röth, D., Krammer, P.H., and Gülow, K. (2014). Dynamin related protein 1-dependent mitochondrial fission regulates oxidative signalling in T cells. *FEBS Lett.* **588**, 1749–1754.
- Samant, S.A., Zhang, H.J., Hong, Z., Pillai, V.B., Sundaresan, N.R., Wolfgeher, D., Archer, S.L., Chan, D.C., and Gupta, M.P. (2014). SIRT3 deacetylates and activates OPA1 to regulate mitochondrial dynamics during stress. *Mol. Cell Biol.* **34**, 807–819.
- Santel, A., Frank, S., Gaume, B., Herrier, M., Youle, R.J., and Fuller, M.T. (2003). Mitofusin-1 protein is a generally expressed mediator of mitochondrial fusion in mammalian cells. *J. Cell Sci.* **116**, 2763–2774.
- Schluns, K.S., Williams, K., Ma, A., Zheng, X.X., and Lefrançois, L. (2002). Cutting edge: requirement for IL-15 in the generation of primary and memory antigen-specific CD8 T cells. *J. Immunol.* **168**, 4827–4831.
- Sebastián, D., Hernández-Alvarez, M.I., Segalés, J., Soriano, E., Muñoz, J.P., Sala, D., Waget, A., Liesa, M., Paz, J.C., Gopalacharyulu, P., et al. (2012). Mitofusin 2 (Mfn2) links mitochondrial and endoplasmic reticulum function with insulin signaling and is essential for normal glucose homeostasis. *Proc. Natl. Acad. Sci. USA* **109**, 5523–5528.
- Sena, L.A., Li, S., Jairaman, A., Prakriya, M., Ezponda, T., Hildeman, D.A., Wang, C.R., Schumacker, P.T., Licht, J.D., Perlman, H., et al. (2013). Mitochondria are required for antigen-specific T cell activation through reactive oxygen species signaling. *Immunity* **38**, 225–236.
- Serasinghe, M.N., Wieder, S.Y., Renault, T.T., Elkholi, R., Ascio, J.J., Yao, J.L., Jabado, O., Hoehn, K., Kageyama, Y., Sesaki, H., and Chipuk, J.E. (2015). Mitochondrial division is requisite to RAS-induced transformation and targeted by oncogenic MAPK pathway inhibitors. *Mol. Cell* **57**, 521–536.
- Smith-Garvin, J.E., Koretzky, G.A., and Jordan, M.S. (2009). T cell activation. *Annu. Rev. Immunol.* **27**, 591–619.
- Taguchi, N., Ishihara, N., Jofuku, A., Oka, T., and Mihara, K. (2007). Mitotic phosphorylation of dynamin-related GTPase Drp1 participates in mitochondrial fission. *J. Biol. Chem.* **282**, 11521–11529.
- Toyama, E.Q., Herzig, S., Courchet, J., Lewis, T.L., Jr., Losón, O.C., Hellberg, K., Young, N.P., Chen, H., Polleux, F., Chan, D.C., and Shaw, R.J. (2016). Metabolism. AMP-activated protein kinase mediates mitochondrial fission in response to energy stress. *Science* **351**, 275–281.
- van der Windt, G.J., Everts, B., Chang, C.H., Curtis, J.D., Freitas, T.C., Amiel, E., Pearce, E.J., and Pearce, E.L. (2012). Mitochondrial respiratory capacity is a critical regulator of CD8+ T cell memory development. *Immunity* **36**, 68–78.
- van der Windt, G.J., O'Sullivan, D., Everts, B., Huang, S.C., Buck, M.D., Curtis, J.D., Chang, C.H., Smith, A.M., Ai, T., Faubert, B., et al. (2013). CD8 memory T cells have a bioenergetic advantage that underlies their rapid recall ability. *Proc. Natl. Acad. Sci. USA* **110**, 14336–14341.
- Vander Heiden, M.G., Cantley, L.C., and Thompson, C.B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033.
- Wai, T., and Langer, T. (2016). Mitochondrial dynamics and metabolic regulation. *Trends Endocrinol. Metab.* **27**, 105–117.
- Wang, R., and Green, D.R. (2012). Metabolic checkpoints in activated T cells. *Nat. Immunol.* **13**, 907–915.
- Wang, D., Wang, J., Bonamy, G.M., Meeusen, S., Brusch, R.G., Turk, C., Yang, P., and Schultz, P.G. (2012). A small molecule promotes mitochondrial fusion in mammalian cells. *Angew. Chem. Int. Ed. Engl.* **51**, 9302–9305.
- Wang, L., Ishihara, T., Ibayashi, Y., Tatsushima, K., Setoyama, D., Hanada, Y., Takeichi, Y., Sakamoto, S., Yokota, S., Mihara, K., et al. (2015). Disruption of mitochondrial fission in the liver protects mice from diet-induced obesity and metabolic deterioration. *Diabetologia* **58**, 2371–2380.
- Williams, M.A., and Bevan, M.J. (2007). Effector and memory CTL differentiation. *Annu. Rev. Immunol.* **25**, 171–192.
- Youle, R.J., and Karbowski, M. (2005). Mitochondrial fission in apoptosis. *Nat. Rev. Mol. Cell Biol.* **6**, 657–663.
- Yu, T., Robotham, J.L., and Yoon, Y. (2006). Increased production of reactive oxygen species in hyperglycemic conditions requires dynamic change of mitochondrial morphology. *Proc. Natl. Acad. Sci. USA* **103**, 2653–2658.
- Zhang, Z., Wakabayashi, N., Wakabayashi, J., Tamura, Y., Song, W.J., Sereda, S., Clerc, P., Polster, B.M., Aja, S.M., Pletnikov, M.V., et al. (2011). The dynamin-related GTPase Opa1 is required for glucose-stimulated ATP production in pancreatic beta cells. *Mol. Biol. Cell* **22**, 2235–2245.

# Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma

Willy Hugo,<sup>1,6,9</sup> Jesse M. Zaretsky,<sup>2,6,9</sup> Lu Sun,<sup>1,6</sup> Chunying Song,<sup>1,6</sup> Blanca Homet Moreno,<sup>3</sup> Siwen Hu-Lieskovan,<sup>3</sup> Beata Berent-Maoz,<sup>3</sup> Jia Pang,<sup>3</sup> Bartosz Chmielowski,<sup>3</sup> Grace Cherry,<sup>3</sup> Elizabeth Seja,<sup>3</sup> Shirley Lomeli,<sup>1,6</sup> Xiangju Kong,<sup>1,6</sup> Mark C. Kelley,<sup>7</sup> Jeffrey A. Sosman,<sup>8</sup> Douglas B. Johnson,<sup>8</sup> Antoni Ribas,<sup>2,3,4,5,6</sup> and Roger S. Lo<sup>1,2,5,6,\*</sup>

<sup>1</sup>Division of Dermatology, Department of Medicine

<sup>2</sup>Department of Molecular and Medical Pharmacology

<sup>3</sup>Division of Hematology & Oncology, Department of Medicine

<sup>4</sup>Division of Surgical Oncology, Department of Surgery

<sup>5</sup>Jonsson Comprehensive Cancer Center

<sup>6</sup>David Geffen School of Medicine

University of California, Los Angeles, CA 90095-1662, USA

<sup>7</sup>Department of Surgery, Vanderbilt-Ingram Cancer Center, Nashville, TN 37232, USA

<sup>8</sup>Department of Medicine, Vanderbilt-Ingram Cancer Center, Nashville, TN 37232, USA

<sup>9</sup>Co-first authors

\*Correspondence: rlo@mednet.ucla.edu

<http://dx.doi.org/10.1016/j.cell.2016.02.065>

## SUMMARY

**PD-1 immune checkpoint blockade provides significant clinical benefits for melanoma patients. We analyzed the somatic mutanomes and transcriptomes of pretreatment melanoma biopsies to identify factors that may influence innate sensitivity or resistance to anti-PD-1 therapy. We find that overall high mutational loads associate with improved survival, and tumors from responding patients are enriched for mutations in the DNA repair gene *BRCA2*. Innately resistant tumors display a transcriptional signature (referred to as the IPRES, or innate anti-PD-1 resistance), indicating concurrent up-expression of genes involved in the regulation of mesenchymal transition, cell adhesion, extracellular matrix remodeling, angiogenesis, and wound healing. Notably, mitogen-activated protein kinase (MAPK)-targeted therapy (MAPK inhibitor) induces similar signatures in melanoma, suggesting that a non-genomic form of MAPK inhibitor resistance mediates cross-resistance to anti-PD-1 therapy. Validation of the IPRES in other independent tumor cohorts defines a transcriptomic subset across distinct types of advanced cancer. These findings suggest that attenuating the biological processes that underlie IPRES may improve anti-PD-1 response in melanoma and other cancer types.**

## INTRODUCTION

PD-1 immune checkpoint blockade therapy induces a high rate of anti-melanoma response and provides unprecedented clinical

benefits (Hamid et al., 2013; Topalian et al., 2012). This therapeutic approach has also been shown to be active against a growing list of human malignancies, and clinical testing of combinations of PD-1 (or PD-L1) with other treatment targets has already begun (Sharma and Allison, 2015). However, effective clinical use of anti-PD-1 agents is encumbered by a high rate of innate resistance (60%–70%) in advanced metastatic melanoma. The mechanistic basis for the variation in response patterns or in long-term clinical benefits (i.e., survival) remains poorly explained.

In melanoma, the extent of pretreatment and especially treatment-induced intra-tumoral T cell infiltration correlates with clinical responses (Tumeh et al., 2014), supporting unleashing of tumor-specific T cells as the primary mechanistic basis of anti-PD-1 therapy. Preliminary retrospective analyses of clinical data hinted at prior failure of mitogen-activated protein kinase (MAPK)-targeted therapy being a negative factor for subsequent response to immune checkpoint blockade in melanoma (Puzanov et al., 2015, *Pigment Cell Melanoma Res.*, abstract; Ramanujam et al., 2015, *Pigment Cell Melanoma Res.*, abstract; Simeone et al., 2015, *Pigment Cell Melanoma Res.*, abstract). In this context, acquired resistance to MAPK-targeted therapy has been correlated with depletion of intra-tumoral T cells, exhaustion of CD8 T cells, and loss of antigen presentation (Hugo et al., 2015).

At the genomic level, the overall mutation load has been correlated with clinical responses to anti-PD-1 therapy and linked to smoking in non-small-cell lung cancer or mismatch repair deficiency in colon cancer. (Le et al., 2015; Rizvi et al., 2015). Whether it was non-small-cell lung tumors on anti-PD-1 therapy or melanoma tumors on anti-CTLA-4 therapy (Snyder et al., 2014; Van Allen et al., 2015), the range of somatic mutation and neoepitope loads of the responding pretreatment tumors overlapped significantly with that of the non-responding tumors, despite statistically significant differences in their medians. As

such, the lack of high mutational loads did not necessarily preclude clinical responses, and conversely, the presence of high mutational loads does not always correlate with responses. Thus, additional genomic or non-genomic features seem likely to contribute to anti-PD-1 response patterns. Here, we sought to assess omic-scale features related to clinical response and survival patterns in order to gain insights into potential strategies for patient stratification and identification of anti-PD-1 combinatorial therapies.

## RESULTS AND DISCUSSION

### High Mutational Load Does Not Associate with Tumor Response but Correlates with Improved Patient Survival

We analyzed the whole-exome sequences (WES) of 38 pretreatment (pembrolizumab and nivolumab) melanoma tumors (responding,  $n = 21$ ; non-responding,  $n = 17$ ; total 34 of 38 pretreatment; 4 of 38 early on-treatment; 14 of 38 patients with prior MAPK inhibitor (MAPKi) treatment; Table S1A) and patient-matched normal tissues for germline references. Responding pretreatment tumors were derived from patients who went on to have complete or partial responses or stable disease control (with mixed responses excluded) in response to anti-PD-1 therapy. Non-responding tumors were derived from patients who had progressive disease. These response patterns were based on irRECIST (Hoos et al., 2015; Wolchok et al., 2009). We also analyzed the transcriptomes through RNA-seq of responding ( $n = 15$ ) and non-responding ( $n = 13$ ) pretreatment tumors (total 27 of 28 pretreatment; 1 of 28 early on-treatment) with available high-quality RNA. WES achieved a median of 140 $\times$  coverage in both tumor and normal tissues (Table S1B). We detected a median of 489 non-synonymous somatic mutations in the 38 tumors (range, 73 to 3,985, which is similar to that in a different set of melanoma tissues; Van Allen et al., 2015).

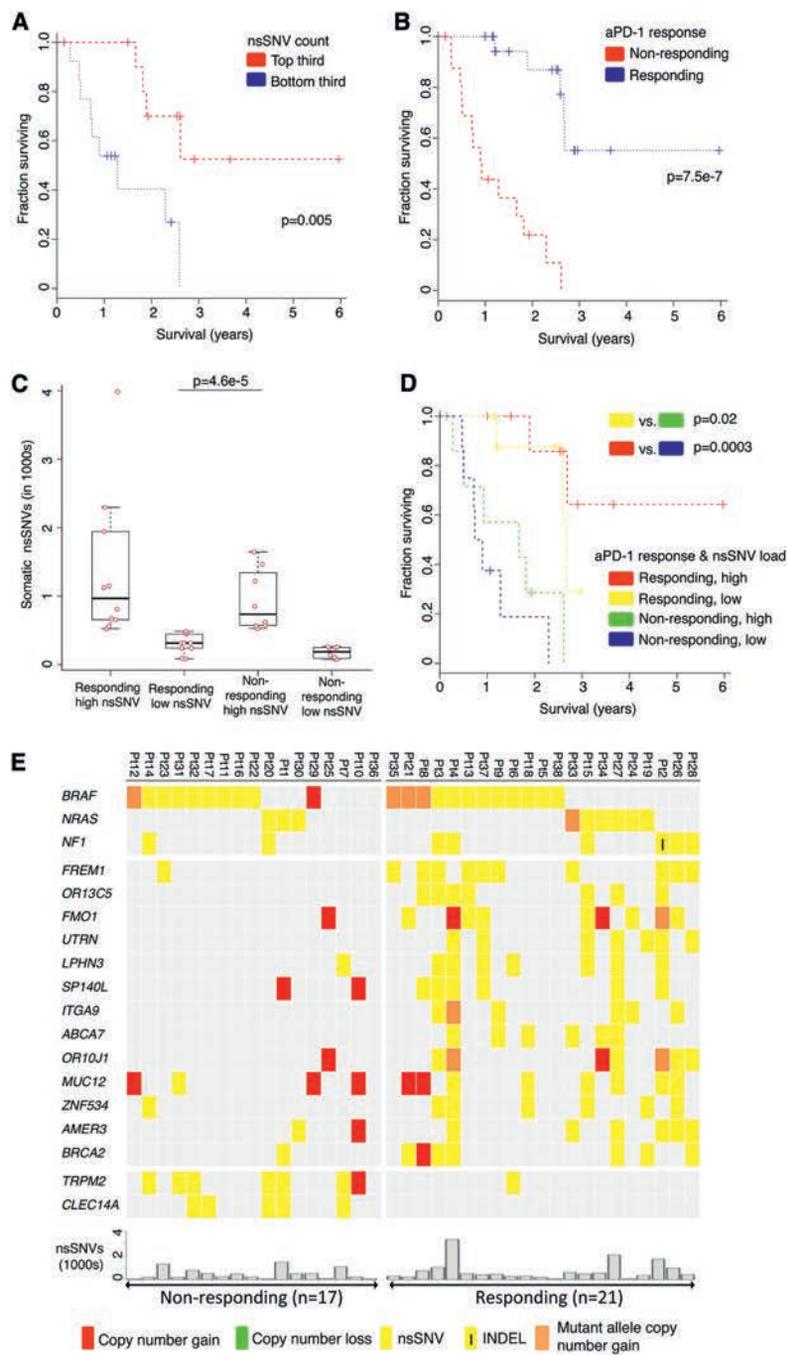
We found that responding pretreatment tumors on anti-PD-1 therapy harbored more non-synonymous single nucleotide variants (nsSNVs) compared to the non-responding tumors, albeit the statistical significance cutoff was not met (median nsSNVs responding = 495 and non-responding = 281,  $p = 0.30$ , Mann-Whitney; Table S1B). Increased predicted HLA (human leukocyte antigen) class I and class II neopeptide loads were also detected in the responding pretreatment tumors, although these differences were not statistically significant either (median HLA class I neopeptides responding = 231 and non-responding = 156,  $p = 0.41$ ; median HLA class II neopeptides responding = 130 and non-responding = 95,  $p = 0.36$ , Mann-Whitney; Table S1B). Even when we considered only expressed nsSNV and neopeptide loads, the statistical significance of the differences between the responding versus non-responding tumors was not augmented. The comparison of these two groups of tumors was not likely biased by small differences in mean tumor purities or depth of sequencing (Figures S1A and S1B). The numbers of predicted HLA class I and II neopeptides were strongly correlated with the number of nsSNVs (Figure S1C). We did not identify any recurrent predicted neopeptide or experimentally validated neoantigens (Table S1C). Previous work analyzing melanoma tumors sampled prior to anti-CTLA-4 antibody therapy had associated responses with a tetrapeptide signature (Snyder

et al., 2014). However, we did not observe enrichment of this peptide motif in the pretreatment tumors that responded to anti-PD-1 therapy (Figure S1D). Likewise, analysis of an independent cohort of 110 melanoma tumors' pre-anti-CTLA-4 therapy also did not yield enrichment of this tetrapeptide motif among responding tumors (Van Allen et al., 2015).

In addition to examining the relationship between non-synonymous somatic mutational loads in pretreatment tumors and anti-tumor responses elicited by anti-PD-1 antibodies, we also examined the potential influences of somatic mutational loads on clinical benefits derived from anti-PD-1 immunotherapy as reflected by patient survival. Notably, a mutational load in the top third (compared to the bottom third) was significantly associated with improved survival (Figure 1A). We also observed a trend toward higher mutational load being associated with better survival among melanoma patients not treated with anti-PD-1 antibodies (Cancer Genome Atlas Network, 2015), although this association did not reach statistical significance (Figure S1E), suggesting that the prognostic power of a high mutational load is augmented in the setting of anti-PD-1 therapy. As expected, a positive association between objective tumor responses and survival was highly statistically significant (Figure 1B). Moreover, when we divided each non-responding and responding tumor group into sub-groups with low or high mutational loads (i.e., below or above the median total somatic nsSNVs of each response group; Figure 1C), patients with responding tumors of low mutation loads significantly outlived patients with non-responding tumors of high mutation loads (Figure 1D). This is despite the fact that mutational loads of these two groups were significantly different, with no overlap across the two distributions (Figure 1C). Hence, factors beyond the mutational load also influence shorter-term tumor response patterns and longer-term patient survival.

### Enrichment for *BRCA2* Mutations in Anti-PD-1 Responsive Melanoma

We then sought to identify mutations (nsSNVs and small insertion-and-deletions [indels]; Table S1D) that (1) were recurrently and selectively associated with either responding or non-responding tumors (recurrence  $\geq 25\%$  in one group and, at most, one hit in the other group) and (2) occurred in genes at rates higher than background rates (Fisher's exact test, false discover rate [FDR]-corrected  $p \leq 0.05$ ; Figure 1E; Table S1E). The background mutation rate of each gene was calculated from the WES data of 469 melanoma tumors (Hodis et al., 2012; Cancer Genome Atlas Network, 2015). Analysis of copy-number variations (CNVs) did not identify any recurrent alteration exclusive to either group. *BRCA2* harbored nsSNVs in 6 of 21 responding tumors (28%), but only 1 of 17 non-responding tumors (6%; Figure 1E). With a background mutational rate estimated at 6% (28 of 469 melanoma tumors), *BRCA2* was significantly more frequently mutated in the responding tumors than expected (Fisher  $p = 0.002$ , odds ratio = 6.2). The pattern of mutations in disparate *BRCA2* protein domains suggested loss-of-function mutations (Figure 1F): one in the N-terminal NPM1-interacting region, one in the POLH-interacting domain, and four in the helical domain critical for FANCD2 interaction. Intriguingly, the somatic mutational load of the tumors with *BRCA2* nsSNVs was



**Figure 1. Mutational Correlates of Innate Sensitivity to Anti-PD-1 Therapy**

(A) Overall survival of anti-PD-1-treated patients whose melanoma tumors harbored high (top third) versus low (bottom third) mutational (somatic nsSNVs) loads; p values, log-rank test.

(B) Overall survival of anti-PD-1-treated melanoma patients whose pretreatment tumors responded (n = 20) or did not respond (n = 17); p value, log-rank test.

(C) Total number of nsSNVs detected in anti-PD-1 responding and non-responding melanoma tumors harboring high (above the respective group’s median) or low (below the group’s median) mutational loads; p value, Mann Whitney test.

(D) Overall survival of anti-PD-1-treated melanoma patients harboring high (above the group median) or low (below the group median) mutational loads and whose pretreatment tumors responded or did not respond; p value, log-rank test.

(E) Recurrent exomic alterations (nsSNVs and small INDELs) in pretreatment tumors of responding versus non-responding patients on anti-PD-1 therapy. Copy-number alterations were annotated for the same genes as a reference. Top, mutations recurrent in responding versus non-responding tumors (recurrence in 25% in one group and at most one occurrence in the opposite group, Fisher’s exact test, FDR-corrected  $p \leq 0.05$  on enrichment against the background mutation frequency). Bottom, the total nsSNV load of each melanoma tumor.

(F) Schematics of impact of missense mutations in the BRCA2 protein and its domains.

(G) Total number of nsSNVs detected in melanomas with or without BRCA2 non-synonymous mutations; p value, Mann Whitney test.

See also Table S1 and Figure S1.

significantly higher than that with wild-type *BRCA2* in this cohort of tumors (Figure 1G), as well as two additional cohorts of melanoma tumors (Figure S1F). Thus, *BRCA2* loss-of-function mutations, which are expected to produce defects in homologous recombination and double-stranded DNA break repair (Holloman, 2011), may produce specific mutational signatures or unknown effects (e.g., induction of cell death) that contribute to anti-PD-1 responsiveness.

### Co-enriched Transcriptomic Signatures in a Major Subset of Anti-PD-1 Resistant Melanoma

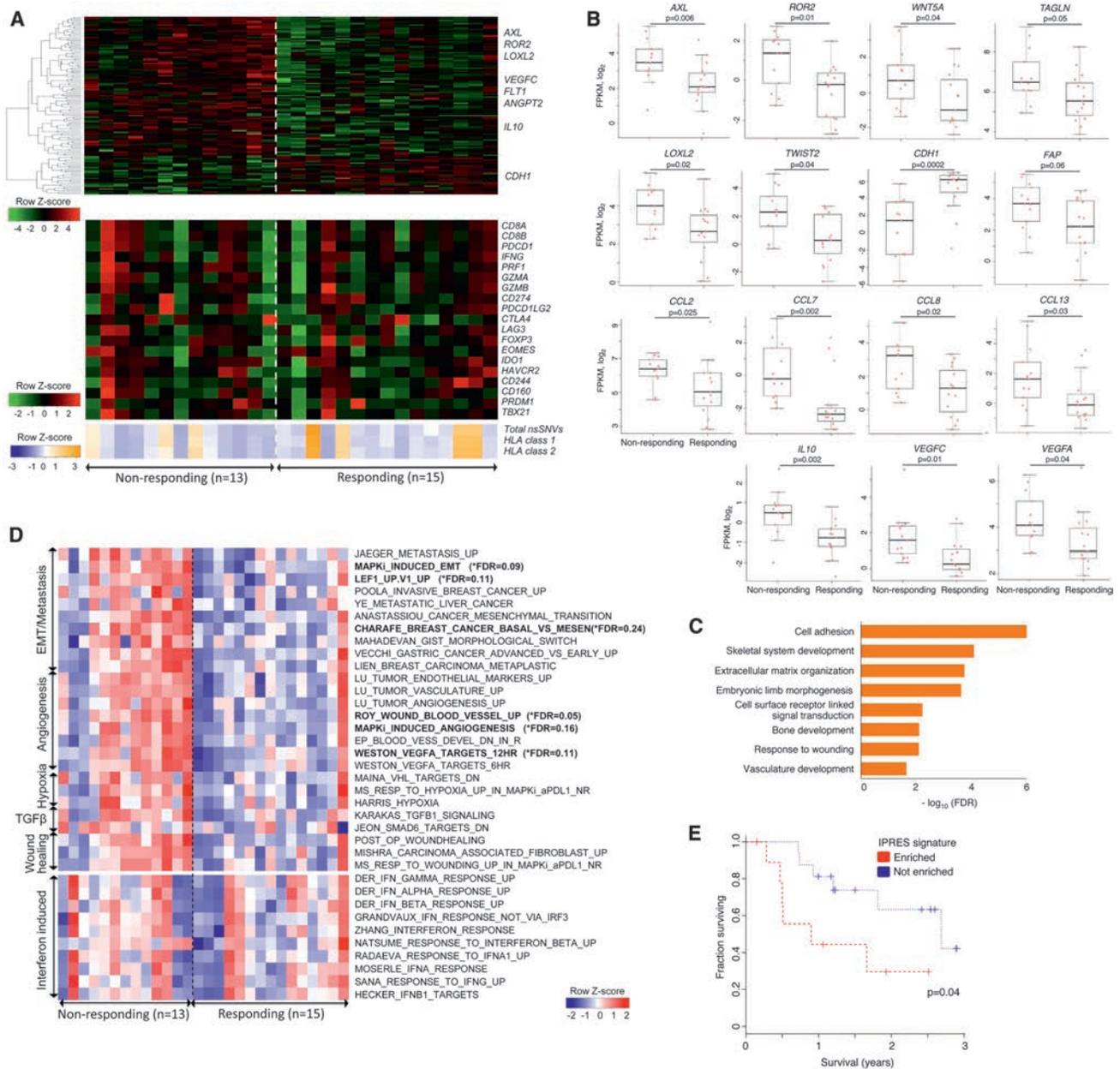
We then addressed whether transcriptomic features would differentiate between responding ( $n = 15$ ) versus non-responding ( $n = 13$ ) tumors sampled prior to anti-PD-1 therapy (total 27 of 28 pretreatment tumors and 1 of 28 early on-treatment). We compared the transcriptomes of the two tumor groups using two approaches: (1) analysis of differentially expressed genes (DEGs; Figures 2A, top, and 2B) across the two aggregate groups (Table S2A) coupled with GO term enrichment analysis of DEGs (Figure 2C) and (2) differential signature enrichment based on single-sample gene set variance analysis, or GSVA, scores using publicly available (C2 chemical and genetic perturbation, C6 oncogenic, and C7 immunologic subsets of the Molecular Signature Database, Broad Institute) and self-curated (see below) perturbation-induced gene signatures (Figure 2D; Table S2B).

From analysis of DEGs (cutoff, 2-fold difference between the absolute medians of normalized expressions in the two groups; nominal Mann-Whitney  $p \leq 0.1$ ), we made observations suggesting that mesenchymal and inflammatory tumor phenotypes may be associated with innate anti-PD-1 resistance. First, 693 genes were differentially expressed between the responding versus non-responding pretreatment tumors, and the transcriptomes of non-responding tumors were dominated by relative gene up-expression events compared with the transcriptomes of responding tumors (Figure 2A, top, showing only genes whose differential expression met nominal Mann-Whitney  $p \leq 0.05$ ; Table S2A). Second, DEGs that were expressed higher in non-responding pretreatment tumors included mesenchymal transition genes (*AXL*, *ROR2*, *WNT5A*, *LOXL2*, *TWIST2*, *TAGLN*, *FAP*), immunosuppressive genes (*IL10*, *VEGFA*, *VEGFC*), and monocyte and macrophage chemotactic genes (*CCL2*, *CCL7*, *CCL8* and *CCL13*; Figures 2A and 2B). In addition to mesenchymal genes, genes associated with wound healing and angiogenesis, which are considered T cell suppressive (Motz and Coukos, 2011; Schäfer and Werner, 2008; Voron et al., 2014), were expressed higher among non-responding relative to responding pretreatment tumors. Interestingly, a recent study using a mouse melanoma model showed that *VEGFA* and *CCL2* expression was associated with innate anti-PD-1 resistance (Peng et al., 2016). *CDH1*, which is typically down-expressed by mesenchymal cancer cells, was also down-expressed by non-responding (versus responding) pretreatment tumors. Third, genes with putative roles in modulating immune checkpoint sensitivity were not differentially expressed between responding versus non-responding tumor groups (Figures 2A, bottom, and S2A). *GZMA*, *PRF1* (CD8 T cell cytolytic score), *PDCD1LG2* (*PD-L2*), and *CTLA4* were expressed higher in the pretreatment mel-

noma tumors of patients who derived benefit from CTLA-4 antibodies (Van Allen et al., 2015). However, these genes, along with other T cell related genes such as *CD8A/B*, *PD-L1*, *LAG3* (T cell checkpoint genes), and *IFNG*, did not present higher expression in anti-PD-1-responsive tumors (Figures 2A, bottom, and S2A). Similarly, we did not observe higher enrichment of multiple interferon signatures in the anti-PD-1-responsive group (Figure 2D, bottom). Previously, an interferon gamma signature was found to be differentially up-expressed in the pretreatment tumor biopsies from responding patients when a restricted set of immune genes were analyzed (Ribas et al., 2015, J. Clin. Oncol., abstract). However, the technical approach may not be comparable to our whole tumor transcriptomic approach. We did note that the expression levels of HLA class I genes (*HLA-A*, *-B*, or *-C*) trended higher among the responding tumors, although the differences were not statistically significant. Lastly, the complete loss of *PTEN* was reported to promote resistance to immune checkpoint blockade (Peng et al., 2016), but there was only one case of homozygous *PTEN* deletion (with nearly undetectable *PTEN* mRNA expression; Figure S2A) in our cohort (in the non-responsive sub-group), limiting our ability to draw meaningful associations in this dataset. Generally, we did not observe a statistically significant difference in *PTEN* expression between anti-PD-1 responding versus non-responding tumors. Thus, individual gene-based expression analysis suggested mesenchymal and T cell suppressive inflammatory or angiogenic tumor phenotypes as being associated with innate anti-PD-1 resistance.

We then queried biological processes represented by DEGs. While gene ontology (GO) enrichment analysis of genes up-expressed among responding tumors produced no significantly enriched terms, genes up-expressed among non-responding tumors were enriched for cell adhesion, extracellular matrix (ECM) organization, wound healing, and angiogenesis (FDR-adjusted  $p$  values of GO gene sets shown in Figure 2C). Using independently derived perturbation-based transcriptomic signatures (Molecular Signature Database; Table S2C), we sought after differentially enriched processes in the responding versus non-responding pretreatment tumors (cutoff, 10% difference between the absolute medians of GSVA scores in the two groups; FDR-corrected Welch  $t$  test  $p \leq 0.25$ ). Gene sets meeting these standard cutoffs formed the core set (Figure 2D, top, in bold) from which we compiled additional concurrently enriched (nominal Welch  $t$  test  $p \leq 0.1$ ) and functionally related gene sets (Figure 2D, top; Table S2B). We considered these statistically weaker gene set enrichments biologically meaningful given the functional coherence of these gene signatures with the core signatures (Subramanian et al., 2005).

Importantly, a group of 26 transcriptomic signatures were co-enriched en bloc in 9 of 13 non-responding versus 1 of 15 responding pre-anti-PD-1 tumors (see Experimental Procedures). Co-enrichment of these signatures, collectively referred to as the innate anti-PD-1 resistance (IPRES) signature, again indicated heightened mesenchymal transition, angiogenesis, hypoxia, and wound healing. The concurrence of a tumor cell mesenchymal phenotype with an angiogenesis- and wound-healing-related inflammatory microenvironment has been documented in the literature (Chen et al., 2014, 2015; Mak et al.,



**Figure 2. Transcriptomic Signatures of Innate Resistance to Anti-PD-1 Therapy**

(A) Top, heatmap showing differentially expressed genes in the pretreatment tumors derived from patients who responded versus who did not respond to anti-PD-1 treatment (gene expression with inter-quartile range [IQR]  $\geq 2$ ; median fold-change [FC] difference  $\geq 2$ ; Mann Whitney  $p \leq 0.05$ ). Middle, mRNA expression levels of genes with hypothetical roles in modulating response patterns to anti-PD-1 therapy. Bottom, overall number of nsSNVs and HLA class 1 and 2 neoepitopes (predicted).

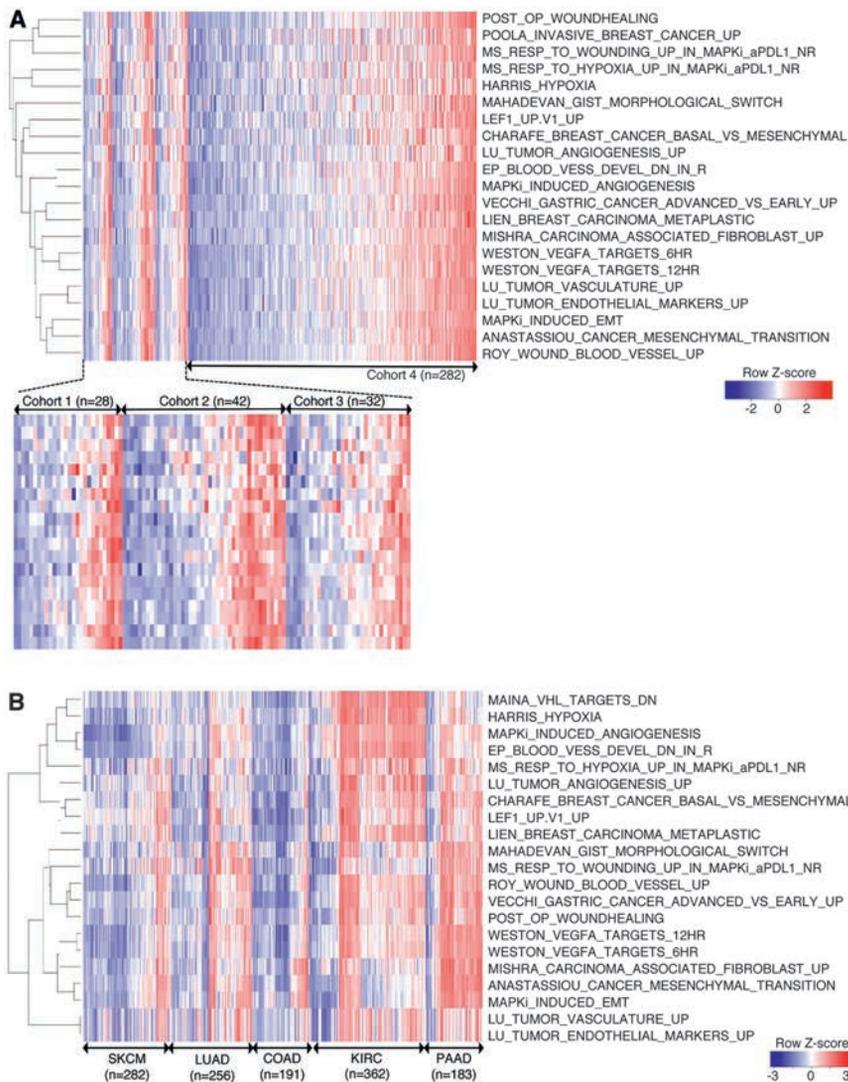
(B) mRNA levels of genes (which control tumor cell mesenchymal transition, tumor angiogenesis, and macrophage and monocyte chemotaxis) that were differentially expressed between the responding versus non-responding pretreatment tumors; p values, Mann Whitney test.

(C) GO enrichment of genes that were expressed higher in the non-responding tumors.

(D) Heatmap showing the gene set variance analysis (GSVA) scores of gene signatures differentially enriched in the responding versus non-responding pre-anti-PD-1 tumors (absolute median GSVA score difference  $\geq 10\%$ , FDR-corrected Welch t test  $p \leq 0.25$  or nominal Welch t test  $p \leq 0.1$ ). For comparison, enrichment scores of interferon signatures are also displayed.

(E) Overall survival of anti-PD-1-treated melanoma patients with presence (n = 10) or absence (n = 16) of co-enriched innate anti-PD-1 resistance (IPRES) signatures; p value, log-rank test.

See also Table S2 and Figure S2.



**Figure 3. Co-enrichment of Innate Anti-PD-1 Resistance-Associated Signatures Defines a Transcriptomic Subset in Melanoma and Multiple Cancers**

(A) Heatmap showing GSVA scores of IPRES signatures across four independent RNA-seq datasets derived from metastatic melanoma. Cohort 1, pretreatment (anti-PD-1) tumors; cohort 2, pretreatment (anti-CTLA-4) tumors; cohort 3, pretreatment (MAPKi) tumors; cohort 4, TCGA cutaneous melanoma (metastatic only).

(B) Heatmap showing GSVA scores of IPRES signatures across TCGA RNA-seq datasets (metastatic melanoma, SKCM; lung adenocarcinoma, LUAD; colon adenocarcinoma, COAD; kidney clear cell carcinoma, KIRC; and pancreatic adenocarcinoma, PAAD). See also Figure S3.

was an exclusive feature of our cohort, we queried three additional cohorts of metastatic melanoma-derived RNA-seq (Hugo et al., 2015; Cancer Genome Atlas Network, 2015; Van Allen et al., 2015), including a cohort consisting of only <sup>V600</sup>BRAF mutant melanomas (cohort 3; Hugo et al., 2015). We found that IPRES content signatures co-enriched not only in the same tumors, but also in about a third of total samples in each of the four independent transcriptomic datasets (cohort 1 from this study, 10 IPRES-enriched tumors of 28 total tumors; cohort 2, 15 of 42; cohort 3, 11 of 32; cohort 4, 90 of 282; Figure 3A). Considering 126 among 384 total tumors as the background prevalence for co-enrichment of IPRES content signatures in metastatic melanoma, we determined that this

2016). Interestingly, this set of 26 IPRES signatures included signatures induced by MAPKi treatment of melanoma tumors and cell lines (Table S2C). We have shown recently that MAPKi treatment of melanoma cells induces transcriptome-wide reprogramming, leading to concurrent phenotype switches (C.S., L.S., W.H., and R.S.L., unpublished data). Notably, MAPKi-induced signatures of mesenchymal-invasive transition, angiogenesis, and wound-healing signatures were detected in the residual melanoma tumors from patients on MAPKi therapy, suggesting that induction of these signatures may negatively impact responsiveness to combinatorial anti-PD-1/L1 therapy.

### IPRES Signatures Define a Transcriptomic Subset across Cancers

The observations that IPRES content signatures were co-enriched in the same tumors (Figure 2D) and that MAPKi induced these signatures concurrently (Table S2C) implied co-regulated tumor phenotypes that together define a transcriptomic subset. To evaluate whether co-enrichment of IPRES content signatures

IPRES-enriched transcriptomic subset was over-represented among the anti-PD-1 non-responding pretreatment tumors (Fisher  $p = 0.013$ , odds ratio = 4.6) and under-represented among the responding pretreatment tumors (Fisher  $p = 0.04$ , odds ratio = 0.15) within cohort 1. In contrast, co-enrichment of IPRES signatures was neither over- nor under-represented among the responding or non-responding pre-anti-CTLA-4 melanoma tumors in cohort 2 (Figure S2B; Van Allen et al., 2015), which suggests that mechanisms of innate resistance to anti-PD-1 and anti-CTLA-4 are not necessarily similar.

Furthermore, co-enrichment of the IPRES signatures defined a transcriptomic subset within not only melanoma, but also all major common human malignancies analyzed (Figure 3B). The IPRES-enriched transcriptomic subset of certain cancers, such as pancreatic adenocarcinoma, made up the majority of tumors. Within a side-by-side comparison, only 6 of 69 primary cutaneous melanomas showed co-enrichment of IPRES signatures, in contrast to 90 of 282 metastatic (The Cancer Genome Atlas [TCGA]) melanomas ( $p = 3.9 \times 10^{-5}$ , odds ratio = 0.2; Figure S3),

consistent with mesenchymal transition and metastasis gene sets among IPRES signatures. Thus, co-enrichment of IPRES signatures defines a distinct transcriptomic program that exists across distinct cancer histologies.

This study highlights the utility of both exome and transcriptome sequencing data generated from pretreatment tumor samples for the identification of potential determinants of response to anti-PD-1. Although the overall somatic mutational loads of anti-PD-1-responsive melanoma tumors were not significantly higher than those of non-responsive tumors, higher mutational loads associated significantly with better survival after anti-PD-1 therapy. This finding is still consistent with the notion that neoepitopes derived from somatic non-synonymous mutations are critical for deriving clinical benefits from anti-PD-1 therapy in melanoma. However, objective tumor responses, although strongly associated with survival benefits, did not appear to be driven overwhelmingly by the overall somatic mutational loads. That is to say, a relatively low mutational load did not preclude a tumor response. This is consistent with findings from gastrointestinal cancers where low mutational loads did not preclude tumor infiltration by mutation-reactive, class I- and II-restricted T cells (Tran et al., 2015). Thus, overall somatic or predicted neoepitope loads of pretreatment melanoma tumors are not enough to predict response patterns to anti-PD-1 therapy.

In our cohort, responsive tumors were significantly enriched for (likely) loss-of-function mutations in *BRCA2*. As one would predict from the known function of *BRCA2* in DNA repair, *BRCA2*-mutated melanomas harbored higher mutational loads than *BRCA2*-wild-type melanomas. Although it is conceivable that defective *BRCA2*-DNA repair results in specific mutational motifs (rather than or in addition to the general increase in mutational load) that enhance responsiveness, it is also possible that cellular stress resulting from defective DNA repair could lead to increased cell death and anti-tumor immunity. Moreover, these data support the notion that tumor cell phenotypic plasticity (i.e., mesenchymal transition) and the resultant impacts on the microenvironment (e.g., ECM remodeling, cell adhesion, and angiogenesis features of immune suppressive wound healing) are critical barriers to anti-PD-1 responses. The limited number of patients in our melanoma cohort posed certain challenges to our analysis. For example, we relaxed the statistical stringency in single gene-based differential expression analysis (bypassing multiple hypothesis correction) to derive enough genes for GO enrichment analysis. However, converging findings from alternative analysis (i.e., GSEA) of the transcriptome data helped to mitigate potential caveats. Finally, in separate work, we found that mutation-targeted therapy (i.e., MAPKi) induces tumor cell-autonomous changes (e.g., mesenchymal transition; C.S., L.S., W.H., and R.S.L., unpublished data) and upregulates anti-PD-1 resistance-associated processes in residual tumors that have regressed in response to MAPKi treatment. Thus, while our findings in this study necessitate confirmation in independent tissue cohorts, the identification of transcriptomic features associated with anti-PD-1 resistance suggests that mitigation of IPRES-related biological processes may enhance response rates to anti-PD-1 (and anti-PD-1 plus MAPKi) therapy.

## EXPERIMENTAL PROCEDURES

### Tumor Specimens and Profiling

All tissues in this study were obtained with the approval of institutional review boards and patients' consents. All patients received either pembrolizumab or nivolumab as the anti-PD-1 therapy for their metastatic melanoma. 38 melanoma specimens (32 pre-treatment tumors, 2 pretreatment tumor-derived cultures, 3 early on-treatment tumors without response, and 1 early on-treatment tumor with response) and their patient-matched normal tissues were analyzed by WES. Among these 38 samples with WES data, 28 with sufficient RNA quality were also analyzed by RNA sequencing (RNA-seq). This set included another RNA-seq dataset derived from a second-site, pre-treatment tumor biopsy from patient #27. However, this second-site, pre-treatment tumor-derived WES dataset was excluded in our aggregate mutation analysis to avoid double counting two tumor exomes from the same patient (Table S1A).

38 tumor specimens and their respective normal tissues were subjected to WES (Table S1A). WES was performed using pair-end sequencing with read length of  $2 \times 100$  bps based on the Illumina HiSeq2000 platform. RNA from a subset of 28 tumors were pair-end sequenced with read length of  $2 \times 100$  bps (Illumina HiSeq2000). We included two tumors from patient #27 for transcriptomic analyses, but not for mutation and neoepitope analyses, since the tumors may not share the same transcriptomic profile, but they essentially contain the same set of non-synonymous somatic mutations.

### Whole Exome Sequencing

We called single nucleotide variant (SNV) and INDEL as reported (Shi et al., 2014) using a stand-alone version of Oncotator (Ramos et al., 2015). Copy numbers were called using the intersection of the copy-number calls derived from Sequenza (Favero et al., 2015) and VarScan2 (Koboldt et al., 2012). Tumor purities and ploidies (Table S1B) were calculated based on the calls of Sequenza using WES data with default parameters. The impact of *BRCA2* nsSNVs was visualized using the domain information in the INTERPRO protein domain database (Mitchell et al., 2015).

### HLA Types and Neoepitopes

The four-digit HLA class I and II types of each patient were called using ATHLATES (Liu et al., 2013) using the WES-sequencing reads from the normal tissue. To ensure concordance, we manually compared ATHLATES' calls of the normal versus tumor samples and ascertained that there was at least no two-digit HLA typing discrepancy between any normal tumor pair. For each non-synonymous coding mutation from a tumor, we predicted its impact on the patient's HLA class I and II binding using the standalone version of the programs NetMHCpan v2.8 (Hoof et al., 2009; Nielsen et al., 2007) and NetMHCIIpan v3.0 (Karosiene et al., 2013), respectively. Specifically, for HLA class I binding prediction using netMHCpan v2.8, we tested all 9-11-mer peptides containing the mutated amino acids for binding to the patient's HLA-A, -B, and -C. A peptide was defined as a neoepitope based on two criteria: (1) predicted binding affinity  $\leq 500$ nM and (2) rank percentage  $\leq 2\%$  (default cutoff). For HLA class II binding prediction using netMHCIIpan v3.0, we tested the 9-19-mers containing the mutated amino acids for binding to the patient-specific, ATHLATES-predicted DPA-DPB, DQA-DQB, and DRA-DRB allele pairs. We also applied the same predicted binding affinity and rank percentage cutoff as we did for HLA class I to nominate the HLA class-II-binding neoepitopes. Expressed non-synonymous mutations and neoepitopes were defined based on corresponding genes with normalized expression levels  $\geq 1$  (in fragments per kilobase of transcripts per million mapped reads [FPKM]). Statistical differences of nsSNV, HLA class I and II neoepitopes, WES coverages, and tumor purities were computed using two-sided Mann-Whitney test.

### Mutation Recurrence

To estimate the statistical significance of the recurrence of gene mutations in the responding or non-responding tumors, we used an independent batch of 469 melanomas' whole-exome sequence datasets (Hodis et al., 2012; Cancer Genome Atlas Network, 2015) to estimate each gene's background mutation frequency. Significance was computed by Fisher's exact test followed by FDR

adjustment for multiple hypothesis testing. We listed genes that fulfilled the criteria: (1) recurrence in at least 25% of the responder/non-responder, (2) occurrence of at most once in the opposite group, and (3) Fisher's exact test FDR-adjusted  $p$  value  $\leq 0.05$ . These genes were illustrated in Figure 1A, and all genes that fulfilled the first and second criteria and tested for multiple hypotheses were listed in Table S1E. The association between *BRCA2* nsSNVs and overall nsSNV counts were tested using two-sided Mann-Whitney test and validated in independent WES datasets (Hodis et al., 2012; Cancer Genome Atlas Network, 2015).

### RNA-Seq and Gene Set Enrichment

Paired-end transcriptome reads were mapped to the UCSC hg19 reference genome using Tophat2 (Kim et al., 2013). Normalized expression levels of genes were expressed in FPKM values as generated by cuffquant and cuffnorm (Trapnell et al., 2012). The program was run with the option “-frag-bias-correct” and “-multi-read-correct” to improve sensitivity (Roberts et al., 2011). A gene was defined as differentially expressed between the responding and non-responding tumor groups when its median expression differed by at least 2-fold between the groups with a nominal two-sided Mann-Whitney  $p$  value  $\leq 0.1$  (Table S2A). Applying multiple hypothesis correction of FDR  $p \leq 0.25$  only yielded three differentially expressed genes: *ALDH1L2* and *MFAP2* in the non-responding and *CDH1* (E-cadherin) in the responding group. As such, the genes meeting the uncorrected, nominal Mann-Whitney  $p$  value  $\leq 0.1$  that were expressed higher either in the responding or non-responding group were separately analyzed for GO term enrichments using the online functional annotation tools DAVID (Huang et al., 2009a). Enriched GO terms were selected from the GO biological process terms in DAVID's fat database (Huang et al., 2009b). GO terms which were highly overlapping, as defined by functional clustering in DAVID's website, were represented by the terms with the best FDR-adjusted  $p$  values.

To calculate single-sample gene set enrichment, we used the GSEA program (Hänzelmann et al., 2013) to derive the absolute enrichment scores of previously experimentally validated gene signatures as follows: (1) the C2 CGP (chemical and genetic perturbation sets), (2) the C6 and C7 subset of the Molecular Signature Database version 4.0 (Subramanian et al., 2005), (3) self-curated MAPKi-induced gene signatures using cell lines and patient-derived tumors (C.S., L.S., W.H., and R.S.L., unpublished data), (4) post-operation wound signature (Inkeles et al., 2015), and (5) melanoma invasive/proliferative signatures (Hoek et al., 2008). To derive the GSEA score of each signature in each tumor sample, we computed from raw RNA-seq read counts by HTSEQ COUNT program and then normalized them to  $\log_2$  counts per million (CPM) values using EdgeR (McCarthy et al., 2012). We removed batch effects using the edgeR function `RemoveBatchEffect` when we combined RNA-seq data from multiple experiments (Figure 3A). The normalized  $\log_2$  CPM values were then passed on as input for GSEA in the RNA-seq mode. Differentially enriched core gene sets between the responding and non-responding tumor groups were defined by GSEA score differences of  $\geq 10\%$  and FDR-corrected, two-sided Welch  $t$  test  $p$  value  $\leq 0.25$  (we used  $t$  test because the GSEA scores were normally distributed around zero). Two gene sets, INGRAM\_SHH\_TARGETS\_DN and WONG\_ENDMETRIUM\_CANCER\_DN, were not included in the core set because they did not specifically point to a cellular process and/or relate to the other six gene sets in the core set (Table S2B, top eight). We also collected gene sets that met the GSEA score differences of  $\geq 10\%$  and nominal Welch  $t$  test  $p$  value  $\leq 0.1$  (Table S2B) and included those which were concordantly enriched and functionally related to the core gene sets to make up the full list of IPRES signatures (Figure 2D).

To compare co-enrichment of IPRES signatures across multiple melanoma cohorts, we combined and batch-corrected the  $\log_2$  CPM values of four melanoma transcriptome cohorts: (1) our current pre-anti-PD-1 tumors ( $n = 28$ ), (2) pre-anti-CTLA-4 tumors ( $n = 42$ ), (3) pre-MAPKi tumors ( $n = 32$ ), and (4) the metastatic subset of TCGA melanoma ( $n = 282$ ). We row-normalized the GSEA scores of each gene set in the IPRES signature across the samples from the four cohorts. For this comparative study, we excluded the gene sets JAEGER\_METASTASIS\_UP, YE\_METASTASIS\_LIVER\_CANCER, KARAKAS\_TGFB1\_SIGNALING, and JEON\_SMAD6\_TARGETS\_DN from the IPRES set because they showed weaker co-enrichment with rest of the gene sets (see

Figure 2D, top). The IPRES (enrichment) score was defined as the average  $Z$  score across all gene sets in the IPRES signature, and we applied an average  $Z$  score of 0.35 as the cutoff for IPRES signature enrichment in a tumor sample. This resulted in IPRES co-enrichment in nine non-responding tumors and one responding tumor in our anti-PD-1 cohort (this cutoff was chosen because it provided the largest average  $Z$  score separation between the samples with and without IPRES co-enrichment). Since the IPRES score was not comparable across analyses in Figures 3A, 3B, and S3 (since the IPRES score is  $Z$  score-base), we used the 90<sup>th</sup> highest IPRES score in the TCGA metastatic melanoma cohort as the IPRES score cutoff (since there were 90 of 282 tumors showing IPRES co-enrichment in this TCGA metastatic cohort; Figure 3A) for analyses performed to yield Figures 3B and S3. This allowed for a non-parametric comparison across multiple TCGA datasets at the IPRES co-enrichment level established in our anti-PD-1 cohort.

### Source Data

Analysis of differential non-synonymous mutational hits in responders versus non-responders to ipilimumab was based on the mutation calls as reported (Van Allen et al., 2015). We curated published CD8 T cell exhaustion genes (Wherry, 2011) to minimize those likely to be expressed by melanoma cells by excluding genes whose maximum  $\log_2$  FPKM was 1 in an in-house melanoma cell-line-derived RNA-seq database ( $n = 26$  cell lines). This resulted in the inclusion of genes for surface receptors *PDCD1* (*PD-1*), *LAG3*, *HAVCR2* (*Tim-3*), *CD160*, and *CD244* as well as transcription factors *EOMES*, *PRDM1* (*Blimp-1*), and *TBX21* (*T-bet*). We assessed co-enrichment of IPRES content signatures in the (1) anti-CTLA-4 pretreatment cohort (Van Allen et al., 2015), (2) MAPKi pretreatment cohort (Hugo et al., 2015; C.S., L.S., W.H., and R.S.L., unpublished data), (3) TCGA melanoma (metastatic and primary subsets separately analyzed; Cancer Genome Atlas Network, 2015), (4) TCGA pancreatic ductal adenocarcinoma (TCGA, provisional, see <https://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp>), (5) TCGA lung adenocarcinoma (Cancer Genome Atlas Research Network, 2014), (6) TCGA colorectal adenocarcinoma (Cancer Genome Atlas Network, 2012), and (7) TCGA kidney clear cell carcinoma (Cancer Genome Atlas Research Network, 2013).

### ACCESSION NUMBERS

The accession number for the transcriptome data is GEO: GSE78220.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.02.065>.

### AUTHOR CONTRIBUTION

W.H., J.M.Z., L.S., C.S., A.R., and R.S.L. generated, analyzed, and interpreted the data. S.H.-L., B.H.M., B.C., G.C., E.S., M.C.K., J.A.S., D.B.J., A.R., and R.S.L. evaluated patients and provided tissue reagents. B.B.-M., J.P., S.L., and X.K. processed tissues for analysis. All authors contributed to manuscript preparation. W.H. and R.S.L. developed the concepts, supervised the project, and wrote the paper.

### ACKNOWLEDGMENTS

This work has been funded by the NIH (R35 CA197633 to A.R., P01 CA168585 to A.R. and R.S.L., 1R01CA176111 to R.S.L., and K12CA0906525 to D.B.J.), the Ressler Family Foundation (to R.S.L. and A.R.), the Grimaldi Family Fund (to A.R. and R.S.L.), the Burroughs Wellcome Fund (to R.S.L.), the Ian Copeland Melanoma Fund (to R.S.L.), the Wade F.B. Thompson/Cancer Research Institute CLIP Grant (to R.S.L.), the Steven C. Gordon Family Foundation (to R.S.L.), the American Skin Association (to W.H.), the Dr. Robert Vigen Memorial Fund (to A.R.), the Garcia-Corsini Family Fund (to A.R.), the ASCO Conquer Cancer Career Development Award (to D.B.J.), and the American Cancer Society Research Professorship (to J.A.S.). Informed consents were obtained from patients for the research performed in this study.

Received: December 25, 2015  
 Revised: January 30, 2016  
 Accepted: February 29, 2016  
 Published: March 17, 2016

## REFERENCES

- Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330–337.
- Cancer Genome Atlas Network (2015). Genomic Classification of Cutaneous Melanoma. *Cell* 161, 1681–1696.
- Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 499, 43–49.
- Cancer Genome Atlas Research Network (2014). Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511, 543–550.
- Chen, L., Gibbons, D.L., Goswami, S., Cortez, M.A., Ahn, Y.H., Byers, L.A., Zhang, X., Yi, X., Dwyer, D., Lin, W., et al. (2014). Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression. *Nat. Commun.* 5, 5241.
- Chen, L., Heymach, J.V., Qin, F.X., and Gibbons, D.L. (2015). The mutually regulatory loop of epithelial-mesenchymal transition and immunosuppression in cancer progression. *Oncol Immunology* 4, e1002731.
- Favero, F., Joshi, T., Marquard, A.M., Birkbak, N.J., Krzystanek, M., Li, Q., Szallasi, Z., and Eklund, A.C. (2015). Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol.* 26, 64–70.
- Hamid, O., Robert, C., Daud, A., Hodi, F.S., Hwu, W.J., Kefford, R., Wolchok, J.D., Hersey, P., Joseph, R.W., Weber, J.S., et al. (2013). Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N. Engl. J. Med.* 369, 134–144.
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 14, 7.
- Hodis, E., Watson, I.R., Kryukov, G.V., Arold, S.T., Imielinski, M., Theurillat, J.P., Nickerson, E., Auclair, D., Li, L., Place, C., et al. (2012). A landscape of driver mutations in melanoma. *Cell* 150, 251–263.
- Hoek, K.S., Eichhoff, O.M., Schlegel, N.C., Döbbling, U., Kobert, N., Schärer, L., Hemmi, S., and Dummer, R. (2008). In vivo switching of human melanoma cells between proliferative and invasive states. *Cancer Res.* 68, 650–656.
- Holloman, W.K. (2011). Unraveling the mechanism of BRCA2 in homologous recombination. *Nat. Struct. Mol. Biol.* 18, 748–754.
- Hoof, I., Peters, B., Sidney, J., Pedersen, L.E., Sette, A., Lund, O., Buus, S., and Nielsen, M. (2009). NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* 61, 1–13.
- Hoos, A., Wolchok, J.D., Humphrey, R.W., and Hodi, F.S. (2015). CCR 20th Anniversary Commentary: Immune-Related Response Criteria—Capturing Clinical Activity in Immuno-Oncology. *Clin. Cancer Res.* 21, 4989–4991.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009b). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13.
- Hugo, W., Shi, H., Sun, L., Piva, M., Song, C., Kong, X., Moriceau, G., Hong, A., Dahlman, K.B., Johnson, D.B., et al. (2015). Non-genomic and Immune Evolution of Melanoma Acquiring MAPKi Resistance. *Cell* 162, 1271–1285.
- Inkeles, M.S., Scumpia, P.O., Swindell, W.R., Lopez, D., Teles, R.M., Graeber, T.G., Meller, S., Homey, B., Elder, J.T., Gilliet, M., et al. (2015). Comparison of molecular signatures from multiple skin diseases identifies mechanisms of immunopathogenesis. *J. Invest. Dermatol.* 135, 151–159.
- Karosiene, E., Rasmussen, M., Blicher, T., Lund, O., Buus, S., and Nielsen, M. (2013). NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics* 65, 711–724.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.
- Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576.
- Le, D.T., Uram, J.N., Wang, H., Bartlett, B.R., Kemberling, H., Eyring, A.D., Skora, A.D., Luber, B.S., Azad, N.S., Laheru, D., et al. (2015). PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* 372, 2509–2520.
- Liu, C., Yang, X., Duffy, B., Mohanakumar, T., Mitra, R.D., Zody, M.C., and Pfeifer, J.D. (2013). ATHLATES: accurate typing of human leukocyte antigen through exome sequencing. *Nucleic Acids Res.* 41, e142.
- Mak, M.P., Tong, P., Diao, L., Cardnell, R.J., Gibbons, D.L., William, W.N., Skoulidis, F., Parra, E.R., Rodriguez-Canales, J., Wistuba, I.I., et al. (2016). A Patient-Derived, Pan-Cancer EMT Signature Identifies Global Molecular Alterations and Immune Target Enrichment Following Epithelial-to-Mesenchymal Transition. *Clin. Cancer Res.* 22, 609–620.
- McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297.
- Mitchell, A., Chang, H.Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S., et al. (2015). The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* 43, D213–D221.
- Motz, G.T., and Coukos, G. (2011). The parallel lives of angiogenesis and immunosuppression: cancer and other tales. *Nat. Rev. Immunol.* 11, 702–711.
- Nielsen, M., Lundegaard, C., Blicher, T., Lamberth, K., Harndahl, M., Justesen, S., Roder, G., Peters, B., Sette, A., Lund, O., and Buus, S. (2007). NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS ONE* 2, e796.
- Peng, W., Chen, J.Q., Liu, C., Malu, S., Creasy, C., Tetzlaff, M.T., Xu, C., McKenzie, J.A., Zhang, C., Liang, X., et al. (2016). Loss of PTEN promotes resistance to T cell-mediated immunotherapy. *Cancer Discov.* 6, 202–216.
- Ramos, A.H., Lichtenstein, L., Gupta, M., Lawrence, M.S., Pugh, T.J., Sakseena, G., Meyerson, M., and Getz, G. (2015). Oncotator: cancer variant annotation tool. *Hum. Mutat.* 36, E2423–E2429.
- Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., et al. (2015). Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128.
- Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L., and Pachter, L. (2015). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.* 12, R22.
- Schäfer, M., and Werner, S. (2008). Cancer as an overheating wound: an old hypothesis revisited. *Nat. Rev. Mol. Cell Biol.* 9, 628–638.
- Sharma, P., and Allison, J.P. (2015). Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. *Cell* 161, 205–214.
- Shi, H., Hugo, W., Kong, X., Hong, A., Koya, R.C., Moriceau, G., Chodon, T., Guo, R., Johnson, D.B., Dahlman, K.B., et al. (2014). Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy. *Cancer Discov.* 4, 80–93.
- Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J.M., Desrichard, A., Walsh, L.A., Postow, M.A., Wong, P., Ho, T.S., et al. (2014). Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* 371, 2189–2199.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach

- for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102, 15545–15550.
- Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., et al. (2012). Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* 366, 2443–2454.
- Tran, E., Ahmadzadeh, M., Lu, Y.C., Gros, A., Turcotte, S., Robbins, P.F., Gartner, J.J., Zheng, Z., Li, Y.F., Ray, S., et al. (2015). Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* 350, 1387–1390.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578.
- Tumeh, P.C., Harview, C.L., Yearley, J.H., Shintaku, I.P., Taylor, E.J., Robert, L., Chmielowski, B., Spasic, M., Henry, G., Ciobanu, V., et al. (2014). PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515, 568–571.
- Van Allen, E.M., Miao, D., Schilling, B., Shukla, S.A., Blank, C., Zimmer, L., Sucker, A., Hillen, U., Foppen, M.H., Goldinger, S.M., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350, 207–211.
- Voron, T., Marcheteau, E., Pernot, S., Colussi, O., Tartour, E., Taieb, J., and Terme, M. (2014). Control of the immune response by pro-angiogenic factors. *Front. Oncol.* 4, 70.
- Wherry, E.J. (2011). T cell exhaustion. *Nat. Immunol.* 12, 492–499.
- Wolchok, J.D., Hoos, A., O'Day, S., Weber, J.S., Hamid, O., Lebbé, C., Maio, M., Binder, M., Bohnsack, O., Nichol, G., et al. (2009). Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin. Cancer Res.* 15, 7412–7420.

# Progressive Loss of Function in a Limb Enhancer during Snake Evolution

Evgeny Z. Kvon,<sup>1</sup> Olga K. Kamneva,<sup>2</sup> Uirá S. Melo,<sup>1</sup> Iros Barozzi,<sup>1</sup> Marco Osterwalder,<sup>1</sup> Brandon J. Mannion,<sup>1</sup> Virginie Tissières,<sup>3</sup> Catherine S. Pickle,<sup>1</sup> Ingrid Plajzer-Frick,<sup>1</sup> Elizabeth A. Lee,<sup>1</sup> Momoe Kato,<sup>1</sup> Tyler H. Garvin,<sup>1</sup> Jennifer A. Akiyama,<sup>1</sup> Veena Afzal,<sup>1</sup> Javier Lopez-Rios,<sup>3</sup> Edward M. Rubin,<sup>1,4</sup> Diane E. Dickel,<sup>1</sup> Len A. Pennacchio,<sup>1,4,\*</sup> and Axel Visel<sup>1,4,5,6,\*</sup>

<sup>1</sup>MS 84-171, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

<sup>2</sup>Department of Biology, Stanford University, Stanford, CA 94305, USA

<sup>3</sup>Department of Biomedicine, University of Basel, 4058 Basel, Switzerland

<sup>4</sup>U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA

<sup>5</sup>School of Natural Sciences, University of California, Merced, CA 95343, USA

<sup>6</sup>Lead Contact

\*Correspondence: lapennacchio@lbl.gov (L.A.P.), avisel@lbl.gov (A.V.)

<http://dx.doi.org/10.1016/j.cell.2016.09.028>

## SUMMARY

The evolution of body shape is thought to be tightly coupled to changes in regulatory sequences, but specific molecular events associated with major morphological transitions in vertebrates have remained elusive. We identified snake-specific sequence changes within an otherwise highly conserved long-range limb enhancer of *Sonic hedgehog* (*Shh*). Transgenic mouse reporter assays revealed that the *in vivo* activity pattern of the enhancer is conserved across a wide range of vertebrates, including fish, but not in snakes. Genomic substitution of the mouse enhancer with its human or fish ortholog results in normal limb development. In contrast, replacement with snake orthologs caused severe limb reduction. Synthetic restoration of a single transcription factor binding site lost in the snake lineage reinstated full *in vivo* function to the snake enhancer. Our results demonstrate changes in a regulatory sequence associated with a major body plan transition and highlight the role of enhancers in morphological evolution.

## INTRODUCTION

Distant-acting transcriptional enhancers are a major class of tissue-specific regulatory DNA sequences that has been implicated in morphological evolution in vertebrates (Chan et al., 2010; Cooper et al., 2014; Cretekos et al., 2008; Guenther et al., 2014; Guerreiro et al., 2013; Indjeian et al., 2016; Jones et al., 2012; Lopez-Rios et al., 2014; McLean et al., 2011; Prabhakar et al., 2008). Sequence changes in non-coding regulatory DNA are hypothesized to be a main driver of changes in body shape (Britten and Davidson, 1969; Carroll, 2008; King and Wilson, 1975; Wray, 2007), but many aspects of this complex interplay between molecular changes in regulatory sequences and morphological adaptations across the vertebrate tree remain

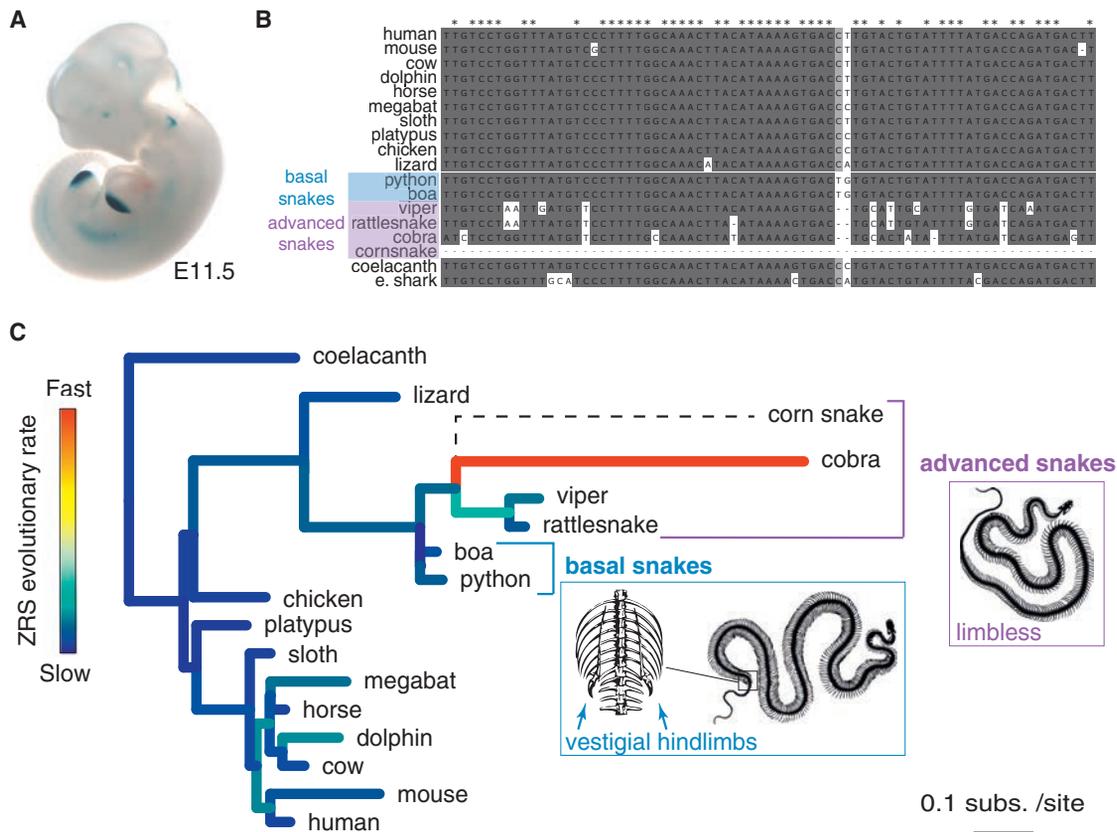
the subject of considerable debate (Hoekstra, 2012; Wittkopp and Kalay, 2011; Wray, 2007).

In the present study, we utilized a series of recently sequenced snake genomes to study the molecular and functional evolution of a critical limb enhancer in snakes and examine its possible role in limb loss. Our analysis focuses on one of the best-studied vertebrate enhancers, the Zone of Polarizing Activity [ZPA] Regulatory Sequence (ZRS, also known as MFCS1) (Lettice et al., 2003, 2008, 2012, 2014; Sagai et al., 2004, 2005; Zeller and Zuniga, 2007). The ZRS is a limb-specific enhancer of the *Sonic hedgehog* (*Shh*) gene that is located at the extreme distance of nearly one million base pairs from its target promoter. During limb development, the enhancer is active in the posterior limb bud mesenchyme (Figure 1A), where its activity is critically required for normal limb development in mouse (Sagai et al., 2005). Single-nucleotide mutations within the ZRS cause limb malformations, such as preaxial polydactyly, in multiple vertebrate species including humans (Hill and Lettice, 2013; Lettice et al., 2003, 2008; VanderMeer and Ahituv, 2011). Surprisingly, we observed that the sequence of this limb enhancer is conserved throughout nearly all examined species in the snake lineage. In basal snakes, which retain vestigial limbs, it is highly conserved, whereas it underwent a rapid increase in substitution rate in advanced snakes, in which all skeletal limb structures have disappeared. Consistent with this, we provide evidence that the snake enhancer progressively lost its *in vivo* function as the body plan evolved from basal to advanced snakes. Finally, we identify a specific subset of nucleotide changes within the enhancer that contribute to its functional degeneration in snakes and show in a mouse model that synthetic reintroduction of just one degraded transcription factor binding site is sufficient to recreate the ancestral function and to rescue normal limb formation *in vivo*.

## RESULTS

### A Critical Limb Enhancer Is Evolutionarily Conserved but Highly Diverged in Snakes

To explore the potential role of the ZRS limb enhancer in snake evolution, we examined the draft genomes of six snake species



**Figure 1. Evolution of a Limb Enhancer across the Vertebrate Tree**

(A) Human ZRS enhancer activity in a mid-gestation (E11.5) mouse embryo. Staining in structures other than limb was not reproducible in additional transgenic embryos and due to ectopic effects.

(B) Comparison of the core ZRS region across 18 different vertebrate species including two basal (blue) and four advanced (purple) snakes. See Figure S1 for full alignment.

(C) Phylogeny of vertebrate species used in the study (based on UCSC [https://genome.ucsc.edu/cgi-bin/hgGateway] and Hsiang et al., 2015; Pyron et al., 2013). Branch length indicates absolute ZRS substitution rate, colors indicate relative ZRS evolutionary rate compared to other embryonic enhancers (see Figure S2 and Method Details). The schematic illustrations of snake skeletons were drawn using images from Romanes (1892), <http://www.zoochat.com/>, and <http://www.skullcleaning.com/> as templates.

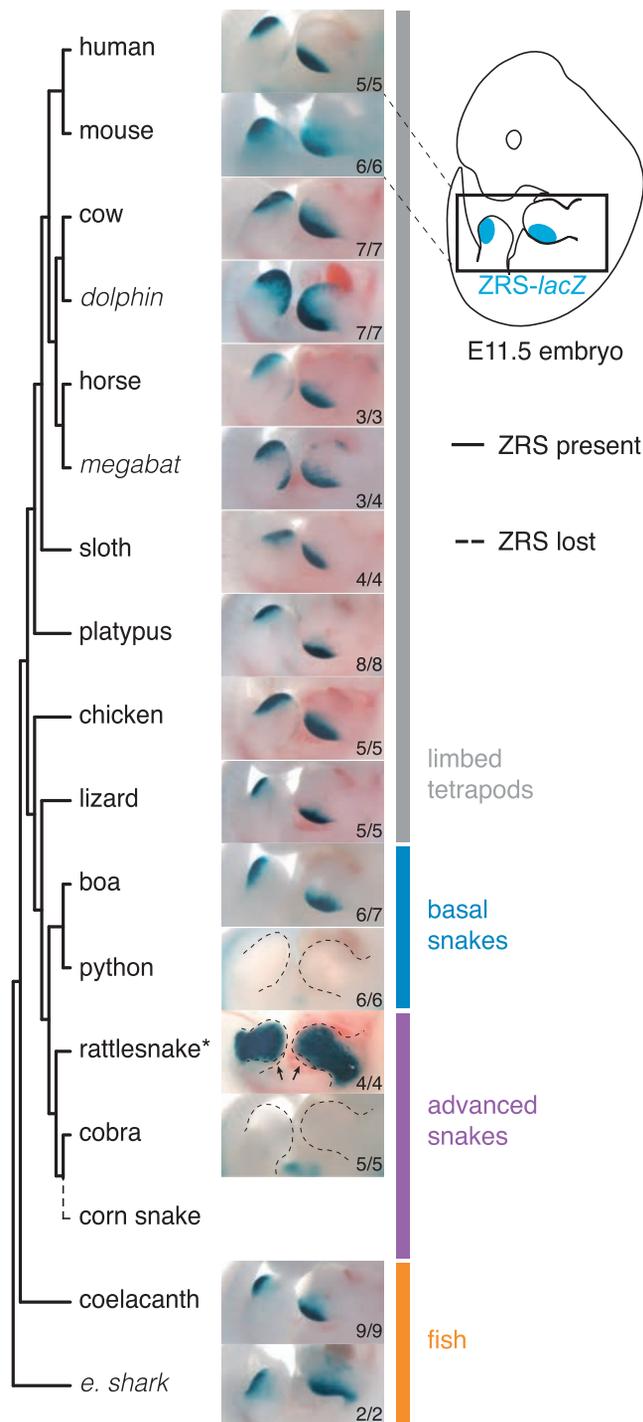
See also Figures S1 and S2.

including the Burmese python (*Python molurus bivittatus*) (Castoe et al., 2013), boa constrictor (*Boa constrictor constrictor*), king cobra (*Ophiophagus hannah*) (Vonk et al., 2013), speckled rattlesnake (*Crotalus mitchellii pyrrhus*), viper (*Vipera berus berus*), and corn snake (*Pantherophis guttatus*) (Ullate-Agote et al., 2014). These species represent different morphological stages within the evolutionary history of snakes (Apesteguía and Zaher, 2006; Martill et al., 2015), from basal snakes (boa and python) that retained a vestigial pelvic girdle and rudimentary hindlimbs, to advanced snakes (viper, rattlesnake, king cobra, and corn snake) that completely lost all skeletal limb structures and represent the majority (>85%) of all extant snake species (Lawson et al., 2005; Pyron et al., 2013). Nearly all of the snake species studied have a ZRS-orthologous sequence (Figures 1B and S1). However, while the ZRS enhancer of basal snakes shares ~80% nucleotide identity with the orthologous region from limbed lizards and shows a substitution rate similar to other vertebrate ZRS orthologs, the ZRS of advanced

snakes displays a substantially increased number of substitutions compared to other enhancers ( $p = 0.012$ , permutation test; Figures 1B, 1C, and S2; Table S4). This fast evolutionary rate clearly distinguishes the ZRS from other limb enhancers, which do not show such an increase in substitutions (Figure S2) (Infante et al., 2015). Thus, while nearly all snake species examined have a ZRS enhancer, a loss of evolutionary constraint on this enhancer coincides with the complete loss of limb structures at the transition from basal to advanced snakes.

### Loss of Region-Specific Limb Enhancer Activity in Snakes

To systematically examine whether the sequence changes observed in different snake ZRS orthologs alter the in vivo function of the enhancer, we used a transgenic mouse enhancer reporter assay (<http://enhancer.lbl.gov/>) (Kothary et al., 1989; Visel et al., 2007). We determined ZRS enhancer activity patterns for 16 different species covering a wide range of jawed vertebrates,



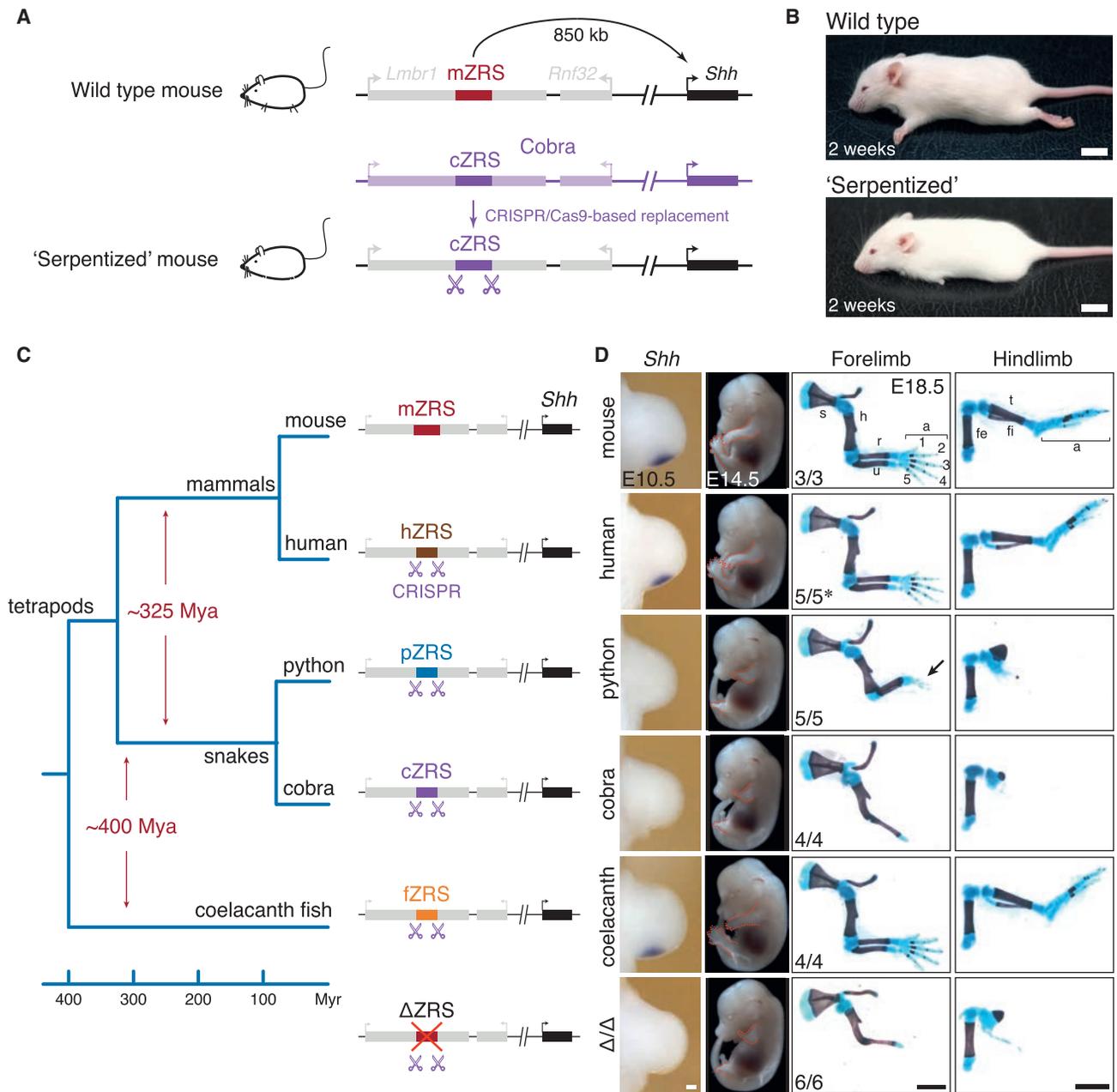
**Figure 2. Comparison of Enhancer Activity across Jawed Vertebrates**

Enhancer activities for 16 different vertebrate species in the limb buds of transgenic E11.5 stage mouse embryos. Numbers of embryos with lacZ activity in the limb over the total number of transgenic embryos screened are indicated. Some species (marked in italics) were active in the ZPA of the limb buds but had additional activity expanded anteriorly (*dolphin* and *megabat*) or proximally (*elephant shark*). \*The rattlesnake ZRS enhancer drives an ectopic reporter activity pattern that does not include the ZPA (arrows point to the ZPA area without detectable LacZ activity).

including cartilaginous and bony fishes (*elephant shark* and *coelacanth*), four snakes (*boa*, *python*, *rattlesnake*, and *cobra*), and ten limbed tetrapods at mid-gestation (embryonic day [E] 11.5), a time point when the mouse ZRS is active (Figures 1A and 2) (Lettice et al., 2003). The orthologs from nine finned or limbed vertebrates (*coelacanth*, *lizard*, *chicken*, *platypus*, *sloth*, *horse*, *cow*, *mouse*, and *human*) displayed reproducible patterns of activity in the posterior limb bud that were indistinguishable from the activity of the mouse enhancer (Figure 2), confirming the deep conservation of its function across vertebrates with paired appendages (Dahn et al., 2007; Lettice et al., 2003; Sagai et al., 2004). ZRS orthologs from three species were active in the ZPA of mouse limb buds but also had activity expanded anteriorly (*dolphin* and *megabat*) or proximally (*elephant shark*). In contrast, in four out of five basal and advanced snake species examined, either the enhancer activity in the ZPA or the enhancer sequence itself was lost (Figure 2). Among them, the rattlesnake ZRS displayed an ectopic limb activity pattern that did not include the ZPA and may be related to an ~180-bp insertion specifically found in the viper and rattlesnake lineage (Figures 2 and S1). Only the ZRS of *boa*, which diverged from *python* 63–96 million years ago (Esquerré and Scott Keogh, 2016) and among the examined snakes is the one showing the lowest nucleotide substitution rate with respect to that of the *lizard*, retained activity in the ZPA. Notably, the ZRS from all advanced snakes examined (*rattlesnake* and *cobra*) completely lost ZPA-specific activity.

### Snake Enhancer Knockin Causes Severe Limb Truncation in Mice

To assess the extent to which the observed activity changes in transgenic reporter assays affect vertebrate limb morphology in vivo, we employed CRISPR/Cas9 genome editing to generate a series of knockin (KI) mice where the functionally critical 1.3-kb core region of the ZRS (Figure S3) was replaced with the orthologous sequences of the same length from other species. We first replaced the mouse ZRS with the orthologs from *human* (73% sequence identity to the mouse ZRS) and *coelacanth* (57% sequence identity to the human ZRS), whose last common ancestor lived approximately 400 million years ago. Both the human and *coelacanth* orthologs resulted in *Shh* expression at the onset of limb bud formation that was indistinguishable from wild-type and rescued the formation of fully developed limbs (Figures 3 and S4G–S4J), indicating that despite considerable evolutionary distance between mammals and fish, the enhancers of mouse, human, and *coelacanth* are largely functionally interchangeable. In contrast, replacing the mouse ZRS with the orthologous *cobra* sequence resulted in a complete loss of *Shh* expression and a truncated limb phenotype, affecting both the fore- and hindlimbs, that is indistinguishable from the phenotype caused by deletion of the mouse enhancer (Figures 3, S3, and S4G) (Sagai et al., 2005). This result confirms that despite recognizable sequence conservation, the *cobra* sequence lacks limb enhancer function and is therefore unable to support limb development. The less diverged *python* ZRS resulted in a similar but a slightly milder phenotype. While most skeletal forelimb and hindlimb elements distal of the stylopod:zeugopod junction were also severely affected, the *python* ZRS resulted in formation of



**Figure 3. Limb Phenotypes of Knockin Mice with ZRS Orthologs from Other Vertebrate Species**

(A) CRISPR/Cas9-mediated replacement of the mouse ZRS sequence with an orthologous sequence from cobra. Schematic of the mouse *Shh* locus is shown at the top. The ZRS is located in the intron of the *Lmbr1* gene (intron-exon structure not shown), 850 kb away from the promoter of *Shh*. A homologous locus from king cobra with the cobra ZRS enhancer (cZRS) is indicated in purple. A CRISPR/Cas9-modified “serpentized” mouse *Shh* locus is shown below. See also Figures S4A–S4F and Method Details. Gene diagram not to scale.

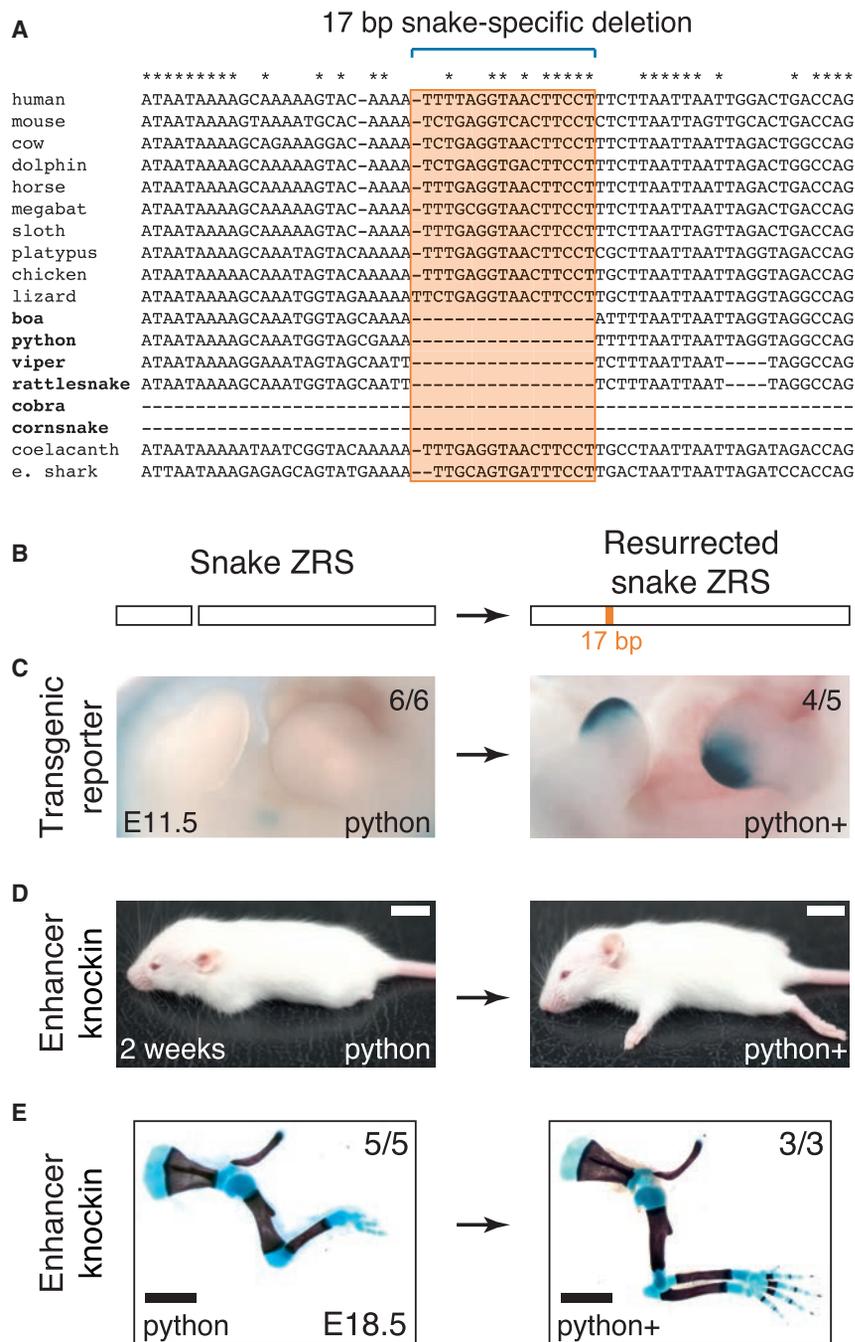
(B) Gross phenotypes of ZRS<sup>WT/Δ</sup> (top) and serpentized ZRS<sup>cZRS/Δ</sup> (bottom) mice. Scale bars, 10 mm.

(C and D) Limb phenotypes of knockin mice with ZRS orthologs from other vertebrate species.

(C) Phylogeny and approximate divergence estimates (Amemiya et al., 2013; Hsiang et al., 2015; Wright et al., 2015) are shown on the left. Schematic mouse *Shh* loci with the ZRS replaced by orthologs from human (hZRS), python (pZRS), cobra (cZRS), and coelacanth fish (fZRS) are shown.

(D) Comparative *Shh* mRNA in situ hybridization analysis in knockin mouse embryos during forelimb bud development (first column). Per knockin line, the *Shh* transcript distribution was assessed in at least three independent mouse embryos. See Figure S4G for hindlimb bud analysis of *Shh* expression. Corresponding whole-mount E14.5 knockin mouse embryos (second column) and skeletal preparations at E18.5 (third and fourth columns) are shown; s, scapula; h, humerus; r, radius; u, ulna; fe, femur; fi, fibula; t, tibia; a, autopod. The genotypes of the embryos are ZRS<sup>WT/Δ</sup> (mouse), ZRS<sup>hZRS/Δ</sup> (human), ZRS<sup>pZRS/Δ</sup> (python), ZRS<sup>cZRS/Δ</sup> (cobra), and ZRS<sup>fZRS/Δ</sup> (coelacanth fish). Arrow points to rudimentary digits in ZRS<sup>pZRS/Δ</sup> embryos. Bottom embryo shows E14.5

(legend continued on next page)



### Figure 4. Resurrection of Snake Limb Enhancer Function In Vivo

(A) Snake-specific deletion in the ZRS. An alignment of the central ZRS region for 18 vertebrates, including six snakes, is shown. Asterisks indicate nucleotides that are conserved in limbed tetrapods and fish.

(B) A 17-bp sequence is able to resurrect python ZRS enhancer function.

(C) Shown are the wild-type (left) and modified (right) python ZRS in vivo enhancer activities in the limb buds of transgenic E11.5 mouse embryos. Numbers of embryos with lacZ activity in the limb over the total number of transgenic embryos screened are indicated.

(D) The resurrected allele is able to rescue limb development when knocked into the mouse genome in place of the wild-type ZRS. Shown are gross phenotypes of ZRS<sup>ZRS/Δ</sup> (python, left) and ZRS<sup>ZRS(py)/Δ</sup> (python+, right) mice at 2 weeks of age. Scale bars, 10 mm.

(E) Skeletal preparations from E18.5 knockin mice are shown. See Figures S5B and S5C for more detailed skeletal phenotypes. Scale bars, 2 mm. See also Figure S5.

in situ hybridization indeed revealed very weak levels of *Shh* transcript in the posterior forelimb bud of python ZRS knockin mouse embryos (data not shown). Taken together our data indicate that both snake enhancers tested lost their ability to induce normal limb development in mice despite the much shorter evolutionary distance between mammals and snakes than between mammals and lobe-finned fish.

### In Vivo Resurrection of a Distant-Acting Snake Limb Enhancer

To identify specific nucleotide changes within the enhancer that may have led to its loss of activity in snakes, we examined the snake sequences in detail. While multiple nucleotide differences are observed between snakes and limbed lizards (Figure S1), one small deletion of 17 bp stood out because it affected a region of the ZRS that was highly conserved across all examined tetrapods and fish (Figure 4A).

two to three rudimentary digits in the forelimb and a slightly enlarged ossification resembling a rudimentary zeugopod (Figure 3D). This result may be due to residual enhancer activity that was not detected in transgenic reporter assays (Figure 2). Consistent with this possibility, prolonged staining after RNA

Although it represents less than 10% of all sequence changes between the snake and lizard ZRS, this deletion is the only sequence that is deleted in all snakes but present in all examined limbed vertebrates and fish (Figures 4A and S1). To directly test whether this small snake-specific deletion contributed to the loss

gross and limb skeletal phenotypes of the ZRS<sup>Δ/Δ</sup> KO mice (see Figure S3 for details). Numbers of embryos that exhibited representative limb phenotype over the total number of embryos with the genotype are indicated. \*Three of five mouse embryos displayed mild digit number variation (see Figures S4H–S4J). Scale bars, 0.1 mm (left column), 2 mm (columns 3 and 4). See also Figures S3 and S4.



increase in substitution rate in snakes (Figure 5D). Taken together, our results implicate the loss of the E1 ETS site as well as potentially other ETS and homeodomain transcription factor binding sites in the loss of function of this limb enhancer in snakes.

## DISCUSSION

In the present study, we demonstrate an increased rate of sequence changes, as well as progressive *in vivo* loss of function for a distant-acting limb enhancer in snakes. Decreased sequence conservation and loss of enhancer function were most pronounced in advanced snakes, which have lost all skeletal limb structures. The only snake genome in which no ZRS sequence was detected belonged to the corn snake. Our results indicate that the previously reported loss of the ZRS enhancer in Japanese rat snakes (Sagai et al., 2004), a member of the same subfamily (*Colubrinae*) as corn snakes, is not representative of snakes in general but affects only a small subset of advanced snakes where it occurred after the morphological loss of all limb structures (Figures 1B and S1). Across the snake species examined, the progressive sequence degeneration of the enhancer correlated with its loss of activity in transgenic reporter assays. In contrast, across all limbed tetrapods and fish examined, the enhancer activity was highly conserved. Remarkably, even a ZRS ortholog from fish (coelacanth), which shares less sequence similarity with the human ortholog than with the python ortholog (57% versus 59%), was sufficient for normal limb development despite the major morphological differences between mammalian limbs and coelacanth fins.

The molecular basis of loss of limbs in snakes as they evolved from their limbed ancestor has been the subject of extensive speculation (Apesteguía and Zaher, 2006; Cohn and Tickle, 1999; Di-Poï et al., 2010; Infante et al., 2015; Lopez-Rios, 2016; Martill et al., 2015; Sagai et al., 2004; Tchernov et al., 2000; Zeller et al., 2009). Our genomic enhancer replacement experiments in mice conclusively demonstrate that the loss of function in a single enhancer observed in snakes is sufficient to cause severe limb reduction in mice, raising the possibility that ZRS deactivation contributed to the loss of limbs in the snake lineage. However, changes in other sequences involved in limb development must also have occurred in snakes. These changes could for example involve regulation of *Hox* genes that act upstream of *Shh* (Cohn and Tickle, 1999; Di-Poï et al., 2010; Head and Polly, 2015), or other genes that are critical for initiation of limb development (e.g., Min et al., 1998; Rallis et al., 2003; Sekine et al., 1999; Tanaka et al., 2002). Notably, following the morphological disappearance of limbs, any sequence required exclusively for limb development is no longer subject to negative selection and is expected to degrade over time. This is exemplified by the reduction in the transgenic reporter activity of other serpentine limb enhancers whose phenotypic impact on limb development remains to be determined (Guerreiro et al., 2016; Infante et al., 2015). In the case of the ZRS, the enhancer activity observed in a basal snake (boa, Figure 2) suggests that the sequence degeneration of the ZRS in snakes started in conjunction with or, more likely, after other disruptive molecular events contributing to the loss of limbs. Consequently, we do not expect that the reintroduction

of a fully functional ZRS into a snake genome alone would be sufficient to induce the formation of fully or even partially developed limbs in snakes.

While we deliberately focused on a locus with strong pre-existing evidence for function from human disease and mouse genetics studies (reviewed in Hill and Lettice, 2013; VanderMeer and Ahituv, 2011), an increasing number of unbiased genome-wide enhancer data across closely and distantly related animal species (Acemel et al., 2016; Arnold et al., 2014; Cotney et al., 2013; Eckalbar et al., 2016; Gehrke et al., 2015; He et al., 2011; Prescott et al., 2015; Reilly et al., 2015; Villar et al., 2015; Xiao et al., 2012) creates a rapidly growing list of candidate lineage- and species-specific enhancers. A major challenge is the identification of the subsets of these enhancers that functionally contribute to morphological and other phenotypic diversity. Our study provides an example how genome editing-enabled enhancer replacement makes it possible to recapitulate the functional erosion of a regulatory sequence across evolution through *in vivo* experiments. As genome-editing tools are becoming increasingly available, we expect that this approach will be useful to routinely study the phenotypes associated with evolutionary changes in other regulatory sequences associated with morphological adaptations in vertebrates.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Phylogenetic Analysis
  - *In Vivo* Transgenic Reporter Assays
  - Generation of Enhancer Knockout and Knockin Mice Using CRISPR/Cas9
  - *In Situ* Hybridization
  - Skeletal Preparations
  - Sample Selection and Blinding
  - Motif Analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Substitution Rates in TF Motifs
  - Differences in Evolutionary Rates

## SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and six tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.09.028>.

An audio PaperClip is available at <http://dx.doi.org/10.1016/j.cell.2016.09.028#mmc2>.

## AUTHOR CONTRIBUTIONS

E.Z.K., D.E.D., E.M.R., A.V., and L.A.P. conceived the project. E.Z.K. and O.K. performed the phylogenetic analysis. E.Z.K. and U.S.M. cloned transgenic reporter and targeting vectors. E.Z.K., B.J.M., I.P.-F., C.S.P., T.H.G., M.K., E.A.L., J.A.A., and V.A. carried out transgenic validation. E.Z.K. performed the enhancer knockout and knockin studies. V.T., J.L.-R., M.O., and E.Z.K. performed *in situ* hybridization (ISH). I.B. and E.Z.K. performed motif analysis.

E.Z.K., A.V., and L.A.P. wrote the manuscript with input from the remaining authors.

## ACKNOWLEDGMENTS

This work was supported by NIH grants R01HG003988, U54HG006997, and U01DE024427 (to A.V. and L.A.P.) and by the Olga Mayenfisch Foundation and the University of Basel (to J.L.-R.). E.Z.K. is supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation funded by a Howard Hughes Medical Institute. M.O. was supported by a Swiss National Science Foundation (SNSF) fellowship. We thank T. Castoe for providing the draft assembly sequence of the Boa genome and J. Doudna for providing a plasmid containing a human-optimized Cas9 gene. We also thank A. Stark for helpful suggestions and comments on the manuscript. Research was conducted at the E.O. Lawrence Berkeley National Laboratory and performed under Department of Energy Contract DE-AC02-05CH11231, University of California.

Received: June 19, 2016

Revised: August 7, 2016

Accepted: September 15, 2016

Published: October 20, 2016

## REFERENCES

- Acemel, R.D., Tena, J.J., Irastorza-Azcarate, I., Marlétaz, F., Gómez-Marín, C., de la Calle-Mustienes, E., Bertrand, S., Diaz, S.G., Aldea, D., Aury, J.M., et al. (2016). A single three-dimensional chromatin compartment in amphioxus indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nat. Genet.* **48**, 336–341.
- Amemiya, C.T., Alföldi, J., Lee, A.P., Fan, S., Philippe, H., Maccallum, I., Braasch, I., Manousaki, T., Schneider, I., Rohner, N., et al. (2013). The African coelacanth genome provides insights into tetrapod evolution. *Nature* **496**, 311–316.
- Apestequía, S., and Zaher, H. (2006). A Cretaceous terrestrial snake with robust hindlimbs and a sacrum. *Nature* **440**, 1037–1040.
- Arnold, C.D., Gerlach, D., Spies, D., Matts, J.A., Sytnikova, Y.A., Pagani, M., Lau, N.C., and Stark, A. (2014). Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat. Genet.* **46**, 685–692.
- Attanasio, C., Nord, A.S., Zhu, Y., Blow, M.J., Li, Z., Liberton, D.K., Morrison, H., Plajzer-Frick, I., Holt, A., Hosseini, R., et al. (2013). Fine tuning of craniofacial morphology by distant-acting enhancers. *Science* **342**, 1241006.
- Blow, M.J., McCulley, D.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2010). ChIP-Seq identification of weakly conserved heart enhancers. *Nat. Genet.* **42**, 806–810.
- Britten, R.J., and Davidson, E.H. (1969). Gene regulation for higher cells: A theory. *Science* **165**, 349–357.
- Bronson, S.K., Plaehn, E.G., Kluckman, K.D., Hagaman, J.R., Maeda, N., and Smithies, O. (1996). Single-copy transgenic mice with chosen-site integration. *Proc. Natl. Acad. Sci. USA* **93**, 9067–9072.
- Capellini, T.D., Di Giacomo, G., Salsi, V., Brendolan, A., Ferretti, E., Srivastava, D., Zappavigna, V., and Selleri, L. (2006). Pbx1/Pbx2 requirement for distal limb patterning is mediated by the hierarchical control of Hox gene spatial distribution and Shh expression. *Development* **133**, 2263–2273.
- Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* **134**, 25–36.
- Castoe, T.A., de Koning, A.P.J., Hall, K.T., Card, D.C., Schield, D.R., Fujita, M.K., Ruggiero, R.P., Degner, J.F., Daza, J.M., Gu, W., et al. (2013). The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. *Proc. Natl. Acad. Sci. USA* **110**, 20645–20650.
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552.
- Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Jr., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* **327**, 302–305.
- Cohn, M.J., and Tickle, C. (1999). Developmental basis of limblessness and axial patterning in snakes. *Nature* **399**, 474–479.
- Cooper, K.L., Sears, K.E., Uygur, A., Maier, J., Baczkowski, K.-S., Brosnahan, M., Antczak, D., Skidmore, J.A., and Tabin, C.J. (2014). Patterning and post-patterning modes of evolutionary digit loss in mammals. *Nature* **511**, 41–45.
- Cotney, J., Leng, J., Yin, J., Reilly, S.K., DeMare, L.E., Emera, D., Ayoub, A.E., Rakic, P., and Noonan, J.P. (2013). The evolution of lineage-specific regulatory activities in the human embryonic limb. *Cell* **154**, 185–196.
- Cretekos, C.J., Wang, Y., Green, E.D., Martin, J.F., Rasweiler, J.J., 4th, and Behringer, R.R. (2008). Regulatory divergence modifies limb length between mammals. *Genes Dev.* **22**, 141–151.
- Dahn, R.D., Davis, M.C., Pappano, W.N., and Shubin, N.H. (2007). Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning. *Nature* **445**, 311–314.
- Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **9**, 772.
- Di-Poi, N., Montoya-Burgos, J.I., Miller, H., Pourquié, O., Milinkovitch, M.C., and Duboule, D. (2010). Changes in Hox genes' structure and function during the evolution of the squamate body plan. *Nature* **464**, 99–103.
- Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J.A., and McMahon, A.P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417–1430.
- Eckalbar, W.L., Schlebusch, S.A., Mason, M.K., Gill, Z., Parker, A.V., Booker, B.M., Nishizaki, S., Muswamba-Nday, C., Terhune, E., Nevenon, K.A., et al. (2016). Transcriptomic and epigenomic characterization of the developing bat wing. *Nat. Genet.* **48**, 528–536.
- Esquerré, D., and Scott Keogh, J. (2016). Parallel selective pressures drive convergent diversification of phenotypes in pythons and boas. *Ecol. Lett.* **19**, 800–809.
- Gehrke, A.R., Schneider, I., de la Calle-Mustienes, E., Tena, J.J., Gomez-Marín, C., Chandran, M., Nakamura, T., Braasch, I., Postlethwait, J.H., Gómez-Skarmeta, J.L., et al. (2015). Deep conservation of wrist and digit enhancers in fish. *Proc. Natl. Acad. Sci. USA* **112**, 803–808.
- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345.
- Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018.
- Guenther, C.A., Tasic, B., Luo, L., Bedell, M.A., and Kingsley, D.M. (2014). A molecular basis for classic blond hair color in Europeans. *Nat. Genet.* **46**, 748–752.
- Guerreiro, I., Nunes, A., Woltering, J.M., Casaca, A., Nóvoa, A., Vinagre, T., Hunter, M.E., Duboule, D., and Mallo, M. (2013). Role of a polymorphism in a Hox/Pax-responsive enhancer in the evolution of the vertebrate spine. *Proc. Natl. Acad. Sci. USA* **110**, 10682–10686.
- Guerreiro, I., Gitto, S., Nóvoa, A., Codourey, J., Nguyen Huynh, T.H., Gonzalez, F., Milinkovitch, M.C., Mallo, M., and Duboule, D. (2016). Reorganisation of Hoxd regulatory landscapes during the evolution of a snake-like body plan. *eLife* **5**, 347.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321.
- He, Q., Bardet, A.F., Patton, B., Purvis, J., Johnston, J., Paulson, A., Gogol, M., Stark, A., and Zeitlinger, J. (2011). High conservation of transcription factor binding and evidence for combinatorial regulation across six *Drosophila* species. *Nat. Genet.* **43**, 414–420.
- Head, J.J., and Polly, P.D. (2015). Evolution of the snake body form reveals homoplasy in amniote Hox gene function. *Nature* **520**, 86–89.

- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589.
- Hill, R.E., and Lettice, L.A. (2013). Alterations to the remote control of Shh gene expression cause congenital abnormalities. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120357.
- Hoekstra, H.E. (2012). Genomics: Stickleback is the catch of the day. *Nature* **484**, 46–47.
- Hsiang, A.Y., Field, D.J., Webster, T.H., Behlke, A.D.B., Davis, M.B., Racicot, R.A., and Gauthier, J.A. (2015). The origin of snakes: Revealing the ecology, behavior, and evolutionary history of early snakes using genomics, phenomics, and the fossil record. *BMC Evol. Biol.* **15**, 87.
- Indjeian, V.B., Kingman, G.A., Jones, F.C., Guenther, C.A., Grimwood, J., Schmutz, J., Myers, R.M., and Kingsley, D.M. (2016). Evolving new skeletal traits by cis-regulatory changes in bone morphogenetic proteins. *Cell* **164**, 45–56.
- Infante, C.R., Mihala, A.G., Park, S., Wang, J.S., Johnson, K.K., Lauderdale, J.D., and Menke, D.B. (2015). Shared enhancer activity in the limbs and phallus and functional divergence of a limb-genital cis-regulatory element in snakes. *Dev. Cell* **35**, 107–119.
- Jolma, A., Yan, J., Whittington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. *Cell* **152**, 327–339.
- Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., et al.; Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61.
- Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780.
- King, M.C., and Wilson, A.C. (1975). Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116.
- Kmita, M., Turchini, B., Zákány, J., Logan, M., Tabin, C.J., and Duboule, D. (2005). Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature* **435**, 1113–1116.
- Kothary, R., Clapoff, S., Darling, S., Perry, M.D., Moran, L.A., and Rossant, J. (1989). Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Development* **105**, 707–714.
- Lawson, R., Slowinski, J.B., Crother, B.I., and Burbrink, F.T. (2005). Phylogeny of the Colubroidea (Serpentes): New evidence from mitochondrial and nuclear genes. *Mol. Phylogenet. Evol.* **37**, 581–601.
- Lettice, L.A., Heaney, S.J.H., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735.
- Lettice, L.A., Hill, A.E., Devenney, P.S., and Hill, R.E. (2008). Point mutations in a distant sonic hedgehog cis-regulator generate a variable regulatory output responsible for preaxial polydactyly. *Hum. Mol. Genet.* **17**, 978–985.
- Lettice, L.A., Williamson, I., Wiltshire, J.H., Peluso, S., Devenney, P.S., Hill, A.E., Essafi, A., Hagman, J., Mort, R., Grimes, G., et al. (2012). Opposing functions of the ETS factor family define Shh spatial expression in limb buds and underlie polydactyly. *Dev. Cell* **22**, 459–467.
- Lettice, L.A., Williamson, I., Devenney, P.S., Kilanowski, F., Dorin, J., and Hill, R.E. (2014). Development of five digits is controlled by a bipartite long-range cis-regulator. *Development* **141**, 1715–1725.
- Lopez-Rios, J. (2016). The many lives of SHH in limb development and evolution. *Semin. Cell Dev. Biol.* **49**, 116–124.
- Lopez-Rios, J., Duchesne, A., Speziale, D., Andrey, G., Peterson, K.A., Germann, P., Únal, E., Liu, J., Floriot, S., Barbey, S., et al. (2014). Attenuated sensing of SHH by Ptch1 underlies evolution of bovine limbs. *Nature* **511**, 46–51.
- Martill, D.M., Tischlinger, H., and Longrich, N.R. (2015). EVOLUTION. A four-legged snake from the Early Cretaceous of Gondwana. *Science* **349**, 416–419.
- May, D., Blow, M.J., Kaplan, T., McCulley, D.J., Jensen, B.C., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., et al. (2011). Large-scale discovery of enhancers from human heart tissue. *Nat. Genet.* **44**, 89–93.
- McLean, C.Y., Reno, P.L., Pollen, A.A., Bassan, A.I., Capellini, T.D., Guenther, C., Indjeian, V.B., Lim, X., Menke, D.B., Schaar, B.T., et al. (2011). Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* **471**, 216–219.
- Min, H., Danilenko, D.M., Scully, S.A., Bolon, B., Ring, B.D., Tarpley, J.E., DeRose, M., and Simonet, W.S. (1998). Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to Drosophila branchless. *Genes Dev.* **12**, 3156–3161.
- Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E. (2014). CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* **42**, W401–W407.
- Ovchinnikov, D. (2009). Alcian blue/alizarin red staining of cartilage and bone in mouse. *Cold Spring Harb. Protoc.* Published online March 2009. <http://dx.doi.org/10.1101/pdb.prot5170>.
- Panman, L., Galli, A., Lagarde, N., Michos, O., Soete, G., Zuniga, A., and Zeller, R. (2006). Differential regulation of gene expression in the digit forming area of the mouse limb bud by SHH and gremlin 1/FGF-mediated epithelial-mesenchymal signalling. *Development* **133**, 3419–3428.
- Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D., et al. (2006). In vivo enhancer analysis of human conserved non-coding sequences. *Nature* **444**, 499–502.
- Prabhakar, S., Visel, A., Akiyama, J.A., Shoukry, M., Lewis, K.D., Holt, A., Plajzer-Frick, I., Morrison, H., Fitzpatrick, D.R., Afzal, V., et al. (2008). Human-specific gain of function in a developmental enhancer. *Science* **321**, 1346–1350.
- Prescott, S.L., Srinivasan, R., Marchetto, M.C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F.H., Swigut, T., and Wysocka, J. (2015). Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. *Cell* **163**, 68–83.
- Pyron, R.A., Burbrink, F.T., and Wiens, J.J. (2013). A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evol. Biol.* **13**, 93.
- Rallis, C., Bruneau, B.G., Del Buono, J., Seidman, C.E., Seidman, J.G., Nissim, S., Tabin, C.J., and Logan, M.P.O. (2003). Tbx5 is required for forelimb bud formation and continued outgrowth. *Development* **130**, 2741–2751.
- Reilly, S.K., Yin, J., Ayoub, A.E., Emera, D., Leng, J., Cotney, J., Sarro, R., Rakic, P., and Noonan, J.P. (2015). Evolutionary genomics. Evolutionary changes in promoter and enhancer activity during human corticogenesis. *Science* **347**, 1155–1159.
- Romanes, G.J. (1892). Darwin and after Darwin: An Exposition of the Darwinian Theory and a Discussion of post-Darwinian Questions (The Open Court Publishing Company).
- Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T., and Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). *Mamm. Genome* **15**, 23–34.
- Sagai, T., Hosoya, M., Mizushima, Y., Tamura, M., and Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. *Development* **132**, 797–803.
- Sekine, K., Ohuchi, H., Fujiwara, M., Yamasaki, M., Yoshizawa, T., Sato, T., Yagishita, N., Matsui, D., Koga, Y., Itoh, N., and Kato, S. (1999). Fgf10 is essential for limb and lung formation. *Nat. Genet.* **21**, 138–141.
- Shimodaira, H., and Hasegawa, M. (2001). CONSEL: For assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246–1247.
- Tanaka, M., Münsterberg, A., Anderson, W.G., Prescott, A.R., Hazon, N., and Tickle, C. (2002). Fin development in a cartilaginous fish and the origin of vertebrate limbs. *Nature* **416**, 527–531.

- Tchernov, E., Rieppel, O., Zaher, H., Polcyn, M.J., and Jacobs, L.L. (2000). A fossil snake with limbs. *Science* 287, 2010–2012.
- Ullate-Agote, A., Milinkovitch, M.C., and Tzika, A.C. (2014). The genome sequence of the corn snake (*Pantherophis guttatus*), a valuable resource for EvoDevo studies in squamates. *Int. J. Dev. Biol.* 58, 881–888.
- VanderMeer, J.E., and Ahituv, N. (2011). cis-regulatory mutations are a genetic cause of human limb malformations. *Dev. Dyn.* 240, 920–930.
- Villar, D., Berthelot, C., Aldridge, S., Rayner, T.F., Lukk, M., Pignatelli, M., Park, T.J., Deaville, R., Erichsen, J.T., Jasinska, A.J., et al. (2015). Enhancer evolution across 20 mammalian species. *Cell* 160, 554–566.
- Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser—A database of tissue-specific human enhancers. *Nucleic Acids Res.* 35, D88–D92.
- Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854–858.
- Vonk, F.J., Casewell, N.R., Henkel, C.V., Heimberg, A.M., Jansen, H.J., McCleary, R.J.R., Kerckamp, H.M.E., Vos, R.A., Guerreiro, I., Calvete, J.J., et al. (2013). The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc. Natl. Acad. Sci. USA* 110, 20651–20656.
- Wang, H., Yang, H., Shivalilla, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918.
- Wittkopp, P.J., and Kalay, G. (2011). Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13, 59–69.
- Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* 8, 206–216.
- Wright, A.M., Lyons, K.M., Brandley, M.C., and Hillis, D.M. (2015). Which came first: The lizard or the egg? Robustness in phylogenetic reconstruction of ancestral states. *J. Exp. Zool. B Mol. Dev. Evol.* 324, 504–516.
- Xiao, S., Xie, D., Cao, X., Yu, P., Xing, X., Chen, C.-C., Musselman, M., Xie, M., West, F.D., Lewin, H.A., et al. (2012). Comparative epigenomic annotation of regulatory DNA. *Cell* 149, 1381–1392.
- Yang, H., Wang, H., Shivalilla, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370–1379.
- Yang, H., Wang, H., and Jaenisch, R. (2014). Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protoc.* 9, 1956–1968.
- Zeller, R., and Zuniga, A. (2007). Shh and Gremlin1 chromosomal landscapes in development and disease. *Curr. Opin. Genet. Dev.* 17, 428–434.
- Zeller, R., López-Ríos, J., and Zuniga, A. (2009). Vertebrate limb bud development: Moving towards integrative analysis of organogenesis. *Nat. Rev. Genet.* 10, 845–858.

# Domestication and Divergence of *Saccharomyces cerevisiae* Beer Yeasts

Brigida Gallone,<sup>1,2,3,4,11</sup> Jan Steensels,<sup>1,2,11</sup> Troels Prah,<sup>5</sup> Leah Soriaga,<sup>6</sup> Veerle Saels,<sup>1,2</sup> Beatriz Herrera-Malaver,<sup>1,2</sup> Adriaan Merlevede,<sup>1,2</sup> Miguel Roncoroni,<sup>1,2</sup> Karin Voordeckers,<sup>1,2</sup> Loren Miraglia,<sup>8</sup> Clotilde Teiling,<sup>9</sup> Brian Steffy,<sup>9</sup> Maryann Taylor,<sup>10</sup> Ariel Schwartz,<sup>6</sup> Toby Richardson,<sup>6</sup> Christopher White,<sup>5</sup> Guy Baele,<sup>7</sup> Steven Maere,<sup>3,4,\*</sup> and Kevin J. Verstrepen<sup>1,2,12,\*</sup>

<sup>1</sup>Laboratory for Genetics and Genomics, Centre of Microbial and Plant Genetics (CMPG), KU Leuven, Kasteelpark Arenberg 22, 3001 Leuven, Belgium

<sup>2</sup>Laboratory for Systems Biology, VIB, Bio-Incubator, Gaston Geenslaan 1, 3001 Leuven, Belgium

<sup>3</sup>Department of Plant Systems Biology, VIB, 9052 Gent, Belgium

<sup>4</sup>Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Gent, Belgium

<sup>5</sup>White Labs, 9495 Candida Street, San Diego, CA 92126, USA

<sup>6</sup>Synthetic Genomics, 11149 North Torrey Pines Road, La Jolla, CA 92037, USA

<sup>7</sup>Department of Microbiology and Immunology, Rega Institute, KU Leuven, 3000 Leuven, Belgium

<sup>8</sup>Encinitas Brewing Science, 141 Rodney Avenue, Encinitas, CA 92024, USA

<sup>9</sup>Illumina, 5200 Illumina Way, San Diego, CA 92122, USA

<sup>10</sup>Biological & Popular Culture (BioPop), 2205 Faraday Avenue, Suite E, Carlsbad, CA 92008, USA

<sup>11</sup>Co-first author

<sup>12</sup>Lead Contact

\*Correspondence: steven.maere@psb.vib-ugent.be (S.M.), kevin.verstrepen@biw.vib-kuleuven.be (K.J.V.)

<http://dx.doi.org/10.1016/j.cell.2016.08.020>

## SUMMARY

Whereas domestication of livestock, pets, and crops is well documented, it is still unclear to what extent microbes associated with the production of food have also undergone human selection and where the plethora of industrial strains originates from. Here, we present the genomes and phenomes of 157 industrial *Saccharomyces cerevisiae* yeasts. Our analyses reveal that today's industrial yeasts can be divided into five sublineages that are genetically and phenotypically separated from wild strains and originate from only a few ancestors through complex patterns of domestication and local divergence. Large-scale phenotyping and genome analysis further show strong industry-specific selection for stress tolerance, sugar utilization, and flavor production, while the sexual cycle and other phenotypes related to survival in nature show decay, particularly in beer yeasts. Together, these results shed light on the origins, evolutionary history, and phenotypic diversity of industrial yeasts and provide a resource for further selection of superior strains.

## INTRODUCTION

Since prehistoric times, humans have exploited the capacity of the common baker's yeast *Saccharomyces cerevisiae* to convert sugars into ethanol and desirable flavor compounds to obtain foods and beverages with a prolonged shelf-life, enriched sensorial palate, improved digestibility, and an euphoriant effect

due to the presence of ethanol (Michel et al., 1992; Steensels and Verstrepen, 2014). Whereas the use of pure cultures started well after the pioneering work of Pasteur and Hansen in the 19<sup>th</sup> century, early brewers, winemakers, and bakers had already learned that inoculating unfermented foods with a small portion of fermented product resulted in fast and more predictable fermentations. This so-called “backslopping” might have resulted in yeast lineages that grew continuously in these man-made environments and lost contact with their natural niches, providing a perfect setting for domestication. However, strong evidence for this hypothesis is still missing and it remains unclear whether industrial yeast diversity is shaped by selection and niche adaptation (domestication) or neutral divergence caused by geographic isolation and limited dispersal (Goddard and Greig, 2015; Warringer et al., 2011).

Domestication is defined as human selection and breeding of wild species to obtain cultivated variants that thrive in man-made environments, but behave suboptimally in nature. Typical signs of domestication, including genome decay, polyploidy, chromosomal rearrangements, gene duplications, and phenotypes resulting from human-driven selection, have been reported in crops, livestock, and pets (Driscoll et al., 2009; Purugganan and Fuller, 2009). Several studies have recently investigated the *S. cerevisiae* population by sequencing the genomes of hundreds of different strains, providing a first glimpse of the complex evolution of this species (Almeida et al., 2015; Borneman et al., 2011, 2016; Liti et al., 2009; Magwene et al., 2011; Schacherer et al., 2009; Strope et al., 2015). However, most of these studies focused primarily on yeasts from wild and clinical habitats and often include only a limited set of industrial strains, mainly originating from wine. Moreover, most studies use haploid derivatives instead of natural strains and can therefore not explore typical patterns of domestication like polyploidy, aneuploidy,

and heterozygosity. The use of haploids also excludes a large fraction of industrial *S. cerevisiae* strains that have lost the ability to sporulate, such as the vast majority of beer yeasts. Nevertheless, some studies already revealed signs of domestication in wine strains, such as an increased resistance to copper (present in grapevine pesticides) and sulfite (used as a preservative in wine) (Pérez-Ortín et al., 2002; Warringer et al., 2011). An in-depth investigation of strains originating from other industrial niches is still lacking.

Here, we describe the high-quality sequencing, de novo assembly, annotation, and extensive phenotyping of 157 *S. cerevisiae* strains used for the industrial production of beer, wine, bread, spirits, saké, and bioethanol, in their natural ploidy. Our data reveal that industrial yeasts are genetically and phenotypically distinct from wild strains and stem from only a limited set of ancestral strains that have been adapting to man-made environments. They further diversified into five clades: one including Asian strains such as saké yeasts, one mostly containing wine yeasts, a mixed clade that contains bread and other yeasts, and two separate families of beer yeasts. While most clades lack strong geographical substructure, one of the beer clades contains geographically isolated subgroups of strains used in continental Europe (Belgium/Germany), the United Kingdom, and a recent sublineage of United States beer yeasts that diverged from the British subclade during colonization. Interestingly, these beer yeast lineages exhibit clear and profound hallmarks of domestication, more so than the other lineages. The shift from variable, complex, and often harsh environments encountered in nature to more stable and nutrient-rich beer medium favored specialized adaptations in beer yeasts, but also led to genome decay, aneuploidy, and loss of a functional sexual cycle. Specifically, we find evidence for active human selection, demonstrated by convergent evolution for efficient fermentation of beer-specific carbon sources, mainly through mutations and duplications of the *MAL* (maltose) genes, as well as nonsense mutations in *PAD1* and *FDC1*, which are involved in the production of 4-vinyl guaiacol (4-VG), an undesirable off-flavor in beer. Our results further suggest that beer yeast domestication was initiated hundreds of years ago, well after the first reported beer production, but before the discovery of microbes. Together, our results reveal how today's industrial yeasts are the outcome of centuries of human domestication and provide a new resource for further selection and breeding of superior variants.

## RESULTS

### Niche and Geography Drive Diversification

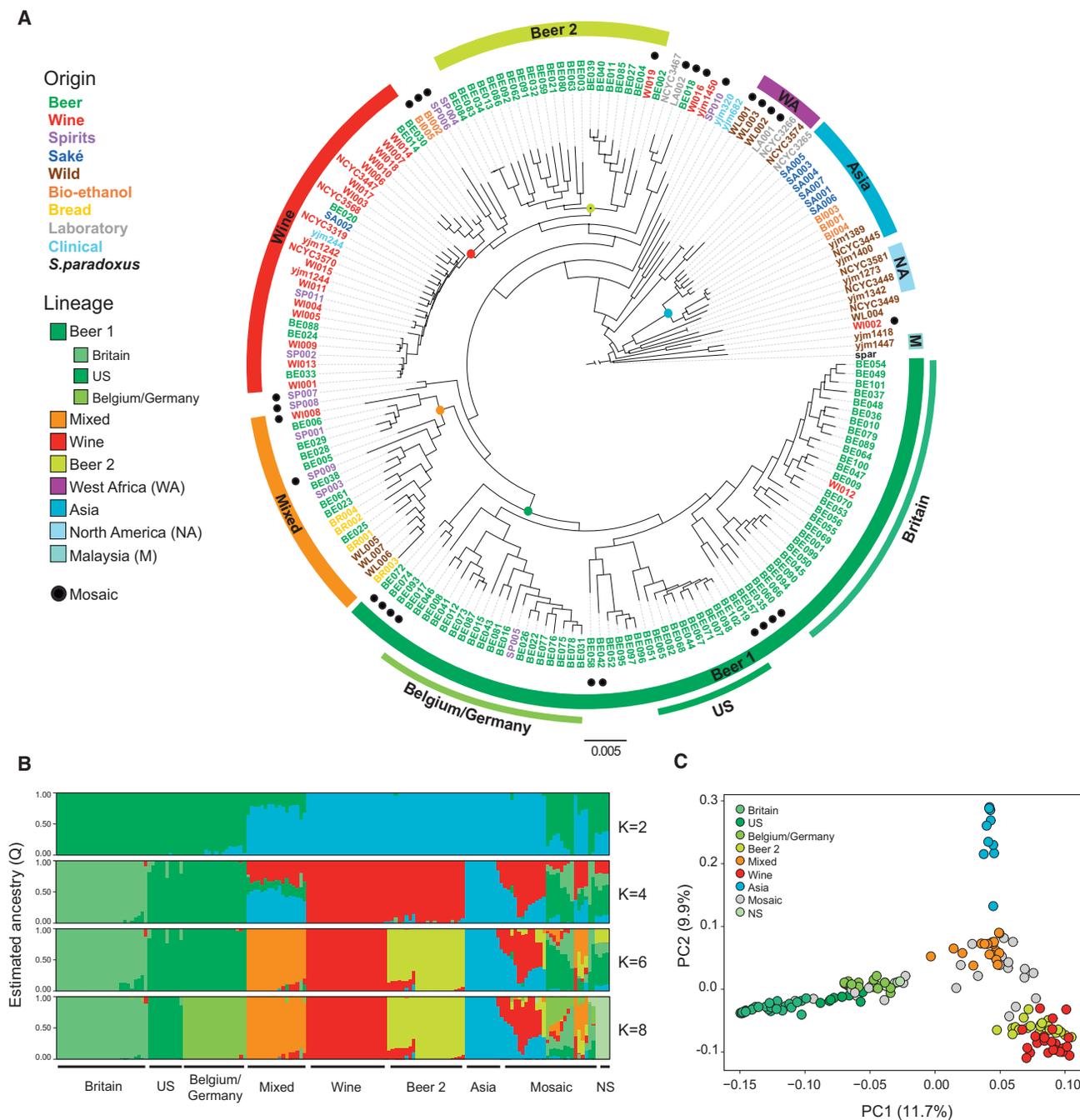
To examine the evolutionary history of industrial yeasts, we sequenced the genomes of 157 *S. cerevisiae* isolates originating from various sources in their natural ploidy to a median coverage of 135× (min = 26×, max = 403×) (for details on data analysis, see STAR Methods). This collection includes 102 industrial beer strains, 19 wine strains, 11 spirit strains, 7 saké strains, 7 strains isolated from spontaneous fermentations, 5 bioethanol strains, 4 bread strains, and 2 laboratory strains (Table S1). Interestingly, ten of these *S. cerevisiae* beer strains are used for commercial production of lager beers, which were believed to be

exclusively produced by strains of the genetically related *Saccharomyces pastorianus*. After de novo assembly of each of the genomes, we inferred a maximum-likelihood phylogenetic tree based on codon alignments for 2,020 concatenated single-copy nuclear genes shared by each of the 157 isolates and the outgroup species *Saccharomyces paradoxus* (Figure S1A). Additionally, we included a representative set of 24 previously sequenced strains belonging to the main established lineages of the *S. cerevisiae* phylogeny (Liti et al., 2009; Strobe et al., 2015), extending the number of strains to 181 (Figure 1A). Trees constructed from the original and extended datasets are congruent and show five main lineages that contain the majority of industrial yeasts: Wine (bootstrap support 100%), Beer 1 (86%), Beer 2 (56%), Asia (100%), and a Mixed lineage (99%) containing yeasts used in different industries. Three of these lineages (Beer 1, Beer 2, and Mixed) were not previously described.

Next, we studied the population structure in a filtered set of 53,929 polymorphic sites accounting for 2,454,052 SNPs across all strains, using the Bayesian model-based clustering approach implemented in fastStructure (Raj et al., 2014) (Figures 1B and S1B). This analysis yields a population structure that is highly consistent with the major lineages defined in the phylogeny and identifies mosaicism in 17% of the strains (in which the estimated ancestry  $Q < 0.8$  for  $K = 8$  ancestral populations). The population structure is further supported by a principal component analysis (PCA) on the same SNP data (Figure 1C).

Further analysis of the phylogeny and population structure reveals that the evolutionary divergence of industrial yeasts is shaped by both their industrial application and geographical origin. First, most yeasts cluster together according to the industry in which they are used and are clearly separated from the wild or clinical yeasts that have previously been sequenced. This was further confirmed by constructing a larger phylogeny, based on nine genomic regions, that includes the vast majority of all sequenced *S. cerevisiae* strains, 450 isolates in total (Figure S1C; Table S2). Wine and saké yeasts cluster in the previously identified Wine and Asia lineages (Liti et al., 2009). The majority of beer yeasts (85.3%) are found in two main lineages (Beer 1 and Beer 2) that are only distantly related. The Mixed clade harbors 7.8% of all beer strains (most of which are atypical beer yeasts that are used for bottle refermentation of strong Belgian ales) and contains all bread strains. Interestingly, spirit strains lack this clear phylogenetic relationship, as they are highly mosaic and scattered throughout the tree, suggesting that these strains might be the result of breeding by modern-day yeast companies that sell yeasts for spirits production. Moreover, because spirit yeasts are typically not re-used after fermentation, they likely had less opportunity to diverge into a separate clade.

Within and between the lineages, we also observed geographical patterns. For example, most saké yeasts form a monophyletic group and cluster together with wild isolates and bioethanol strains from China, while South American bioethanol strains are closely related to strains used to produce cachaça, a Brazilian sugarcane spirit. Moreover, the Beer 1 clade consists of three separate subpopulations, each reflecting geographically distinct groups: Belgium/Germany, Britain, and the United States. The absence of genetic admixture among these subpopulations indicates that these strains diverged allopatrically after the initial split



**Figure 1. Phylogeny and Population Structure of Industrial *S. cerevisiae* Strains**

(A) Maximum likelihood phylogenetic tree of all *S. cerevisiae* strains sequenced in this project supplemented with a representative set of 24 previously sequenced strains (Liti et al., 2009; Strobe et al., 2015) and using *Saccharomyces paradoxus* as an outgroup. Black dots on nodes indicate bootstrap support values <70%. Color codes indicate origin (names) and lineage (circular bands). The basal splits of the five industrial lineages are indicated with colored dots. Mosaic strains identified in this study are indicated with black dots next to the strain codes. Branch lengths reflect the average number of substitutions per site. Scale bar, 0.005 substitutions per site.

(B) Population structure identified in the 157 surveyed strains. The vertical axis depicts the fractional representation of resolved populations (colors) within each strain (horizontal axis, strains listed in Figure S1C) for K = 2, 4, 6, and 8 assumed ancestral populations (where K = 8 maximizes the marginal likelihood and best explains the data structure). Mosaic strains (i.e., strains that possess <80% ancestry from a single population) are visualized as a separate group.

(C) Principal component projection, using the same set of SNPs as in Figure 1B. Colors represent different populations. WA, West Africa; NA, North America; M, Malaysia; NS, not specified.

See also Figure S1 and Tables S1, S2, and S8.

**Table 1. Genetic Diversity within Each Subpopulation of Industrial *S. cerevisiae* Strains**

Subpopulation	Number of Strains	Analyzed Sites	Segregating Sites	$\pi$	$\Theta_w$
Britain	26	12,018,937	101,881	3.13E-03	1.88E-03
United States	10	11,973,239	72,559	2.31E-03	1.72E-03
Belgium/Germany	18	12,017,007	108,560	3.12E-03	2.19E-03
Mixed	17	12,043,532	132,188	4.35E-03	2.69E-03
Wine	24	12,052,956	114,133	1.59E-03	2.15E-03
Beer 2	21	12,063,361	142,745	2.95E-03	2.77E-03
Asia	10	12,035,745	99,879	2.39E-03	2.36E-03

The number of strains per subpopulation, the amount of analyzed and segregating sites, as well as nucleotide diversity ( $\pi$ ) and population mutation rate (Watterson's  $\theta$ ,  $\Theta_w$ ) are indicated.

(Figure 1B). Moreover, the high nucleotide diversity within each of the Beer 1 sublineages exceeds that within the Wine population, suggesting that the split did not happen recently (Table 1). Compared to Beer 1, Beer 2 is more closely related to the Wine lineage and includes 20.6% of all brewing strains. However, in contrast to the Beer 1 group, the Beer 2 lineage lacks geographic structure and contains yeasts originating from Belgium, the United Kingdom, the United States, Germany, and Eastern Europe. The presence of two major genetically distinct sources of beer yeasts hints toward two independent European domestication events, one of which is at the origin of both the Wine and Beer 2 clade.

### Remarkable Structural Variation in Beer Yeasts

Variation in genome structure, such as polyploidy, aneuploidy, large segmental duplications, and copy-number variations (CNVs), have repeatedly been found in association with domestication and adaptation to specific niches in experimentally evolved microbes (Bergström et al., 2014; Borneman et al., 2011; Dunham et al., 2002; Dunn et al., 2012; Pavelka et al., 2010; Rancati et al., 2008; Selmecki et al., 2009; Voordeckers et al., 2015) and in association with domestication of higher organisms (Purugganan and Fuller, 2009).

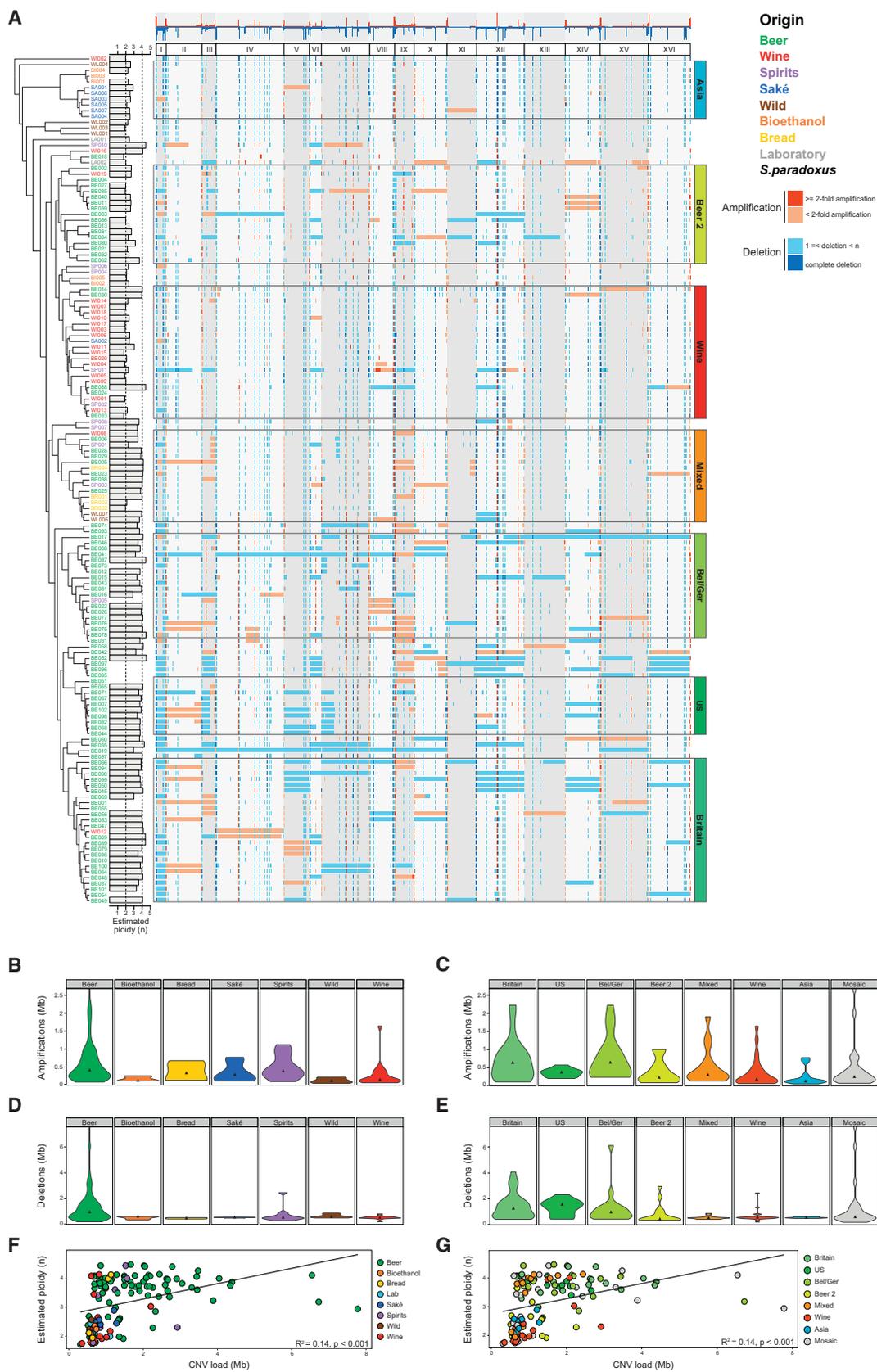
Sequencing the yeast strains in their natural ploidy allowed analysis of gross chromosomal rearrangements and aneuploidies (Figure 2A). We detected a staggering 15,288 deletion and amplification events across all strains, covering on average 1.57 Mb per strain. The size of the regions ranges from complete chromosomes (resulting in aneuploidies) to small local variations of a few kilobases (kb), all of which we will refer to as "CNVs." The extent of deletions significantly exceeds that of amplifications, respectively 1.07 Mb and 0.50 Mb on average per strain (2.15-fold difference, Wilcoxon signed rank test,  $p < 0.001$ ). We observed significant variation among strains originating from different industries in the total frequency of CNV events (ANOVA F test,  $p < 0.001$ ) and the fraction of the genome affected (ANOVA F-test,  $p < 0.001$ ) (Figure S2). Pairwise comparisons of subpopulations and industries show no significant differences in the load of amplifications between strains from different industries or subpopulations, but we detected significant differences in the load of deletions between strains from the wine (median = 0.51 Mb) and beer (median = 0.94 Mb) industry (Tukey honest significant difference [HSD],  $p < 0.05$ ) (Figures 2B–2E). This high incidence of CNV in beer strains goes together with a high

incidence of polyploidy and aneuploidy ( $R^2 \sim 0.14$ ,  $p < 0.001$ ; average genome content of 3.52, SD = 0.67, Figures 2A, 2F, and 2G), which is linked to extensive chromosomal loss and general genome instability (Sheltzer et al., 2011).

CNVs are not uniformly spread across the genome. Considering subtelomere lengths of 33 kb (Brown et al., 2010), on average 39.7% of subtelomeric nucleotide positions are affected by CNV events compared to 9.54% of non-subtelomeric nucleotide positions (4.1-fold difference, Wilcoxon signed-rank test,  $p < 0.001$ ). However, not all subtelomeres are equally prone to CNV: most variability is detected in ChrI, ChrVII, ChrVIII, ChrIX, ChrX, ChrXII, ChrXV, and ChrXVI (Figure 2A). Gene ontology (GO) enrichment analysis reveals that genes involved in nitrogen and carbon metabolism, ion transport, and flocculation are most heavily influenced by CNVs (Table S3), which is in line with previous results (Bergström et al., 2014; Dunn et al., 2012). Interestingly, some CNVs seem linked to specific environments (Table S4), suggesting that CNVs may underlie niche adaptation. For example, many genes involved in uptake and breakdown of maltose (present in saké medium, main carbon source in beer, but absent from grape must) are amplified in beer and saké-related subpopulations, while they are often lost in strains from the Wine subpopulation (false discovery rate [FDR]  $q$  value  $< 0.001$ ).

### Relaxed Selection on Sex and Survival in Nature

Apart from selection for industrial traits, domestication is also characterized by relaxed selection and potential loss of costly traits that are not beneficial in the man-made environment (Driscoll et al., 2009). In order to chart the phenome of our collection and investigate signs of selection for some traits and loss of others, 82 phenotypes, such as aroma production, sporulation characteristics, and tolerance to osmolytes, acids, ethanol, and low and high temperatures, were measured in all strains (Figures 3A and S3; Table S5). Hierarchical clustering of the phenotypes resolves the main phylogenetic lineages and reveals a moderate correlation between genotype and phenotype distances between strains (Spearman correlation  $\sim 0.33$ ), which is further increased (Spearman correlation  $\sim 0.36$ ) when mosaic strains, for which genetic distance has no straightforward evolutionary interpretation, are omitted (Figure 3A). Moreover, the clustering splits the collection into two main phenotypic subgroups: one largely overlapping with the Beer 1 clade that contains the majority of the Belgium/Germany, United States, and



(legend on next page)

Britain beer yeasts as well as mosaic strains containing major genome fractions of these subpopulations and a second one where the remaining genetic subpopulations are strongly over-represented (Fisher's exact test, Bonferroni corrected  $p < 0.001$ ). Overall, strains from the Beer 1 clade perform poorly in general stress conditions that are not usually encountered in the brewing environment (Figure S3; Table S6). In contrast, strains from the Wine subpopulation show superior performance in general stress conditions, which likely reflects the high-sugar and high-alcohol environments encountered in wine-making, as well as survival in potentially nutrient-poor and harsh natural environments in between the grape harvest seasons.

*Saccharomyces cerevisiae* is a facultative sexual organism. While its main mode of reproduction is clonal, sporadic sporulation can help to survive periods of stress (Briza et al., 1990). It has been shown that in yeast, sexual reproduction is beneficial when adapting to new, harsh niches, but plays a lesser role in more favorable environments (Goddard et al., 2005; McDonald et al., 2016). Our data show that there are large systematic differences in the reproductive lifestyle of yeasts inhabiting different industrial niches: 44.4% of the Beer 1 population is obligate asexual, while this trait ranges between 0% and 21% in the other populations (Figure 4A) and is absent in wild strains. Furthermore, over 80% of the non-mosaic Beer 1 strains that are able to sporulate show little or no spore viability (Figure 4B). Additionally, beer yeast lineages generally show a high level of heterozygosity, especially Beer 1. Compared to the Wine clade for example, strains from the Beer 1 and Beer 2 clade have on average 5.10-fold (Tukey HSD,  $p < 0.001$ ) and 2.04-fold (Tukey HSD,  $p = 0.06$ ) more heterozygous sites, respectively (Figures 4C, S4A, and S4B). The lack of genetic admixture suggests that this heterozygosity was acquired during long periods of asexual reproduction, rather than through outbreeding. Further analysis of the correlation between sexual lifestyle and genome structure shows that spore viability is weakly anticorrelated with the heterozygosity level ( $R^2 \sim 0.17$ ;  $p < 0.001$ ) and the fraction of the genome associated with large (>20 kb) amplifications and deletions ( $R^2 \sim 0.16$ ;  $p < 0.001$ ), while sporulation efficiency is only significantly anticorrelated with the latter ( $R^2 \sim 0.19$ ;  $p < 0.001$ ) (Figures 4D–4G).

Together, this indicates that the genome of beer yeasts, but not wine yeasts, show signs of decay and loss of survival skills outside a specific man-made environment, probably caused by their long (estimated >75,000 generations) and uninterrupted growth in rich medium.

### Selection for Industrial Phenotypes

A key hallmark of domestication is phenotypic adaptation to artificial, man-made niches and accentuation of traits desirable for humans. Phenotypic evaluation of the strains for industrially relevant traits (including aroma production, ethanol production, and fermentation performance) shows that many strains harbor phenotypic signatures linked to their industrial application. The ability to accumulate high concentrations of ethanol, for example, seems tightly linked to industrial niche. Beer 1 strains typically generate only 7.5%–10% v/v of ethanol, while strains used for the production of high-alcohol products like saké, spirits, wine, and especially bioethanol, can produce up to 14.5% v/v (Figure 3B; Table S6).

With the exception of a few wine yeast characteristics (see earlier) (Figures 3C and 3D), it remains unclear whether genetic and phenotypic variation between *S. cerevisiae* lineages is primarily caused by human-driven selection and domestication, or if neutral genetic drift or non-human selection are involved. To assess this further, we compared the phenotypic behavior of different subpopulations for two industrially relevant traits for which the genetic underpinnings are largely known, namely maltotriose fermentation and the production of 4-vinyl guaiacol (4-VG), the main compound responsible for phenolic off-flavors (POF). Beer yeasts show a significantly higher capacity to metabolize maltotriose, a carbon source specifically found in beer medium (Figure 3E; Table S6). Efficient utilization of maltotriose correlates with the presence of a specific allele (*AGT1*) of the sugar transporter *MAL11*, known to show high affinity for maltotriose (phenotypic variability explained by SNPs in *MAL11*  $\sim 77.40\%$ , SE 0.5%). This allele is only present in Beer 1 subpopulations and some mosaic strains, while the complete *MAL1* locus (including the *MAL11* gene) is absent in the Wine subpopulation (Table S7). Interestingly, strains of the Beer 2 subpopulation are generally able to ferment maltotriose but contain various frameshift mutations in *MAL11* and show a reduced CNV for the complete *MAL1* locus, suggesting that other, yet unknown mechanisms facilitate maltotriose uptake in this lineage, and maltotriose metabolism evolved convergently in the Beer 1 and Beer 2 lineages.

Yeasts used for the production of alcoholic beverages ideally should not produce undesirable aromas. Although tolerated in some specialty beers, the presence of 4-VG, a compound with a spicy, clove-like aroma, is generally undesired in saké, wine, and most beer styles. Two genes, phenylacrylic acid decarboxylase (*PAD1*) and ferulic acid decarboxylase (*FDC1*), both

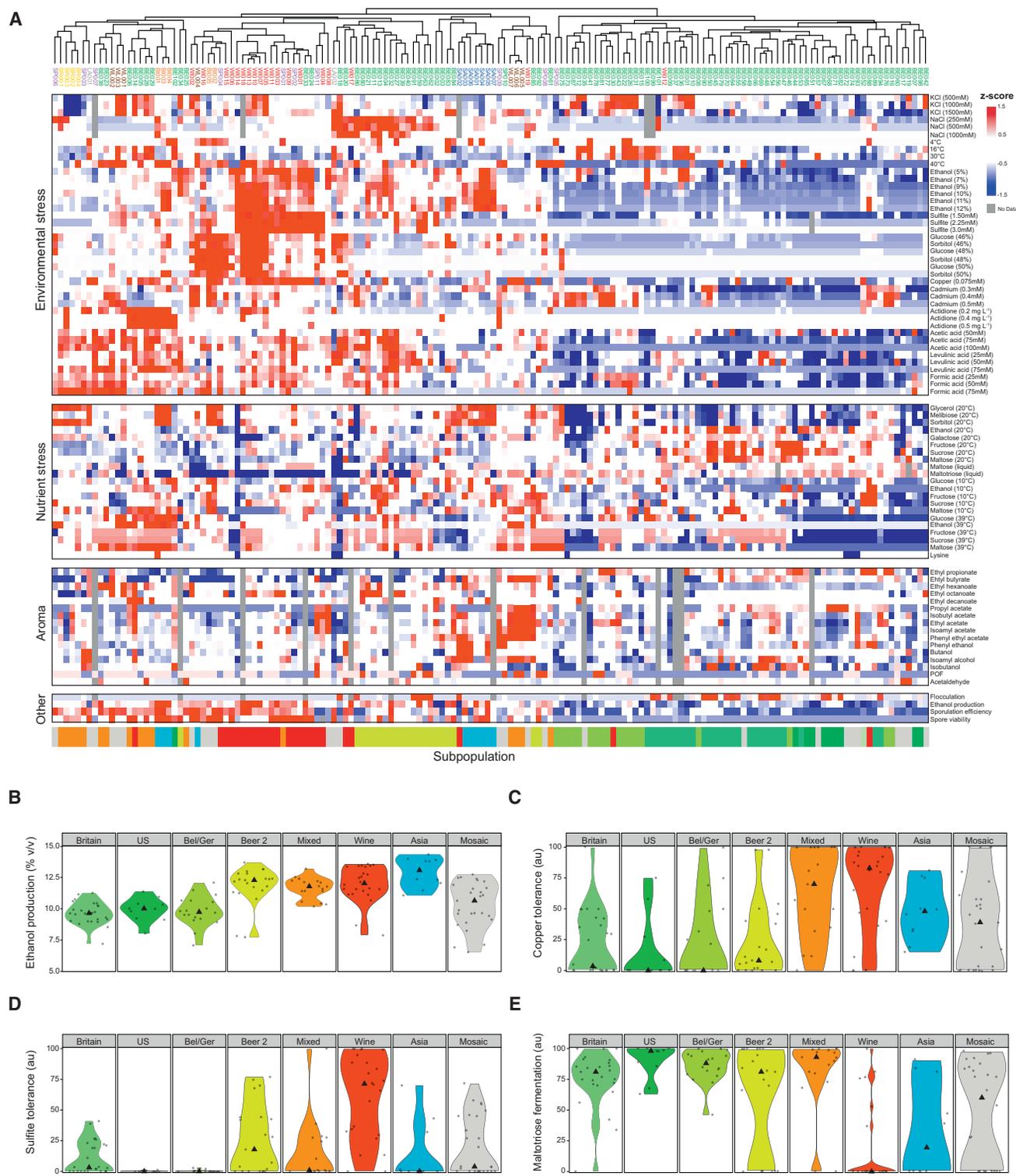
### Figure 2. Ploidy and Copy-Number Variation in Industrial *S. cerevisiae* Strains

(A) Genome-wide visualization of copy-number variation (CNV) profiles, with the aggregate profile across all strains depicted on the top. Estimates for the nominal ploidy ( $n$ ) values of the strains are represented by a bar chart next to the strain codes. Heat map colors reflect amplification (red shades) or deletion (blue shades) of genomic fragments. A distinction is made between completely deleted fragments (dark blue) and fragments of which at least one copy is still present (light blue). Similarly, highly amplified fragments (copy number  $\geq 2$ -fold the basal ploidy) are depicted in dark red, while low and moderately amplified fragments (copy number <2-fold the basal ploidy) are depicted in orange. For strains with no estimated ploidy available, colors are only indicative of the presence of amplifications (orange) or deletions (light blue). Roman numbers indicate chromosome number. Strains are clustered according to their genetic relatedness as determined in Figure S1A. Origin (name colors) and population (colored rectangles) are indicated on the figure.

(B–E) Violin plots describing the density of amplifications and deletions across different industries and subpopulations. Triangles indicate the median within each group.

(F and G) Correlations between levels of CNV load (Mb) and estimated ploidy ( $n$ ), by industry and subpopulations.

See also Figure S2 and Tables S3 and S4.



**Figure 3. Trait Variation of Industrial *S. cerevisiae* Strains**

(A) Heat map representation of phenotypic diversity within industrial *S. cerevisiae* strains. Phenotypic values are calculated as Z scores (normalized values) and colored according to the scale on the right. Missing values are represented by gray shadings. Strains are hierarchically clustered based on phenotypic behavior. Strain names are colored according to geographical origin, as in Figure 1A. The corresponding subpopulation of each strain is indicated by the colored bar below the figure, according to the color code of Figure 1B.

(B) Ethanol production (depicted as % v/v<sup>-1</sup>) of all strains from different subpopulations in fermentation medium containing 35% glucose.

(legend continued on next page)

located as a cluster in the subtelomeric region of ChrIV, control 4-VG production. *PAD1* encodes a flavin prenyltransferase that catalyzes the formation of a flavin-derived cofactor, which is required by Fdc1 for decarboxylation of the precursor ferulic acid (White et al., 2015). Pad1 and Fdc1 help to detoxify phenylacrylic acids found in plant cell walls (Mukai et al., 2010). Therefore, it would be expected that, unless there is counterselection, activity of these genes is retained. Interestingly, phenotypic profiling reveals that many industrial strains have lost the ability to produce 4-VG, while it is generally retained in wild strains, as well as in bakery and bioethanol strains (Figure 5B). In these cases, 4-VG production is likely less detrimental, either because the flavor disappears during baking, or the product is not destined for consumption. Sequence analysis shows that many industrial strains, especially beer and saké strains, acquired loss-of-function mutations (SNPs and/or frameshift Indels) in *PAD1* and/or *FDC1*, while this was never observed in strains from natural environments or bioethanol production (Figure 5A). Moreover, different sublineages acquired different disruptive mutations, hinting to the presence of diverse convergent adaptive strategies in response to human selection against 4-VG production.

To investigate the origin and the maintenance of the phenotypic diversity in 4-VG production, we used Bayesian inference to reconstruct the ancestral phenotypic state in the two key genes *PAD1* and *FDC1* using BEAST (Drummond et al., 2012) (Figure 5C). Shifts from 4-VG<sup>+</sup> to 4-VG<sup>-</sup> and vice versa occurred frequently after the initial split from *S. paradoxus*. In both the *PAD1* and the *FDC1* trees, an early subclade containing most Beer 1 strains acquired loss-of-function mutations at the base of the clade, suggesting that already very early during domestication of the Beer 1 lineage, a 4-VG<sup>-</sup> variant was derived from the 4-VG<sup>+</sup> ancestor. Several other loss- and gain-of-function mutations occurred across both trees, most notably the loss-of-function mutation in *FDC1* of the Asian saké (but not bioethanol) strains.

Interestingly, a strong incongruence between single gene trees and the strain phylogeny is present for three beer strains used in the production of German Hefeweizen beers (BE072, BE074, and BE093). Hefeweizen (wheat) beer is a traditional German beer style and one of the few styles where a high 4-VG level is desirable because it contributes to the typical smoky, spicy aroma of these beers. Phylogenetically, Hefeweizen yeasts cluster within the Beer 1 lineage, but they are shown to be highly mosaic, containing genomic fragments of all three Beer 1 subclades (mainly from Belgium/Germany). Only a small fraction (~8%–13%) of the genome originates from the Wine subpopulation, but this fraction includes the subtelomeric region of ChrIV, containing a functional *PAD1* and *FDC1* allele. This suggests that hybridization between different domesticated subpopulations yielded variants combining the typical traits of beer yeasts, including maltotriose fermentation, with a particular trait

from a wine strain (4-VG production) that is only desirable in special beer styles.

### Creating Superior Hybrid Yeasts through Marker-Assisted Breeding

Apart from yielding insight into the origins of today's industrial yeasts, our results also open new routes for the creation of new superior strains. The availability of genomic data and the increasing number of polymorphisms that are known to contribute to industrially relevant phenotypes enables rapid DNA-based selection of superior segregants and hybrids in large-scale breeding schemes. Such marker-assisted breeding is already intensively used for crop and livestock breeding, because it circumvents labor-intensive and time-consuming phenotyping. As proof-of-concept, we combined our genomic and phenotypic data to obtain new hybrids with altered aromatic properties using marker-assisted breeding. Specifically, a 4-VG producing beer strain harboring a heterozygous loss-of-function mutation in *FDC1* (strain BE027) was selected and sporulated to obtain segregants. Next, the *FDC1* allele of the segregants was genotyped using mismatch PCR. Two segregants, one harboring the loss-of-function allele and one harboring the functional allele, were crossed with segregants of SA005, an Asian saké strain with a homozygous non-functional *FDC1* allele, resulting in hybrids with good beer fermentation characteristics but drastically different aroma profiles (4-VG<sup>+</sup> versus 4-VG<sup>-</sup>) that suit specific beer styles (Figure 5D).

### Domestication Predates Microbe Discovery

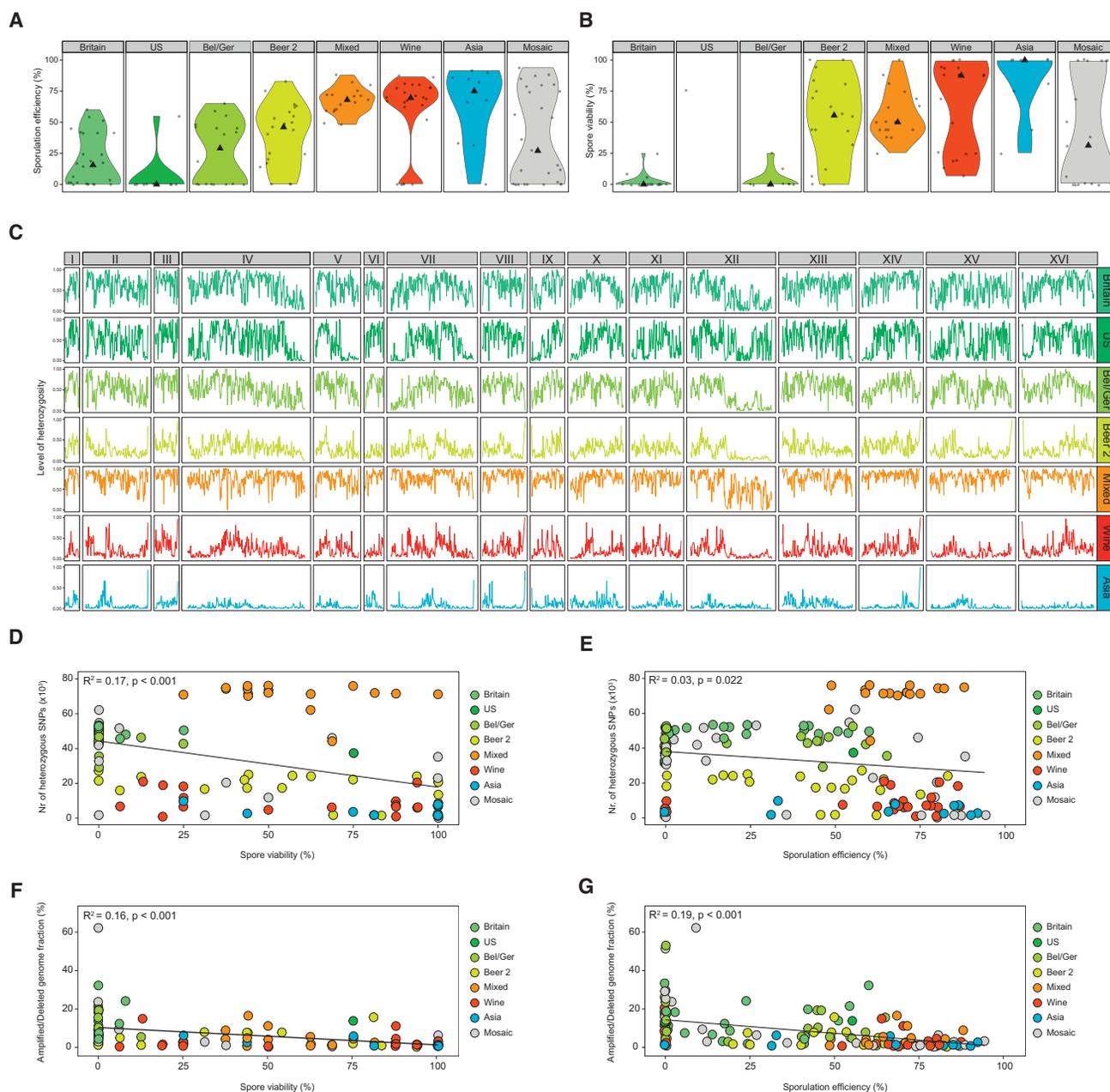
Despite its wide use in industry and as a model organism, little is known about the ecology and evolutionary history of *S. cerevisiae*. Moreover, because early brewers, winemakers, and bakers were unaware of the existence of yeast, there is no record of how yeasts made their way into these processes, nor how yeasts were propagated and shared. As a result, it has proven difficult to estimate when specific industrial lineages originated. Moreover, current demographic and molecular clock models of *S. cerevisiae* employ the experimentally determined mutation rate of the haploid lab strain S288c in rich growth medium (Lynch et al., 2008), while it is known that the mutation rate is heavily influenced by the genetic background (Filteau et al., 2015), ploidy (Sheltzer et al., 2011), growth speed (van Dijk et al., 2015), and environmental stress (Voordeckers et al., 2015), factors that are likely very different for industrial, wild, and lab yeasts. However, our dataset, and specifically the Beer 1 clade, provides a strong tool for dating beer yeast divergence. First, given the absence of a functional sexual cycle and lack of admixture, exclusively clonal reproduction can be assumed. Second, our data show that United States beer yeasts are related closest to European beer yeasts, suggesting that they were imported from Europe during colonization, rather than stemming from indigenous wild United States yeasts. More specifically, United States beer yeasts seem phylogenetically most closely related to British beer yeasts (Figure 1A), which is

(C) Growth of all strains from different subpopulations on medium supplemented with 0.075 mM copper, relative to growth on medium without copper.

(D) Growth of all strains from different subpopulations on medium supplemented with 2.25 mM sulfite, relative to growth on medium without sulfite.

(E) Growth of all strains from different subpopulations in medium containing 1% w/v<sup>-1</sup> maltotriose as the sole carbon source, relative to growth on medium with 1% w/v<sup>-1</sup> glucose. au, arbitrary units; Bel/Ger, Belgium/Germany.

See also Figure S3 and Tables S5, S6, and S7.



**Figure 4. The Reproductive Lifestyle of Industrial *S. cerevisiae* Strains**

(A) Violin plots depicting sporulation efficiency of all strains from different subpopulations.

(B) Violin plots depicting spore viability of all sporulating strains from different subpopulations.

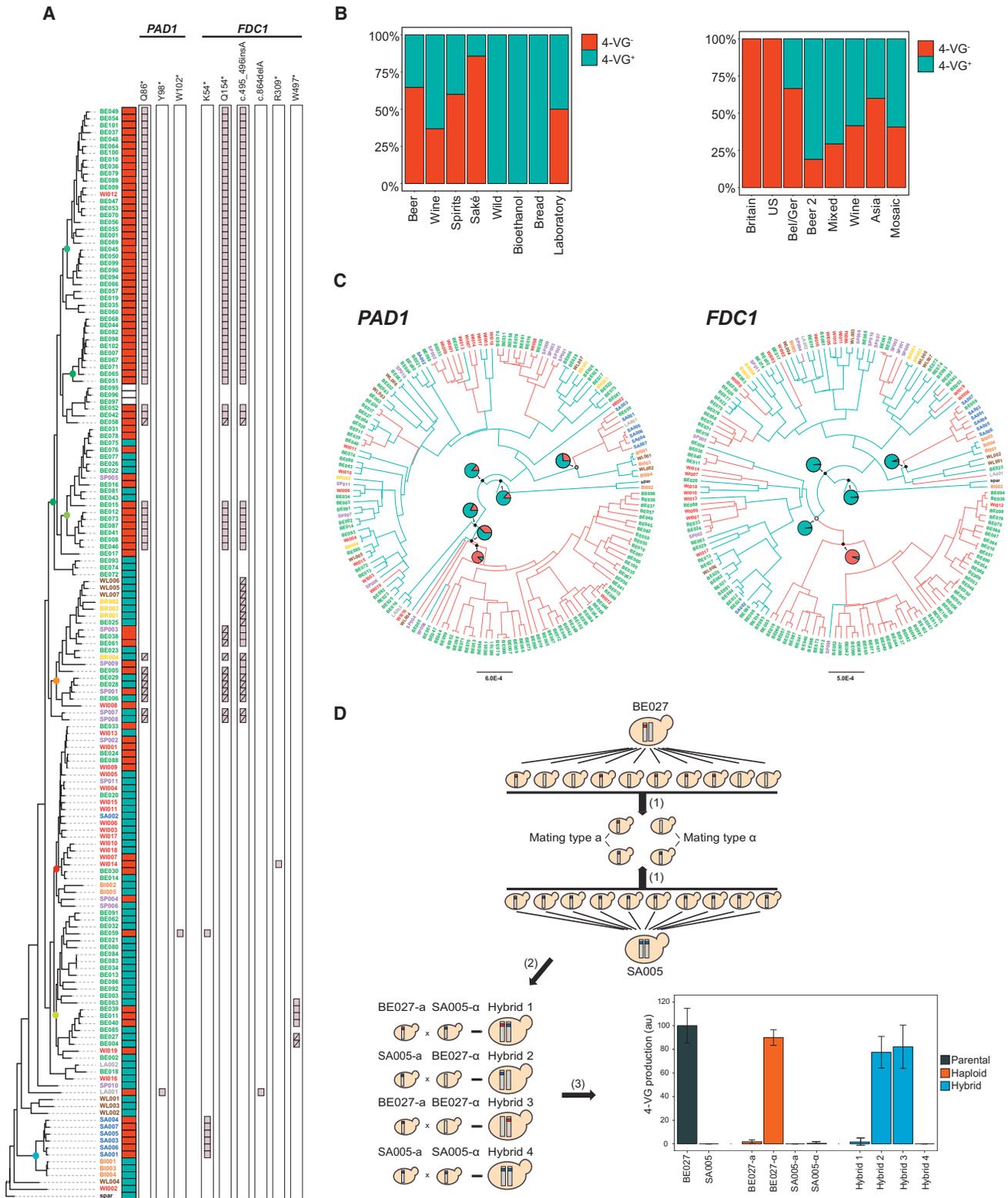
(C) Visualization of the level of heterozygosity across the genome of the different subpopulations, calculated as the ratio of heterozygous/homozygous SNPs in 10 kb windows.

(D–G) Scatter plots depicting the correlation between the number of heterozygous loci and spore viability (D) or sporulation efficiency (E), and the correlation between the fraction of the genome subjected to large (>20 kb) structural variation and spore viability (F) or sporulation efficiency (G). Dot colors indicate subpopulations similar to the color code of Figure 1B.

See also Figure S4.

confirmed by the average per-site nucleotide divergence ( $d_{xy}$ ), which is significantly lower between Britain and the United States (average  $d_{xy} = 1.97 \times 10^{-3}$ ) than between Belgium/Germany and the United States strains (average  $d_{xy} = 2.26 \times 10^{-3}$ ) (Wilcoxon

signed-rank test,  $p < 0.001$ ) (Table S8). This suggests that the origin of the United States brewing strains can be traced back to the introduction of beer culture in the United States by early 17<sup>th</sup> century British settlers (Van Wieren, 1995). Third, in contrast



**Figure 5. Production of 4-Vinyl Guaiacol by Industrial *S. cerevisiae* Strains**

(A) Distribution of loss-of-function SNPs and frame-shift mutations in *FDC1* and *PAD1* of industrial *S. cerevisiae* strains. Gray boxes indicate the presence of the loss-of-function mutation, diagonal bars indicate heterozygosity at this site. Strains are clustered according to the strain phylogeny and strain names are colored according to their origin. Basal splits of the five industrial lineages are indicated with colored dots (see Figure 1A).

(legend continued on next page)

to wine, beer is not produced seasonally but throughout the whole year, which provides fermenting yeast with a predictable and stable growth environment. Yeast cells undergo about three doublings during one batch of beer fermentation, which takes  $\sim 1$  week. Moreover, brewers typically recycle yeasts from a finished fermentation to inoculate a new batch, which implies that beer yeasts are continuously growing in their industrial niche. Together, these facts make it possible to estimate the number of generations to be  $\sim 150$ /year.

Based on estimates of the number of generations per year and the divergence time between United Kingdom and United States beer strains, we calculated the average mutation rate in a brewing environment to be  $1.61\text{--}1.73\text{E-}08$ /bp/generation. While this value differs from previous assumptions, it is similar to the measured mutation rate in a diploid yeast strain that was subjected to 2 years of artificial evolution in a high-ethanol environment (Voordeckers et al., 2015). Moreover, mutations likely also occur in the second phase of beer fermentations, when cells are no longer dividing, which implies that the mutation rate per generation in industrial conditions should be higher than what is measured under conditions where the cells are dividing frequently, as is usually the case in laboratory experiments (Loewe et al., 2003). Using these data, the last common ancestor of the three major Beer 1 subclades (Belgium/Germany, United Kingdom, and the United States) is estimated to date from AD 1573–1604, suggesting that domestication started around this time. Interestingly, this coincides with the gradual switch from home-centered beer brewing where every family produced their own beer, to more professional large-scale brewing, first in pubs and monasteries and later also in breweries (Hornsey, 2003). The last common ancestor of Beer 2 is estimated to be more recent, between AD 1645–1671. This suggests that beer yeast domestication started before the discovery of microbes and the isolation of the first pure yeast cultures by Emil Hansen in the Carlsberg brewery in 1883, but well after the invention of beer production, estimated to have occurred as early as 3000 BC (Michel et al., 1992). Although it is difficult to assess how many different yeast strains were domesticated and in which industrial context these domestications occurred, the limited number of clades of industrial yeasts and the clear segregation of wild and industrial yeasts suggests that today's industrial yeasts originated from a limited set of ancestral strains, or closely related groups of ancestral strains.

## DISCUSSION

Together, our results show that today's industrial *S. cerevisiae* yeasts are genetically and phenotypically separated from wild

stocks due to human selection and trafficking. Specifically, the thousands of industrial yeasts that are available today seem to stem from only a few ancestral strains that made their way into food fermentations and subsequently evolved into separate lineages, each used for specific industrial applications. Within each cluster, strains are sometimes further subdivided along geographical boundaries, as is the case for the Beer 1 clade, which is divided into three main subgroups. However, further subclustering of beer yeasts according to beer style was generally not observed, which may not be surprising as it is common practice for brewers to use only one yeast strain within their brewery for the production of a wide array of different beers. Notable exceptions are yeasts associated with the few beers that largely depend on very specific yeasts characteristics, such as Hefeweizen beers. Another exception may include those beers for which production is restricted to a specific geographic area, such as Belgian Saisons or British Stouts.

We further show that industrial yeasts were clearly subjected to domestication, which is reflected in their genomes and phenomes. Interestingly, domestication seems strongest in beer yeasts, which demonstrate domestication hallmarks such as decay of sexual reproduction and general stress resistance, as well as convergent evolution of desirable traits like maltotriose utilization. Yeasts from the Beer 1 clade show the clearest signs of domestication, possibly because Beer 2 only diverged more recently from other sublineages. Many of these domestication features may have simply been the result of the yeasts' adaptation to their new industrial niches. However, for some traits, it is likely that humans actively intervened, e.g., by selecting strains that do not produce undesirable off-flavors, which our analysis identifies as *PAD1* or *FDC1* nonsense mutants.

The presence of a strong domestication signature in beer yeast genomes agrees well with the common practices in the brewing industry. Beer yeasts are typically recycled after each fermentation batch, and because beer is produced throughout the year, this implies that beer yeasts are continuously growing in their industrial niche. By contrast, wine yeasts can only grow in wine must for a short period every year, spending the rest of their lives in and around the vineyards or in the guts of insects (Bokulich et al., 2014; Christiaens et al., 2014; Stefanini et al., 2016). During these nutrient-poor periods, wine yeasts likely undergo few mitotic doublings, yet they may undergo sexual cycles and even hybridize with wild yeasts (Stefanini et al., 2016). Moreover, only a very small portion of the yeasts may find their way back into the grape must when the next harvest season arrives, while trillions of cells are being transferred to the next batch during backslipping in beer production. This results in large

(B) Percentage of strains within each origin (left) and population (right) capable of producing 4-vinyl guaiacol (4-VG). Red, 4-VG<sup>-</sup>; turquoise, 4-VG<sup>+</sup>.

(C) Phylogenetic trees and ancestral trait reconstruction of *PAD1* and *FDC1* genes. Branches are colored according to the most probable state of their ancestral nodes, turquoise (4-VG<sup>+</sup>) or red (4-VG<sup>-</sup>). Pie charts indicate probabilities of each state at specific nodes, turquoise (4-VG<sup>+</sup>) or red (4-VG<sup>-</sup>); posterior probability for the same nodes is indicated by a dot: black dot, 90%–100%; gray, 70%; white, 42%. Branch lengths reflect the average numbers of substitutions per site (compare scale bars).

(D) Development of new yeast variants with specific phenotypic features by marker-assisted breeding. Two parent strains (BE027 and SA005) were sporulated and, using genetic markers, segregants with the desired genotype were selected (1). Next, breeding between segregants from different parents (outbreeding) or the same parent (inbreeding) were performed (2). This breeding scheme yields hybrids with altered aromatic properties that can directly be applied in industrial fermentations (3). 4-VG production is shown relative to the production of BE027. Yeast genomes are represented by gray bars, loss-of-function mutations in *FDC1* as red (W497\*) and blue (K54\*) boxes within the gray bars. Error bars represent one SD from the mean.

See also Table S5.

effective population sizes for beer, but not wine, yeasts. The differences in industrial practices between beer brewing and wine-making likely had three important consequences. First, beer yeasts evolved faster than wine strains (Figure 1A). This resulted in a large genetic diversity within beer yeasts, while wine yeasts are genetically more homogeneous (Table 1). Second, after the initial domestication event, some beer yeasts were contained in the brewery and diverged allopatrically, leading to geographically defined subpopulations mirroring human traffic and colonization. Third, beer strains generally lost their ability to reproduce sexually. This, combined with continuous cultivation in a mild growth environment, made them susceptible to genetic drift and fixation of deleterious alleles that would otherwise be purged by evolutionary competition in harsh conditions. Hence, these asexual populations continuously accumulated deleterious mutations in an irreversible manner, a process known as Muller's ratchet (Muller, 1964). We propose that continuous clonal reproduction and relaxed selection for general stress resistance and famine likely allowed genome decay in beer yeasts and resulted in yeasts specialized in thriving in a man-made niche like beer fermentations, but not in natural environments. Both these characteristics (genome decay and niche specialization) are considered to be key characteristics of domestication.

Our study does not only provide insight into the domestication origin of industrial yeasts, it may also help to select and breed new superior strains. The genome sequences, phylogenetic tree, and phenome data can be used to set up marker-assisted breeding schemes similar to those routinely used for the breeding of superior crops and livestock (Takeda and Matsuoka, 2008).

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Strain Collection
- METHOD DETAILS
  - DNA Extraction
  - Library Prep and Whole Genome Sequencing
  - De Novo Assembly
  - Annotation
  - Core Genome Analysis and Identification of Single Copy Genes
  - Reference-Based Alignments and Variant Calling
  - Phylogenetic Analyses
  - Population Structure and Diversity Analysis
  - Time Divergence Estimate
  - Copy-Number Variation Analysis
  - Character Evolution Analysis
  - Determination of Cell Ploidy
  - Phenotypic Analysis
  - Development of Artificial Hybrids
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY
  - Data Resources

## SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and eight tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.08.020>.

An audio PaperClip is available at <http://dx.doi.org/10.1016/j.cell.2016.08.020#mmc9>.

## AUTHOR CONTRIBUTIONS

Conceptualization, B.G., J.S., T.P., L.M., S.M., and K.J.V.; Formal Analysis, B.G., J.S., L.S., M.R., A.M., and K.V.; Investigation, B.G., J.S., T.P., A.M., L.S., V.S., B.H.-M., and M.T.; Resources, L.M., B.S., C.T., C.W., and K.J.V.; Writing, B.G., J.S., S.M., and K.J.V.; Supervision, T.P., T.R., A.S., C.W., G.B., S.M., and K.J.V.

## ACKNOWLEDGMENTS

We thank all K.J.V. and S.M. laboratory members for their help and suggestions. We thank K. Wolfe, S. Oliver, and P. Malcorps for their valuable feedback on the manuscript. We acknowledge D. Bami, A. Bass, J. Kurowski, K. Fortmann, and N. Parker for data collection and data review. Additionally we acknowledge S. Rombauts, Y. Lin, R. de Jonge, and V. Storme for fruitful discussions on data analysis. B.G. acknowledges funding from the VIB International PhD Program in Life Sciences. J.S. acknowledges funding from IWT and KU Leuven. K.J.V. acknowledges funding from an ERC Consolidator grant CoG682009, HFSP program grant RGP0050/2013, KU Leuven NATAR Program Financing, VIB, EMBO YIP program, FWO, and IWT. S.M. acknowledges funding from VIB and Ghent University. The funders had no role in study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

Received: March 24, 2016

Revised: June 8, 2016

Accepted: August 8, 2016

Published: September 8, 2016

## REFERENCES

- Almeida, P., Barbosa, R., Zalar, P., Imanishi, Y., Shimizu, K., Turchetti, B., Legras, J.L., Serra, M., Dequin, S., Couloux, A., et al. (2015). A population genomics insight into the Mediterranean origins of wine yeast domestication. *Mol. Ecol.* **24**, 5412–5427.
- Baele, G., and Lemey, P. (2013). Bayesian evolutionary model testing in the phylogenomics era: matching model complexity with computational efficiency. *Bioinformatics* **29**, 1970–1979.
- Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., and Alekseyenko, A.V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* **29**, 2157–2167.
- Bergström, A., Simpson, J.T., Salinas, F., Barré, B., Parts, L., Zia, A., Nguyen Ba, A.N., Moses, A.M., Louis, E.J., Mustonen, V., et al. (2014). A high-definition view of functional genetic variation from natural yeast genomes. *Mol. Biol. Evol.* **31**, 872–888.
- Bokulich, N.A., Thorngate, J.H., Richardson, P.M., and Mills, D.A. (2014). Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proc. Natl. Acad. Sci. USA* **111**, E139–E148.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.
- Borneman, A.R., Desany, B.A., Riches, D., Affourtit, J.P., Forgan, A.H., Pretorius, I.S., Egholm, M., and Chambers, P.J. (2011). Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet.* **7**, e1001287.
- Borneman, A.R., Forgan, A.H., Kolouchova, R., Fraser, J.A., and Schmidt, S.A. (2016). Whole genome comparison reveals high levels of inbreeding and strain

- redundancy across the spectrum of commercial wine strains of *Saccharomyces cerevisiae*. *G3 (Bethesda)* 6, 957–971.
- Briza, P., Breitenbach, M., Ellinger, A., and Segall, J. (1990). Isolation of two developmentally regulated genes involved in spore wall maturation in *Saccharomyces cerevisiae*. *Genes Dev.* 4, 1775–1789.
- Brown, C.A., Murray, A.W., and Verstrepen, K.J. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr. Biol.* 20, 895–903.
- Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973.
- Christiaens, J.F., Franco, L.M., Cools, T.L., De Meester, L., Michiels, J., Wenseleers, T., Hassan, B.A., Yaksi, E., and Verstrepen, K.J. (2014). The fungal aroma gene *ATF1* promotes dispersal of yeast cells through insect vectors. *Cell Rep.* 9, 425–432.
- Cingolani, P., Platts, A., Wang, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 6, 80–92.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158.
- Dittmar, J.C., Reid, R.J., and Rothstein, R. (2010). ScreenMill: a freely available software suite for growth measurement, analysis and visualization of high-throughput screen data. *BMC Bioinformatics* 11, 353.
- Driscoll, C.A., Macdonald, D.W., and O'Brien, S.J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. *Proc. Natl. Acad. Sci. USA* 106 (Suppl 1), 9971–9978.
- Drummond, A.J., and Suchard, M.A. (2010). Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* 8, 114.
- Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973.
- Dunham, M.J., Badrane, H., Ferea, T., Adams, J., Brown, P.O., Rosenzweig, F., and Botstein, D. (2002). Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 99, 16144–16149.
- Dunn, B., Richter, C., Kvittek, D.J., Pugh, T., and Sherlock, G. (2012). Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res.* 22, 908–924.
- Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10, 48.
- Filteau, M., Hamel, V., Pouliot, M.C., Gagnon-Arsenault, I., Dubé, A.K., and Landry, C.R. (2015). Evolutionary rescue by compensatory mutations is constrained by genomic and environmental backgrounds. *Mol. Syst. Biol.* 11, 832.
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152.
- Goddard, M.R., and Greig, D. (2015). *Saccharomyces cerevisiae*: a nomadic yeast with no niche? *FEMS Yeast Res.* 15, fov009.
- Goddard, M.R., Godfray, H.C.J., and Burt, A. (2005). Sex increases the efficacy of natural selection in experimental yeast populations. *Nature* 434, 636–640.
- Hornsey, I.S. (2003). *A History of Beer and Brewing* (Cambridge: The Royal Society of Chemistry).
- Hutter, S., Vilella, A.J., and Rozas, J. (2006). Genome-wide DNA polymorphism analyses using VariScan. *BMC Bioinformatics* 7, 409.
- Huxley, C., Green, E.D., and Dunham, I. (1990). Rapid assessment of *S. cerevisiae* mating type by PCR. *Trends Genet.* 6, 236.
- Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kozlov, A.M., Aberer, A.J., and Stamatakis, A. (2015). ExaML version 3: a tool for phylogenomic analyses on supercomputers. *Bioinformatics* 31, 2577–2579.
- Kück, P., and Meusemann, K. (2010). FASconCAT: convenient handling of data matrices. *Mol. Phylogenet. Evol.* 56, 1115–1118.
- Kuhn, R.M., Karolchik, D., Zweig, A.S., Trumbower, H., Thomas, D.J., Thakka-pallayil, A., Sugnet, C.W., Stanke, M., Smith, K.E., Siepel, A., et al. (2007). The UCSC genome browser database: update 2007. *Nucleic Acids Res.* 35, D668–D673.
- LANFEAR, R., CALCOTT, B., HO, S.Y.W., and GUINDON, S. (2012). Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol. Biol. Evol.* 29, 1695–1701.
- Legras, J.L., and Karst, F. (2003). Optimisation of interdelta analysis for *Saccharomyces cerevisiae* strain characterisation. *FEMS Microbiol. Lett.* 221, 249–255.
- Lemey, P., Rambaut, A., Drummond, A.J., and Suchard, M.A. (2009). Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5, e1000520.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760.
- Liti, G., Carter, D.M., Moses, A.M., Warringer, J., Parts, L., James, S.A., Davey, R.P., Roberts, I.N., Burt, A., Koufopanou, V., et al. (2009). Population genomics of domestic and wild yeasts. *Nature* 458, 337–341.
- Liu, Y., Schröder, J., and Schmidt, B. (2013). Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. *Bioinformatics* 29, 308–315.
- Loewe, L., Textor, V., and Scherer, S. (2003). High deleterious genomic mutation rate in stationary phase of *Escherichia coli*. *Science* 302, 1558–1560.
- Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C.R., Dopman, E.B., Dickinson, W.J., Okamoto, K., Kulkarni, S., Hartl, D.L., and Thomas, W.K. (2008). A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc. Natl. Acad. Sci. USA* 105, 9272–9277.
- Magwene, P.M., Kayıkcı, Ö., Granek, J.A., Reininga, J.M., Scholl, Z., and Murray, D. (2011). Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 108, 1987–1992.
- McDonald, M.J., Rice, D.P., and Desai, M.M. (2016). Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature* 531, 233–236.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303.
- Michel, R.H., McGovern, P.E., and Badler, V.R. (1992). Chemical evidence for ancient beer. *Nature* 360, 24.
- Mukai, N., Masaki, K., Fujii, T., Kawamukai, M., and Iefuji, H. (2010). *PAD1* and *FDC1* are essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*. *J. Biosci. Bioeng.* 109, 564–569.
- Muller, H.J. (1964). The relation of recombination to mutational advance. *Mutat. Res.* 106, 2–9.
- Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R., Moret, B.M., and Stamatakis, A. (2010). How many bootstrap replicates are necessary? *J. Comput. Biol.* 17, 337–354.
- Pavelka, N., Rancati, G., Zhu, J., Bradford, W.D., Saraf, A., Florens, L., Sanderson, B.W., Hatter, G.L., and Li, R. (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* 468, 321–325.
- Peng, Y., Leung, H., Yiu, S., and Chin, F. (2010). IDBA – a practical iterative de Bruijn Graph de novo assembler. In *Research in Computational Molecular Biology*, B. Berger, ed. (Springer), pp. 426–440.

- Pérez-Ortín, J.E., Querol, A., Puig, S., and Barrio, E. (2002). Molecular characterization of a chromosomal rearrangement involved in the adaptive evolution of yeast strains. *Genome Res.* *12*, 1533–1539.
- Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S.E., and Lercher, M.J. (2014). PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* *31*, 1929–1936.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* *81*, 559–575.
- Purugganan, M.D., and Fuller, D.Q. (2009). The nature of selection during plant domestication. *Nature* *457*, 843–848.
- R Development Core Team (2011). R: A language and environment for statistical computing (R Foundation for Statistical Computing).
- Raj, A., Stephens, M., and Pritchard, J.K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* *197*, 573–589.
- Rambaut, A., Suchard, M.A., Xie, D., and Drummond, A.J. (2014) Tracer v1.6. <http://beast.bio.ed.ac.uk/Tracer>.
- Rancati, G., Pavelka, N., Fleharty, B., Noll, A., Trimble, R., Walton, K., Perera, A., Staehling-Hampton, K., Seidel, C.W., and Li, R. (2008). Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell* *135*, 879–893.
- Ranwez, V., Harispe, S., Delsuc, F., and Douzery, E.J.P. (2011). MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PLoS ONE* *6*, e22594.
- Schacherer, J., Shapiro, J.A., Ruderfer, D.M., and Kruglyak, L. (2009). Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature* *458*, 342–345.
- Selmecki, A.M., Dulmage, K., Cowen, L.E., Anderson, J.B., and Berman, J. (2009). Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. *PLoS Genet.* *5*, e1000705.
- Sheltzer, J.M., Blank, H.M., Pfau, S.J., Tange, Y., George, B.M., Humpton, T.J., Brito, I.L., Hiraoka, Y., Niwa, O., and Amon, A. (2011). Aneuploidy drives genomic instability in yeast. *Science* *333*, 1026–1030.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* *30*, 1312–1313.
- Stanke, M., and Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* *33* (Web Server issue, suppl 2), W465–W467.
- Steensels, J., and Verstrepen, K.J. (2014). Taming wild yeast: potential of conventional and nonconventional yeasts in industrial fermentations. *Annu. Rev. Microbiol.* *68*, 61–80.
- Steensels, J., Meersman, E., Snoek, T., Saels, V., and Verstrepen, K.J. (2014). Large-scale selection and breeding to generate industrial yeasts with superior aroma production. *Appl. Environ. Microbiol.* *80*, 6965–6975.
- Stefanini, I., Dapporto, L., Berná, L., Polsinelli, M., Turillazzi, S., and Cavalieri, D. (2016). Social wasps are a *Saccharomyces* mating nest. *Proc. Natl. Acad. Sci. USA* *113*, 2247–2251.
- Strope, P.K., Skelly, D.A., Kozmin, S.G., Mahadevan, G., Stone, E.A., Magwene, P.M., Dietrich, F.S., and McCusker, J.H. (2015). The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Res.* *25*, 762–774.
- Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* *34* (Web Server issue, suppl 2), W609–W612.
- Takeda, S., and Matsuoka, M. (2008). Genetic approaches to crop improvement: responding to environmental and population changes. *Nat. Rev. Genet.* *9*, 444–457.
- van Dijk, D., Dhar, R., Missarova, A.M., Espinar, L., Blevins, W.R., Lehner, B., and Carey, L.B. (2015). Slow-growing cells within isogenic populations have increased RNA polymerase error rates and DNA damage. *Nat. Commun.* *6*, 7972.
- Van Wieren, D.P. (1995). American Breweries II (Eastern Coast Breweriana Association).
- Voordeckers, K., Kominek, J., Das, A., Espinosa-Cantú, A., De Maeyer, D., Arslan, A., Van Pee, M., van der Zande, E., Meert, W., Yang, Y., et al. (2015). Adaptation to high ethanol reveals complex evolutionary pathways. *PLoS Genet.* *11*, e1005635.
- Wang, Q.-M., Liu, W.-Q., Liti, G., Wang, S.A., and Bai, F.Y. (2012). Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. *Mol. Ecol.* *21*, 5404–5417.
- Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. *Nature* *449*, 54–61.
- Warringer, J., Zörgö, E., Cubillos, F.A., Zia, A., Gjuvslund, A., Simpson, J.T., Forsmark, A., Durbin, R., Omholt, S.W., Louis, E.J., et al. (2011). Trait variation in yeast is defined by population history. *PLoS Genet.* *7*, e1002111.
- White, M.D., Payne, K.A.P., Fisher, K., Marshall, S.A., Parker, D., Rattray, N.J.W., Trivedi, D.K., Goodacre, R., Rigby, S.E.J., Scrutton, N.S., et al. (2015). UbiX is a flavin prenyltransferase required for bacterial ubiquinone biosynthesis. *Nature* *522*, 502–506.
- Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., and Weir, B.S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* *28*, 3326–3328.

# Zika Virus Infection during Pregnancy in Mice Causes Placental Damage and Fetal Demise

Jonathan J. Miner,<sup>1,8</sup> Bin Cao,<sup>5,8</sup> Jennifer Govero,<sup>1,8</sup> Amber M. Smith,<sup>1</sup> Estefania Fernandez,<sup>1,4</sup> Omar H. Cabrera,<sup>6</sup> Charise Garber,<sup>1</sup> Michelle Noll,<sup>1</sup> Robyn S. Klein,<sup>1,2,3,4</sup> Kevin K. Noguchi,<sup>6</sup> Indira U. Mysorekar,<sup>4,5,\*</sup> and Michael S. Diamond<sup>1,3,4,7,\*</sup>

<sup>1</sup>Department of Medicine

<sup>2</sup>Department of Anatomy and Neurobiology

<sup>3</sup>Department of Molecular Microbiology

<sup>4</sup>Department of Pathology and Immunology

<sup>5</sup>Department of Obstetrics and Gynecology

<sup>6</sup>Department of Psychiatry

<sup>7</sup>The Center for Human Immunology and Immunotherapy Programs  
Washington University School of Medicine, Saint Louis, MO 63110, USA

<sup>8</sup>Co-first authors

\*Correspondence: mysorekari@wudosis.wustl.edu (I.U.M.), diamond@borcim.wustl.edu (M.S.D.)

<http://dx.doi.org/10.1016/j.cell.2016.05.008>

## SUMMARY

Zika virus (ZIKV) infection in pregnant women causes intrauterine growth restriction, spontaneous abortion, and microcephaly. Here, we describe two mouse models of placental and fetal disease associated with in utero transmission of ZIKV. Female mice lacking type I interferon signaling (*Ifnar1*<sup>-/-</sup>) crossed to wild-type (WT) males produced heterozygous fetuses resembling the immune status of human fetuses. Maternal inoculation at embryonic day 6.5 (E6.5) or E7.5 resulted in fetal demise that was associated with ZIKV infection of the placenta and fetal brain. We identified ZIKV within trophoblasts of the maternal and fetal placenta, consistent with a trans-placental infection route. Antibody blockade of *Ifnar1* signaling in WT pregnant mice enhanced ZIKV trans-placental infection although it did not result in fetal death. These models will facilitate the study of ZIKV pathogenesis, in utero transmission, and testing of therapies and vaccines to prevent congenital malformations.

## INTRODUCTION

Zika virus (ZIKV) is a mosquito-transmitted flavivirus that was first isolated from a febrile rhesus macaque in Uganda in 1947 and is related to other globally relevant arthropod-transmitted human pathogens including dengue (DENV), yellow fever (YFV), West Nile (WNV), Japanese encephalitis (JEV), and tick-borne encephalitis viruses (Lazear and Diamond, 2016). Over the last decade, ZIKV has emerged from a relatively obscure status to causing large epidemics in Micronesia, French Polynesia, and South and Central America. Although in most instances ZIKV infection results in a self-limiting febrile illness associated with rash and conjunctivitis, severe neurological phenotypes can

occur, including Guillain-Barre syndrome and meningoencephalitis (Carteaux et al., 2016; Oehler et al., 2014). Infection in pregnant women is of major concern, as it is linked to catastrophic fetal abnormalities including microcephaly, spontaneous abortion, and intrauterine growth restriction (IUGR) due to placental insufficiency (Brasil et al., 2016). Because of the growing public health concern, there is an urgent need to establish animal models of intrauterine ZIKV infection that define mechanisms of fetal transmission and facilitate testing of therapeutics and vaccines. Furthermore, an in utero animal model of ZIKV infection would establish causality and satisfy the criteria for proof of teratogenicity (Rasmussen et al., 2016).

In 2015, Brazil experienced a sharp rise in the number of cases of pregnancy-associated microcephaly, and this was linked to a concurrent epidemic of ZIKV infection. Mounting evidence suggests that ZIKV infection in pregnant women causes congenital abnormalities and fetal demise (Brasil et al., 2016; Sarno et al., 2016; Ventura et al., 2016). Initial case descriptions of microcephaly and spontaneous abortion have been bolstered by evidence of viral RNA and antigen in the brains of congenitally infected fetuses and newborns (Martines et al., 2016; Mlakar et al., 2016). These findings were substantiated by a prospective study of a cohort of symptomatic, ZIKV-infected pregnant women in which 29% of fetuses exhibited developmental abnormalities including microcephaly and IUGR, which in a subset of cases resulted in fetal demise or stillbirth (Brasil et al., 2016). Preliminary reports suggest that ZIKV-induced fetal abnormalities can occur in all trimesters of pregnancy although the most severe manifestations are associated with infections in the first and second trimesters (Brasil et al., 2016). Congenital abnormalities associated with ZIKV infection also have been described in French Polynesia (by retrospective analysis) and other Latin American countries (Cauchemez et al., 2016). These findings suggest that ZIKV strains in French Polynesia and Latin America share the potential to cause disease during pregnancy.

Recently, we and others have developed models of ZIKV pathogenesis in adult mice that recapitulated several features of human disease (Aliota et al., 2016; Lazear et al., 2016; Rossi et al.,

2016). Whereas 4- to 6-week-old wild-type (WT) mice did not develop overt clinical illness after infection with a contemporary clinical strain of ZIKV, mice lacking the ability to produce or respond to type I interferon (IFN) (e.g., *Ifnar1*<sup>-/-</sup> mice) developed severe neurological disease that was associated with high viral loads in the brain and spinal cord and substantial lethality. In a complementary approach using WT mice treated with a blocking anti-*Ifnar* antibody (MAR1-5A3), we reported a less severe model of ZIKV pathogenesis that also resulted in replication of ZIKV in several organs (Lazear et al., 2016). These animals, however, survived infection and did not develop neurological signs or neuroinvasive disease.

Given the urgent need to understand the basis for in utero transmission of ZIKV and its pathological consequences, we developed two models of ZIKV infection during pregnancy using *Ifnar1*<sup>-/-</sup> females crossed to WT males as well as pregnant WT females treated with an anti-*Ifnar*-blocking antibody. We found that ZIKV infects pregnant dams and the placenta, and this resulted in damage to the placental barrier and infection of the developing fetus, as well as placental insufficiency and IUGR. In severe cases, ZIKV infection of *Ifnar1*<sup>-/-</sup> females led to fetal demise. When dams were treated with an anti-*Ifnar* antibody, infection of the developing fetus occurred but was less severe and did not cause fetal death. These findings establish models for studying mechanisms of in utero transmission and testing of candidate therapies for preventing congenital malformations. They also highlight the concern that ZIKV infection can occur in fetuses of otherwise healthy-appearing dams with uncertain neurodevelopmental consequences.

## RESULTS

Since the type I interferon (IFN) response prevents efficient replication of ZIKV in peripheral organs of WT mice (Lazear et al., 2016), we initially used *Ifnar1*<sup>-/-</sup> mice to facilitate high levels of ZIKV replication during pregnancy. *Ifnar1*<sup>-/-</sup> female mice were bred with WT males so that resulting fetuses would be heterozygous (*Ifnar1*<sup>+/-</sup>) and thus exhibit a largely intact type I IFN signaling response. In parallel, we developed a second model of ZIKV infection during pregnancy by treating WT pregnant dams with an anti-*Ifnar*-blocking antibody 1 day prior to infection (Figure 1A). Both sets of pregnant mice were inoculated via a subcutaneous route in the footpad with 10<sup>3</sup> focus forming units (FFU) of a clinical isolate from French Polynesia (H/PF/2013) that was passaged in Vero cells. This ZIKV strain is at least 97% identical at the nucleotide level to the sequence of an epidemic strain of ZIKV in Brazil (Calvet et al., 2016; Faria et al., 2016). We confirmed the sequence of our ZIKV H/PF/2013 stock by next-generation sequencing (data not shown), which also allowed us to exclude the presence of adventitious pathogens.

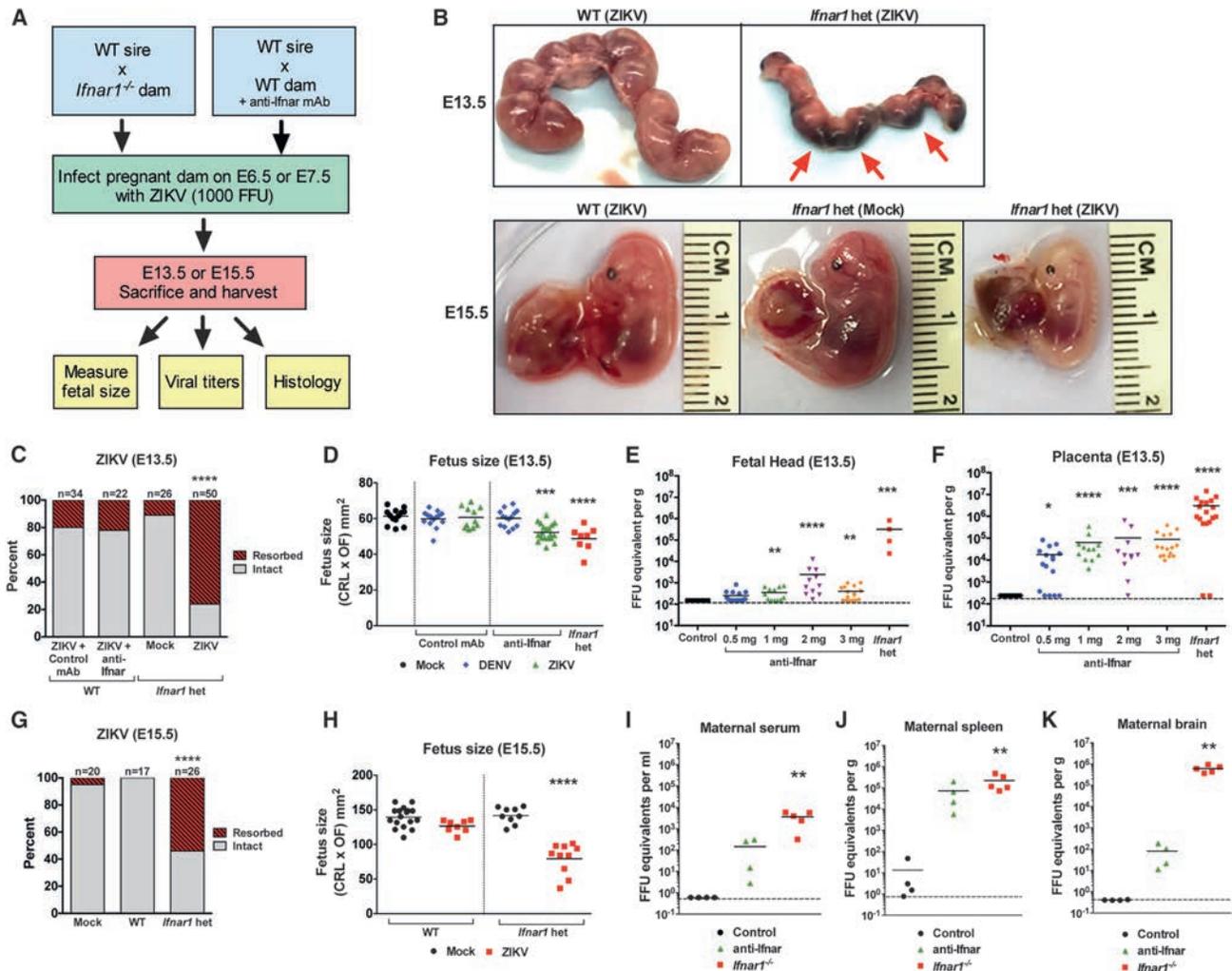
In the *Ifnar1*<sup>-/-</sup> model, pregnant dams mated with WT mice were inoculated on embryonic days 6.5 (E6.5) and E7.5 and sacrificed on E13.5 and E15.5, respectively (Figure 1A). To minimize confounding effects of maternal illness on fetal viability, we evaluated pregnant *Ifnar1*<sup>-/-</sup> mice prior to the onset of disease, which is characterized by hunched posture, fur ruffling, or hind-limb paralysis (Lazear et al., 2016). Individual fetuses were eval-

uated morphologically for size and appearance by measuring the crown-rump length and the occipito-frontal diameter of the fetal head, the latter of which establishes microcephaly in human fetuses (Chervenak et al., 1987; Staples et al., 2016). By E13.5, the majority of ZIKV-infected *Ifnar1*<sup>+/-</sup> heterozygous fetuses had undergone fetal demise and been resorbed, leaving only a placental remnant (Figures 1B, upper, and 1C). The remaining intact *Ifnar1*<sup>+/-</sup> fetuses exhibited significant IUGR (60.2 mm<sup>2</sup> versus 48.7 mm<sup>2</sup>, *p* < 0.0001, Figure 1D). In ZIKV-infected pregnant women, multiple phenotypes have been described including fetal demise, IUGR, and microcephaly (Brasil et al., 2016; Sarno et al., 2016). Although we did not observe isolated microcephaly in this in utero model of ZIKV infection, several other abnormalities were visible in ZIKV-infected *Ifnar1*<sup>+/-</sup> fetuses, including pallor and foci of necrotic tissue in the placenta (Figure 1B).

To determine whether direct infection of the placenta and fetus occurred, we measured ZIKV RNA levels by quantitative real-time RT-PCR (qRT-PCR) as well as infectious virus by plaque assay. High levels of viral RNA and infectious virus were detected within the placenta and also within the fetus head by E13.5 (Figures 1E and 1F and Figures S1A and S1B). As seen with the fetuses from dams infected on E6.5, ZIKV inoculation on E7.5 also resulted in fetal demise and resorption by E15.5 as well as growth restriction (141.8 mm<sup>2</sup> versus 79.5 mm<sup>2</sup>, *p* < 0.0001, Figure 1H) and pallor (Figure 1B) of intact fetuses. As expected from prior studies with *Ifnar1*<sup>-/-</sup> males (Lazear et al., 2016; Rossi et al., 2016), high levels of ZIKV were present in the blood, spleen, and brain of *Ifnar1*<sup>-/-</sup> dams at day 7 after infection (Figures 1I–1K). Of note, the amount of ZIKV RNA within the placenta was ~1,000-fold greater than in maternal serum (Figures 1F and 1I), suggesting that ZIKV replicates preferentially within this tissue.

In our second model of ZIKV infection during pregnancy, WT mice were treated with MAR1-5A3, a blocking anti-*Ifnar* monoclonal antibody (Sheehan et al., 2006), on E5.5, inoculated with ZIKV on E6.5 or E7.5, and fetuses were analyzed on E13.5 or E15.5, respectively (Figure 1A). Although demise was not observed, fetuses exhibited evidence of IUGR compared to control mAb-treated and mock-infected animals (62.3 mm<sup>2</sup> versus 50.2 mm<sup>2</sup>, *p* < 0.005), albeit to a lesser extent than seen in *Ifnar1*<sup>+/-</sup> animals (Figure 1D). In contrast, anti-*Ifnar* mAb-treated mice inoculated subcutaneously with 10<sup>3</sup> FFU of a clinical DENV serotype 3 (DENV-3) isolate that replicates in mice (Pinto et al., 2015; Sarathy et al., 2015) did not exhibit evidence of placental or fetal infection by qRT-PCR or signs of IUGR (Figure 1D and data not shown). These results suggest that ZIKV may have greater tropism for placental cells than other flaviviruses.

The levels of ZIKV RNA detected in WT fetuses were affected by the dose of anti-*Ifnar* mAb administered, with the greatest amounts of ZIKV RNA present in fetuses receiving 2 or 3 mgs of anti-*Ifnar* mAb (Figure 1E). ZIKV RNA persisted in the anti-*Ifnar* mAb-treated fetal heads and bodies at least through E16.5 (Figures S1C and S1D), a critical time in early development of the mouse brain. The placentas in both the *Ifnar1*<sup>-/-</sup> and anti-*Ifnar* antibody models exhibited higher levels of infection than the fetal tissues, and ZIKV RNA accumulation in the placenta was



**Figure 1. Mortality, Viral Burden, and Size of Mouse Fetuses after Maternal Infection with ZIKV**

(A) Schematic depiction of two models of infection during pregnancy. Model 1: WT males were crossed with *Ifnar1*<sup>-/-</sup> dams. Pregnant dams were infected subcutaneously with ZIKV ( $10^3$  FFU) on E6.5 or E7.5 followed by harvest on E13.5, or 15.5, respectively. Model 2: WT males were crossed with WT dams. Pregnant dams treated with 1 mg of an anti-*Ifnar* antibody on days -1, +1, and +3 relative to ZIKV ( $10^3$  FFU) or DENV-3 ( $10^3$  FFU) infection. Mice were sacrificed on E13.5 or E15.5 and fetuses and placentas were harvested for measurements of fetal size by crown-rump length (CRL) and occipito-frontal (OF) diameter.

(B) E13.5 uteri from ZIKV-infected WT and *Ifnar1*<sup>-/-</sup> dams. Most *Ifnar1*<sup>-/-</sup> fetuses carried by *Ifnar1*<sup>-/-</sup> dams died in utero and had undergone resorption, leaving only the residual placenta. In the lower three panels are representative images of fetuses carried by ZIKV-infected WT and mock-infected *Ifnar1*<sup>-/-</sup> dams, the latter of which exhibited growth restriction at E15.5.

(C) Fetus survival on E13.5 after infection with ZIKV on E6.5. Mice were either treated with three 1 mg doses of control or anti-*Ifnar* antibody (left two bars) or untreated mock- or ZIKV-infected *Ifnar1*<sup>-/-</sup> dams (right two bars). Data are representative of at least 3 independent experiments with 1 pregnant female dam per experiment. The n for each group is indicated above each bar. \*\*\*\*p < 0.0001.

(D) Fetus size as assessed by CRL x OF diameter in E13.5 fetuses following E6.5 infection of the indicated pregnant dams with either ZIKV or DENV-3. For these experiments, a total dose of 3 mg of anti-*Ifnar* antibody was used. Bars indicate the mean size of 8-20 fetuses from 2 or 3 independent experiments from fetuses carried by 2 to 3 pregnant dams. \*\*\*p < 0.0005; \*\*\*\*p < 0.0001.

(E and F) Viral burden was measured by qRT-PCR assay from the fetal head and placenta on E13.5 after infection at E6.5. Symbols represent individual fetuses pooled from several independent experiments with the exception of 4 intact *Ifnar1*<sup>+/-</sup> fetal heads that were carried by a single dam. Bars indicate the mean of 4 to 17 mice per group. Dotted lines represent the limit of sensitivity of the assay. \*p < 0.05 \*\*p < 0.005; \*\*\*p < 0.0005; \*\*\*\*p < 0.0001.

(G) Fetus survival on E15.5 after infection with ZIKV on E7.5. Data are representative of at least 2 independent experiments with 1 pregnant female dam per experiment. The n for each group is indicated above each bar. \*\*\*\*p < 0.0001.

(H) Fetus size as assessed by CRL x OF diameter in E15.5 fetuses following E7.5 infection of the indicated pregnant dams with ZIKV. \*\*\*\*p < 0.0001.

(I-K) Viral burden was measured by qRT-PCR assay from maternal serum, spleen, and brain at E13.5. Symbols are derived from individual animals and pooled from 2 or 3 independent experiments. Bars indicate the mean of 4 to 5 mice per group. Dotted lines represent the limit of sensitivity of the assay. See also Figure S1.

independent of the anti-*ifnar* mAb dose above 0.5 mg (Figure 1F). In comparison, mice treated with the isotype control antibody sustained low levels or no detectable ZIKV infection in the placenta, fetal heads or maternal tissues (Figure 1E, 1F, and 1I–1K). Collectively, these data suggest that the mouse placenta is vulnerable to infection with ZIKV, and that high-grade infection may cause placental insufficiency, IUGR, and fetal demise, at least in *Ifnar1*<sup>+/-</sup> animals. Anti-*ifnar* mAb-treated animals sustained less infection and no enhanced lethality although a mild IUGR phenotype was observed.

We evaluated ZIKV localization in the placenta to define whether transmission occurred by a trans-placental route. The mouse placenta is comprised of the maternal decidua and the fetal embryo-derived compartments, including the junctional and labyrinth zones (Figure 2A). Different types of trophoblasts with distinct functions reside within all three layers, including trophoblast giant cells, glycogen trophoblasts, and spongiotrophoblasts. Within the labyrinth zone, fetal capillaries are lined by fetal blood vessel endothelium, which are separated from maternal sinusoids by a layer of mononuclear trophoblasts and a syncytiotrophoblast bilayer (Figure 2A) (Simmons and Cross, 2005; Watson and Cross, 2005). We performed RNA fluorescence in situ hybridization (FISH) coupled with histopathological analysis in ZIKV-infected *Ifnar1*<sup>+/-</sup> placentas and confirmed the presence of ZIKV RNA in different trophoblast cells, including glycogen trophoblasts and spongiotrophoblasts (Figure 2B) and to a lesser extent in mononuclear trophoblasts and syncytiotrophoblasts (data not shown). These findings are consistent with cell culture studies demonstrating ZIKV infection of human trophoblast cell lines (Bayer et al., 2016) and suggest that the mouse model of infection during pregnancy recapitulates features of human disease including placental tropism of ZIKV. We independently confirmed ZIKV infection and replication in two of three human trophoblast cell lines (Figure S2). Transmission electron microscopy of placentas revealed multiple 50 nm dense bodies within the endoplasmic reticulum of the mononuclear trophoblasts (Figure 2C, left), consistent with ZIKV infection of the maternal placenta. As these bodies resemble flavivirus virions (Allison et al., 2003) and were not present in uninfected animals, they are suggestive for the presence of virus. Proximity to non-nucleated maternal erythrocytes (Figure 2C, left) confirmed the location as within the maternal face of the placenta. Consistent with a trans-placental route of infection, we also observed bodies resembling virions within the endoplasmic reticulum of fetal endothelial cells that lined damaged fetal capillaries (Figure 2C, right). The cellular and ultrastructural evidence of ZIKV infection in trophoblasts and fetal endothelium suggests that maternal viremia leads to compromise of the placental barrier by infecting fetal trophoblasts and entering the fetal circulation.

Pathological analysis of ZIKV-infected *Ifnar1*<sup>-/-</sup> (maternal) and *Ifnar1*<sup>+/-</sup> (fetal) placentas showed severe vascular injury characterized by irregularly shaped, reduced fetal capillaries and destruction of the placental microvasculature (Figures 3A and 3B, *Ifnar* het severe). Infected *Ifnar1*<sup>-/-</sup> placentas were smaller, mostly because the labyrinth zone was markedly thinned. In addition, apoptotic trophoblasts were evident in ZIKV-infected placentas (Figure 3A, black arrows). Immunofluorescence stain-

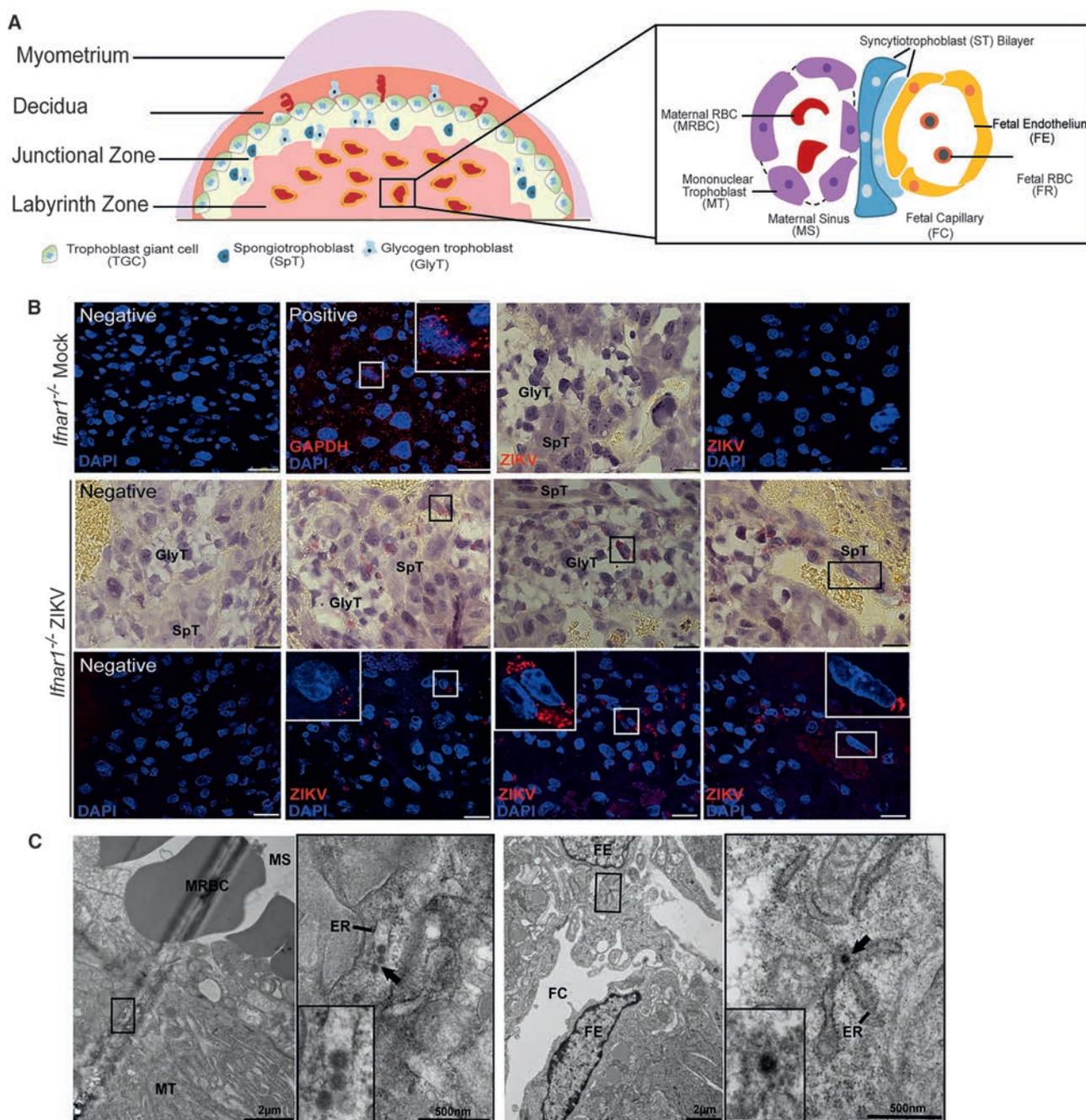
ing of pan-cytokeratin, a pan-trophoblast marker, was diminished in infected *Ifnar1*<sup>-/-</sup> placentas, consistent with evidence of apoptotic trophoblasts (Figure 3B). Apoptosis in trophoblasts can cause disruption of the placental barrier, which compromises protection against pathogens (Robbins and Bakardjiev, 2012). Indeed, ZIKV-infected *Ifnar1*<sup>+/-</sup> placentas contained large numbers of nucleated fetal erythrocytes (Figure 3A, blue arrows), key indicators of fetal stress. Evidence of vascular damage and fewer blood vessels also was reflected by diminished staining of vimentin, a marker of fetal blood vessels in mouse placentas (Figure 3B).

Histopathological assessment of ZIKV-infected *Ifnar1*<sup>+/-</sup> fetal brains demonstrated abundant apoptotic cells within multiple regions at E13.5 (Figures 4A–4D). Activated caspase-3 staining showed low levels of physiological apoptosis in uninfected fetuses (Figures 4E–4H), whereas infected animals had apoptotic cells throughout the midbrain and hindbrain (Figures 4B–4D and 4I). Although we could localize viral RNA in infected placentas, multiple attempts at RNA FISH staining of ZIKV-infected fetal brains did not yield a clear pattern of viral RNA expression (data not shown), despite the recovery of infectious virus (Figures S1A and S1B). Accordingly, we cannot state with certainty whether the enhanced apoptosis within ZIKV-infected fetuses results from infection-induced apoptosis or another process, including ischemia due to placental insufficiency. The presence of numerous apoptotic cells within the developing central nervous system (CNS) coupled with the established neurotropism of ZIKV (Lazear et al., 2016), however, suggests direct infection may contribute.

## DISCUSSION

Epidemiological studies have found that ZIKV infection during pregnancy causes catastrophic neurodevelopmental outcomes in human fetuses, but there currently is no effective treatment or prevention of ZIKV infection other than avoidance of its mosquito vectors. Given the devastating effects of this rapidly emerging infectious disease, small animal models of ZIKV infection during pregnancy are urgently needed to test candidate therapeutics and vaccines that could prevent or mitigate intrauterine infection with ZIKV. We developed two mouse models that support ZIKV replication and trans-placental transmission in pregnant dams: (1) a model of severe disease in pregnant *Ifnar1*<sup>-/-</sup> dams that resulted in fetal demise; and (2) a less severe model of ZIKV pathogenesis in utero using pregnant WT dams that were given anti-*ifnar* antibody prior to and during infection, which resulted in mild IUGR and viral infection within the fetal head during a key period in neurodevelopment.

The placenta acts as a barrier against infections, due to multiple unique structural, cellular, and immune properties. The detrimental effects of congenital viruses on pregnancy and fetal outcomes occur in part because of impaired trophoblast function (Arechavaleta-Velasco et al., 2002). Defective placentas can lead to severe maternal and fetal morbidity and mortality during pregnancy, including spontaneous abortion, stillbirth, preterm birth, IUGR, and other complications. We observed profound pathological changes in ZIKV-infected placentas, including



**Figure 2. ZIKV Infects Trophoblasts and Endothelial Cells within Mouse Placentas**

Pregnant *Ifnar1*<sup>-/-</sup> dams were infected on E7.5 with  $10^3$  FFU of ZIKV via a subcutaneous route and placentas were harvested on E15.5 for histological analysis.

(A) Schematic representation of the mouse placental structure.

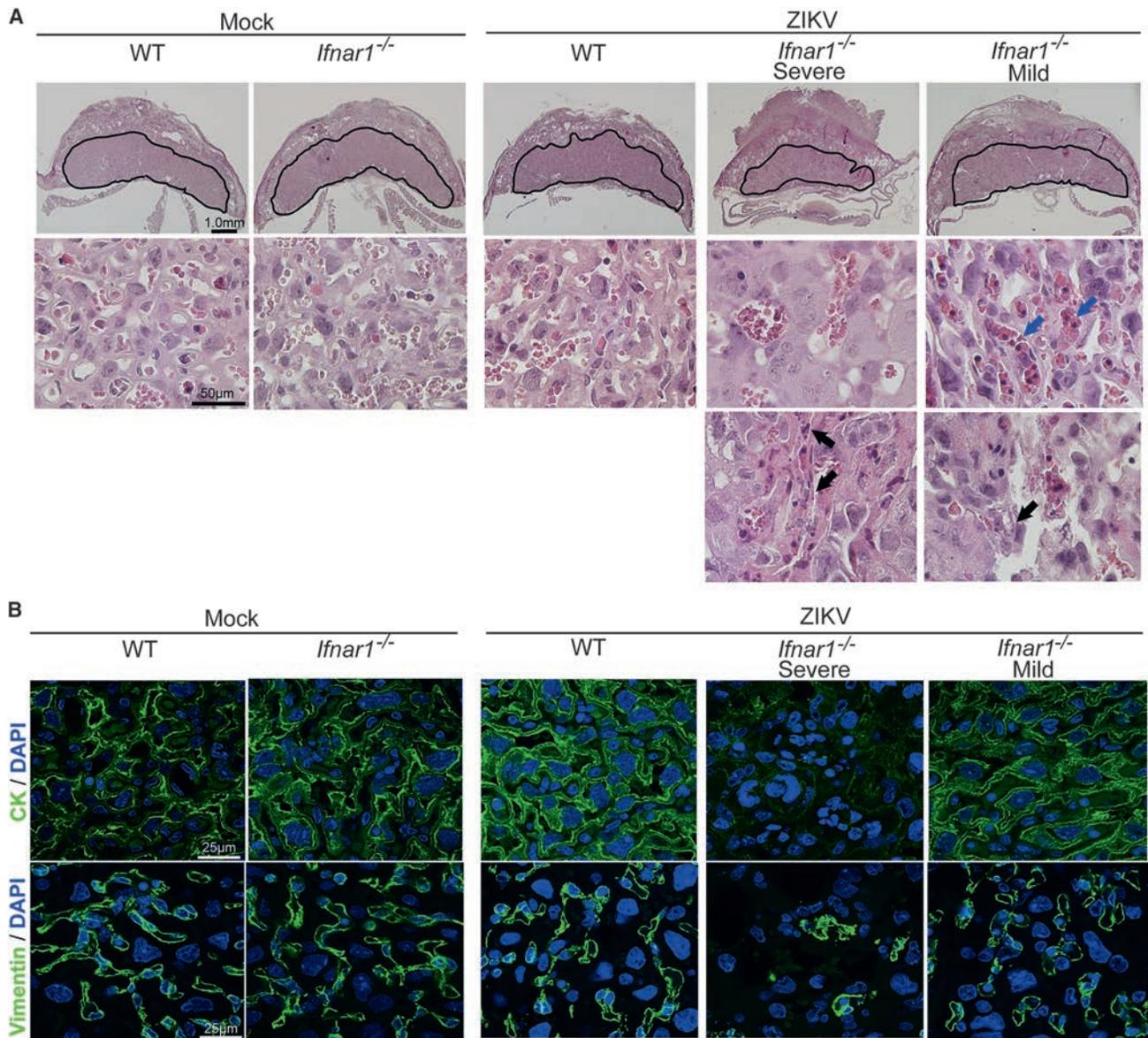
(B) Representative RNA FISH images in uninfected and infected *Ifnar1*<sup>-/-</sup> placentas. Images in each column correspond to the same field of view generated under bright-field or confocal microscopy. Higher magnification of images is displayed as inserts. Scale bar, 25  $\mu$ m.

(C) Transmission electron microscopy images of ZIKV infected *Ifnar1*<sup>-/-</sup> placentas. ZIKV particles were identified within the endoplasmic reticulum in the maternal sinus (left), and in the fetal endothelium (right) lining fetal capillaries in the labyrinth layer.

MT = Mononuclear trophoblast; MS = Maternal sinus; FE = Fetal endothelial cell; FC = Fetal capillary; MRBC = Maternal erythrocyte; and ER = Endoplasmic reticulum. See also Figure S2.

trophoblast apoptosis, abnormal fetal capillary features, and increased fetal nucleated erythrocytes, indicating malfunction of mouse placentas caused by ZIKV infection.

We observed variability in susceptibility to ZIKV infection of different human trophoblast cell lines. Trophoblast cell lines (JEG-3 and HTR-8) originally cultured from choriocarcinoma



### Figure 3. ZIKV Infection Triggers Apoptosis and Vascular Damage in the Placenta

Pregnant dams were infected on E7.5 with  $10^3$  FFU of ZIKV via a subcutaneous route and placentas were harvested on E15.5 for histological analysis.

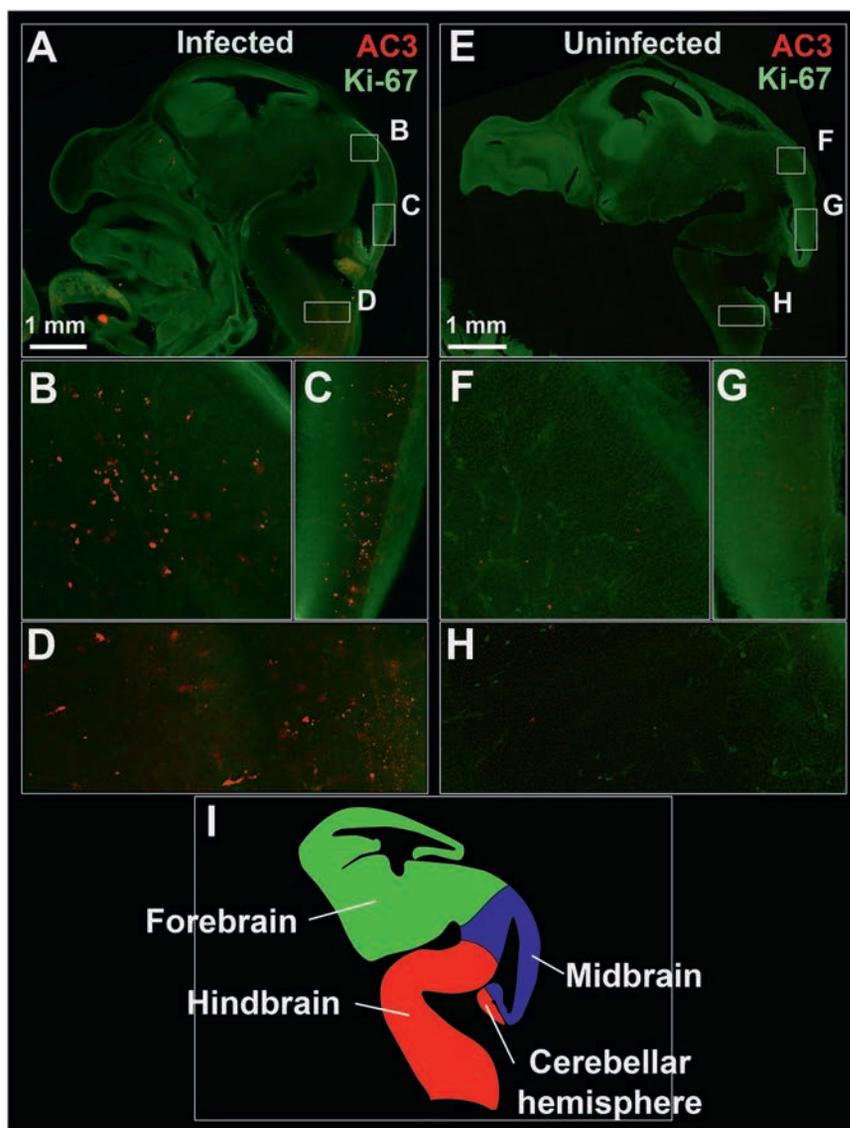
(A) Representative hematoxylin and eosin staining showed pathological features of placentas at E15.5. Labyrinth layers were marked with a solid line on the cross section of mouse placentas. Black arrows indicate apoptotic trophoblasts. Blue arrows indicate increased number of nucleated fetal erythrocytes in fetal capillaries.

(B) Immunofluorescence staining of cytokeratin (CK) and vimentin in mouse placentas. CK, a marker for trophoblasts; vimentin, a marker for the endothelium in fetal capillaries.

See also Figure S2.

explants and first trimester human villous explants, respectively, which exhibit features of extravillous trophoblasts (EVTs) including high invasive capacity and expression of HLA-G, a MHC class II molecule, were susceptible to ZIKV infection. In contrast, a relatively undifferentiated cytotrophoblast cell line (BeWo) was not. Previous studies have shown that EVT are most susceptible to bacterial infections, particularly during the first and second trimesters (Cao and Mysor-

ekar, 2014; Robbins and Bakardjiev, 2012; Zeldovich and Bakardjiev, 2012). Thus, it is possible that in early pregnancy, ZIKV infects EVT and enters the fetal circulation. Placentas nearer to term, which have reduced EVT on the tips of anchoring villi and a more fully developed placental barrier, in general exhibit greater resistance to infection. Indeed, human primary trophoblasts of the villous syncytiotrophoblast phenotype from term placentas were resistant to ZIKV



**Figure 4. ZIKV Infection Is Associated with Evidence of Apoptosis in the Fetal Brain**

Pregnant *Ifnar1*<sup>-/-</sup> dams were infected with 10<sup>3</sup> FFU of ZIKV via a subcutaneous route. Infected (left) or uninfected (right) *Ifnar1*<sup>+/-</sup> E13.5 fetuses were stained with the apoptotic marker activated caspase-3 (AC3; red) and the proliferative marker Ki-67 (green).

(A and E) Sagittal images of representative infected (A) and uninfected (E) fetal heads showing high expression of Ki-67 along brain regions adjacent to ventricles indicative of proliferating neural progenitor cells in the neuroepithelium.

(B-H) Lettered box regions (B-D and F-H) in these images are magnified in corresponding panels below. Higher levels of apoptosis can be seen in the midbrain (B-C) and hindbrain (D) of the infected *Ifnar1*<sup>+/-</sup> fetus. Alternatively, low levels of physiological apoptosis are seen in the absence of infection (F-H).

(I) Diagram depicting the developing E13.5 fetal brain in sagittal view including the forebrain (green), midbrain (blue), and hindbrain (red).

Images are representative of 4-5 sections per fetus from 2 fetuses.

infection due to the production of IFN- $\lambda$  in paracrine manner (Bayer et al., 2016).

Several aspects of our models of ZIKV infection during pregnancy resemble intrauterine infection by ZIKV in humans. Common features included tropism of ZIKV for the placenta, evidence of intrauterine infection, and fetal demise. However, infection during pregnancy in mice did not recapitulate all aspects of human disease, as we did not detect microcephaly, brain calcifications, or absence of individual brain structures, such as the corpus callosum. There are several reasons why ZIKV may not have induced these pathological manifestations in our models. In mice, brain neurogenesis begins around E10 (Finlay and Darlington, 1995), and the brain of a newborn pup is relatively immature at postnatal day 1, akin to the developmental stage of the human brain at mid-gestation (Semple et al., 2013). As the development of the mouse brain includes a major postnatal component, examination of the neurodevelopmental effects of ZIKV

were treated with an anti-*ifnar* antibody, we observed only mild growth restriction in the developing fetus although ZIKV RNA was detectable in the fetal head at both E13.5 and E16.5 after infection. Future behavioral studies may define whether intrauterine infection by ZIKV has long-term neurological effects in mice and serves as a model for evaluation of disease in humans (Staples et al., 2016).

We found that the murine placentas can be infected by ZIKV. Infection of *Ifnar1*<sup>-/-</sup> dams led to severe placental damage and destruction of the microvasculature, which most likely limited blood flow to the developing fetus and caused severe IUGR, ischemia, and fetal demise. In vivo infection of the mouse placenta with ZIKV may provide a model for defining host factors required for or that restrict infection, which could suggest a path for developing therapies to limit placental and intrauterine infection. For example, since IFN- $\lambda$  restricts ZIKV replication within human trophoblasts from term placentas

(Bayer et al., 2016), studies are planned to test the effects of exogenously administered IFN- $\lambda$  on in utero transmission in mice.

Maternal-fetal transmission of pathogens can be mediated by diverse pathways, with the most common being via an ascending route and hemochorial transmission. Viruses from the urogenital tract can disseminate into intrauterine space and colonize the fetal membrane or the placenta in an ascending manner (Edwards et al., 2015). In contrast, viruses from the maternal blood circulation can be transported to the fetomaternal space and infect trophoblasts lining the maternal-fetal interface. These trophoblasts include the EVT<sub>s</sub>, which embed in maternal decidua, endovascular extravillous cytotrophoblasts, and villous trophoblasts, which are bathed in maternal blood (Delorme-Axford et al., 2014). Our studies suggest that ZIKV can infect the placenta through blood-placental transmission and bypass the placental barrier to infect the fetus. Nonetheless, ascending infection routes might be important in sexual transmission during pregnancy. ZIKV has been found in human semen (Musso et al., 2015) and mouse testes (Lazear et al., 2016; Rossi et al., 2016) and can be transmitted sexually from male to female in humans (Musso et al., 2015).

Intrauterine infection with flaviviruses may be an underappreciated phenomenon. Prior to the ZIKV epidemic, there were isolated descriptions of trans-placental infection in humans with other flaviviruses including WNV and JEV (Chaturvedi et al., 1980; Nguyen et al., 2002). These reports suggest that sporadic cases of flavivirus-induced miscarriage or fetal demise might have been unrecognized, although this remains speculative. We are currently testing whether additional variations in ZIKV infection (e.g., dose, route of administration, virus strain, time of infection during pregnancy, and time of analysis) in our mouse model of in utero transmission can recapitulate other morphological abnormalities in the CNS that are described in human disease. Of note, intrauterine infection with Saint Louis encephalitis virus, a less well studied flavivirus, caused severe neurological outcomes in mice. In that model, disease depended on the gestational date of infection (Andersen and Hanson, 1970); mice infected early in gestation survived, whereas those inoculated later developed neurological malformations and died as neonates (Andersen and Hanson, 1970; Andersen and Hanson, 1975). In our studies with DENV and an anti-*ifnar*-blocking antibody, inoculation of mice did not result in placental infection, although it is possible that this was due to a diminished ability of DENV to replicate in mice compared to ZIKV. Experiments that test infection of pregnant animals with additional related viruses may clarify whether placental infection is an underappreciated clinical manifestation of flavivirus pathogenesis.

Treatment and prevention of ZIKV infection will likely require small animal models for testing of vaccines and potential therapies. The mouse models described in our study may be relevant to studying mechanisms of pathogenesis and determining whether vaccines given prior to pregnancy can prevent infection in the developing fetus. Mouse models of ZIKV infection during pregnancy also may provide fundamental insights into how the placental barrier prevents viral infection from the developing fetus, and why this process fails in the context of specific path-

ogens. Finally, our animal model of in utero transmission establishes causality of a fetal syndrome associated with ZIKV infection in mice.

## EXPERIMENTAL PROCEDURES

### Ethics Statement

This study was carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocols were approved by the Institutional Animal Care and Use Committee at the Washington University School of Medicine (Assurance number A3381-01). Inoculations were performed under anesthesia induced and maintained with ketamine hydrochloride and xylazine, and all efforts were made to minimize animal suffering.

### Viruses and Titration

Vero cells (African green monkey kidney epithelial cells) were maintained in DMEM supplemented with 5% fetal bovine serum (Omega) and L-glutamine at 37°C with 5% CO<sub>2</sub>. ZIKV strain H/PF/2013 (French Polynesia, 2013) was provided by the Arbovirus Branch of the Centers for Disease Control and Prevention with permission (X. de Lamballerie, Aix Marseille Université) (Baronti et al., 2014). ZIKV stocks were propagated in Vero cells and titrated by focus-forming assay (FFA). Infected cell foci were detected at 48 hr after infection, following fixation with 1% paraformaldehyde and incubation with 500 ng/mL of flavivirus cross-reactive mouse monoclonal antibody E60 (Oliphant et al., 2006) for 2 hr at room temperature. After incubation for 1 hr with a 1:5,000 dilution of horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG (Sigma), foci were detected by addition of TrueBlue substrate (KPL). Foci were analyzed with a CTL Immunospot instrument. Studies with ZIKV were conducted under biosafety level 2 (BSL2) and animal BSL3 (A-BSL3) containment.

### Mouse Experiments

*Ifnar1*<sup>-/-</sup> mice (Hwang et al., 1995) were backcrossed onto a C57BL/6 background. Mice were bred in a specific-pathogen-free facility at Washington University, or purchased (WT animals) from Jackson Laboratories. Mice were set up for timed-matings and at embryonic days E6.5 or E7.5 were inoculated with ZIKV by subcutaneous (footpad) route with 10<sup>3</sup> FFU of ZIKV in 50  $\mu$ L of PBS. Mice were sacrificed at E13.5, E15.5, or E16.5 depending on the experimental design. Placentas and fetuses were harvested from the infected mice. Fetus size was measured as the crown-rump length  $\times$  occipito-frontal diameter of the head. In some experiments, WT mice were treated with indicated doses of an *Ifnar*-blocking mouse mAb (MAR1-5A3) or isotype control mouse mAb (GIR-208) (produced by Leinco Technologies) (Sheehan et al., 2006; Sheehan et al., 2015) by intraperitoneal injection prior to and after ZIKV or DENV-3 infection.

### Measurement of Viral Burden

ZIKV-infected pregnant mice were euthanized on E13.5, E15.5, or E16.5. Placentas and fetal heads were weighed and homogenized with zirconia beads in a MagNA Lyser instrument (Roche Life Science) in 250 or 500  $\mu$ L of PBS. All homogenized tissues from infected animals were stored at -80°C until virus titration. With some samples, viral burden was determined by plaque assay on Vero cells. Samples were thawed, clarified by centrifugation (2,000  $\times$  g at 4°C for 10 min), and then diluted serially prior to infection of Vero cells. Plaque assays were overlaid with low-melt agarose, and 4 days later were fixed with 10% formaldehyde and stained with crystal violet (Brien et al., 2013). Tissue samples and serum from ZIKV-infected mice were extracted with the RNeasy Mini Kit (tissues) or Viral RNA Mini Kit (serum) (QIAGEN). ZIKV RNA levels were determined by one-step quantitative reverse transcriptase PCR (qRT-PCR) on an ABI 7500 Fast Instrument using standard cycling conditions. Viral burden was expressed on a log<sub>10</sub> scale as viral RNA equivalents per g or per mL after comparison with a standard curve produced using serial 10-fold dilutions of ZIKV RNA. A published primer set was used to detect ZIKV RNA (Lanciotti et al., 2008): Fwd, 5'-CCGCTGCCCAACACAAG-3'; Rev, 5'-CCACTAAGGTCTTTTGCAGACAT-3'; Probe, 5'-/56-FAM/AGCCTACCT/ZEN/TGACAAGCAATCAGACTCAA/3IABkFQ/-3', (Integrated DNA Technologies).

### RNA Fluorescence *In Situ* Hybridization

RNA FISH was performed using ViewRNA ISH Tissue 2-Plex Assay kit (Affymetrix) according to the manufacturer's instructions. Formalin-fixed paraffin-embedded tissue sections were hydrated, heat-treated for 10 min at 90°C, and digested with protease for 20 min. Endogenous alkaline phosphatase was inactivated with 0.2 M HCl and 300 mM NaCl at room temperature for 15 min before the probe hybridization. The probe targeting ZIKV RNA was designed and synthesized by Affymetrix and was based on the ZIKV French Polynesian 2013 genomic RNA sequence (accession number SAMN04592777). Positive controls (probe-targeting housekeeping genes) and no-probe negative controls were included in all samples. Tissues were counterstained with DAPI and Gill's hematoxylin and visualized using standard bright-field and confocal microscopes.

### Immunohistochemistry and Immunofluorescence Imaging

Fetuses were removed by Caesarian-section on E13.5, post-fixed in 4% paraformaldehyde, and sectioned at 70  $\mu$ M with a vibratome. Immunohistochemistry was performed by immersing sections in blocking solution and incubating overnight in primary antibodies raised against activated caspase-3 (AC3; Cell Signaling) and Ki-67 (BD-Biosciences). Subsequently, sections were rinsed and reacted with fluorescent secondary antibodies.

Harvested placentas were fixed in 10% neutral buffered formalin (Fisher) at room temperature and embedded in paraffin. At least three placentas from different litters with the indicated treatments were sectioned and stained with hematoxylin and eosin to assess morphology. For immunofluorescence staining of mouse placentas, deparaffinized tissues were blocked in buffer (1% BSA, 0.3% Triton, 1  $\times$  PBS) for 2 hr and incubated with primary antibodies overnight. The following primary antibodies were used: cytokeratin (1:500, rabbit, DAKO Z0622) and vimentin (1:500, rabbit, Abcam ab92547). After rinsing with PBS, secondary antibody conjugated with Alexa 488 (1:500 in PBS) was applied for 1 hr at room temperature. Samples were counterstained with DAPI (4'6'-diamidino-2-phenylindole, 1:1,000 dilution) for 10 min and mounted in Prolong Gold (Life Technologies). RNA FISH imaging of mouse placentas was acquired using a TSC SPE inverted confocal microscope (Leica) using a 63 $\times$  objective. Histology images were captured by use of a Nikon Eclipse microscope equipped with an Olympus DP71 color camera under 20 $\times$  and 40 $\times$  objectives.

### Transmission Electron Microscopy

Mouse placental samples were fixed in 2% paraformaldehyde and 2.5% glutaraldehyde (Polysciences) in 100 mM sodium cacodylate buffer for 1 hr at room temperature and overnight at 4°C, washed in cacodylate buffer, post-fixed in 1% osmium tetroxide (Polysciences) for 1 hr, then rinsed with distilled H<sub>2</sub>O before en bloc staining with 1% aqueous uranyl acetate (Ted Pella) for 1 hr. Samples were dehydrated in a graded series of ethanol and embedded in Eponate 12 resin (Ted Pella). Sections (90 nm thick) were cut with a Leica Ultracut UCT ultramicrotome, stained with uranyl acetate and lead citrate, and imaged on a JEOL 1200 EX transmission electron microscope (JEOL USA).

### mAb Generation

A more complete description of the anti-ZIKV mAbs will be provided in a forthcoming manuscript (E. Fernandez, D. Platt, and M. Diamond, unpublished data). Briefly, to create ZV-2, *Irf3*<sup>-/-</sup> mice were infected and boosted with 10<sup>3</sup> FFU of ZIKV (MR-766 and H/PF/2013, respectively) and given a final intravenous boost with live 10<sup>6</sup> FFU of ZIKV (H/F/2013) 3 days prior to fusion with P3X63 myeloma cells. Hybridomas secreting antibodies that reacted with ZIKV-infected Vero cells were identified by flow cytometry and cloned by limiting dilution. ZV-2 was purified by protein A affinity chromatography.

### Cell Culture and Infection

JEG-3 and BeWo cells were obtained from ATCC and cultured in F12/DMEM media supplemented with 10% FBS (Life Technologies) at 37°C with 5% CO<sub>2</sub>. HTR-8/SVneo cells were provided by C. H. Graham (Queen's University) and maintained in RPMI 1640 media with 5% FBS. For flow cytometry and viral yield assays, trophoblast cell lines were seeded in 24-well plates at 5  $\times$  10<sup>5</sup> cells per well. For immunofluorescence staining, trophoblast cell lines were

cultured in 4-well chamber slides (Millipore). Cells were counted and infected with ZIKV at a MOI 0.1 for 2 hr, washed twice with warm PBS and cultured with fresh medium. At indicated time points after infection, supernatants were harvested for virus titration, and cells were trypsinized for flow cytometry or fixed for immunofluorescence staining using ZV-2, a ZIKV-specific mAb.

### Data Analysis

All data were analyzed with GraphPad Prism software. For viral burden analysis, the log titers and levels of viral RNA were analyzed by the Mann-Whitney test or ANOVA. A p value of < 0.05 indicated statistically significant differences. Significance of the survival rates were assessed by Chi-square test. The morphological measurements were assessed by a Mann-Whitney test or by a 2-way ANOVA.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.05.008>.

### AUTHOR CONTRIBUTIONS

J.J.M., J.G., B.C., K.K.N., R.S.K., I.U.M., and M.S.D. designed experiments. J.J.M., B.C., J.G., A.M.S., O.H.C., E.F., C.G., M.N., and K.K.N. performed the experiments. J.J.M., B.C., I.U.M., and M.S.D. analyzed the data. J.J.M., B.C., I.U.M., and M.S.D. wrote the first draft of the paper; all authors edited the manuscript.

### ACKNOWLEDGMENTS

This work was supported by grants from the NIH (R01 AI073755 and R01 AI104972 to M.S.D., R01 HD052664 to K.K.N., and R01 NS052632 to R.S.K.) and the Intellectual and Developmental Disabilities Research Center at Washington University (NIH/NICHD U54 HD087011). J.J.M., E.F., and D.J.P. were supported by a Rheumatology Research Foundation Scientist Development Award, NIH Pre-doctoral training grant award (T32 AI007163), and the NIH Research Education Program (R25 HG006687), respectively. The work also was supported by a Preventing Prematurity Initiative grant from the Burroughs Wellcome Fund and a Prematurity Research Initiative Investigator award from the March of Dimes (to I.U.M.). We thank Wandy Betty for her TEM assistance, Fredrik Kraus for his expertise in placental pathology, and Justin Richner for advice with establishing the RNA FISH assays. Finally, we acknowledge Xavier de Lamballerie (Emergence des Pathologies Virales, Aix-Marseille Université, Marseille, France) and the European Virus Archive goes Global (EVAg) for consenting to the use of H/PF/2013 ZIKV strain for this study under a material transfer agreement with the EVAg partner, Aix-Marseille Université. M.S.D. is a consultant for Inbios, Visterra, and Takeda Pharmaceuticals and on the Scientific Advisory Boards of Moderna and OraGene.

Received: April 24, 2016

Revised: May 2, 2016

Accepted: May 2, 2016

Published: May 11, 2016

### REFERENCES

- Aliota, M.T., Caine, E.A., Walker, E.C., Larkin, K.E., Camacho, E., and Osorio, J.E. (2016). Characterization of Lethal Zika Virus Infection in AG129 Mice. *PLoS Negl. Trop. Dis.* 10, e0004682.
- Allison, S.L., Tao, Y.J., O'Riordain, G., Mandl, C.W., Harrison, S.C., and Heinz, F.X. (2003). Two distinct size classes of immature and mature subviral particles from tick-borne encephalitis virus. *J. Virol.* 77, 11357–11366.
- Andersen, A.A., and Hanson, R.P. (1970). Experimental transplacental transmission of St. Louis encephalitis virus in mice. *Infect. Immun.* 2, 320–325.
- Andersen, A.A., and Hanson, R.P. (1975). Intrauterine infection of mice with St. Louis encephalitis virus: immunological, physiological, neurological, and behavioral effects on progeny. *Infect. Immun.* 12, 1173–1183.

- Arechavaleta-Velasco, F., Koi, H., Strauss, J.F., 3rd, and Parry, S. (2002). Viral infection of the trophoblast: time to take a serious look at its role in abnormal implantation and placentation? *J. Reprod. Immunol.* *55*, 113–121.
- Baronti, C., Piorkowski, G., Charrel, R.N., Boubis, L., Leparco-Goffart, I., and de Lamballerie, X. (2014). Complete coding sequence of zika virus from a French polynesia outbreak in 2013. *Genome Announc.* *2*, e00500–e00514.
- Bayer, A., Lennemann, N.J., Ouyang, Y., Bramley, J.C., Morosky, S., Marques, E.T., Jr., Cherry, S., Sadovsky, Y., and Coyne, C.B. (2016). Type III Interferons Produced by Human Placental Trophoblasts Confer Protection against Zika Virus Infection. *Cell Host Microbe* *16*, 30100–30107.
- Brasil, P., Pereira, J.P., Jr., Raja Gabaglia, C., Damasceno, L., Wakimoto, M., Ribeiro Nogueira, R.M., Carvalho de Sequeira, P., Machado Siqueira, A., Abreu de Carvalho, L.M., Cotrim da Cunha, D., et al. (2016). Zika Virus Infection in Pregnant Women in Rio de Janeiro - Preliminary Report. *N. Engl. J. Med.* Published online March 4, 2016. <http://dx.doi.org/10.1056/NEJMoa1602412>.
- Brien, J.D., Lazear, H.M., and Diamond, M.S. (2013). Propagation, quantification, detection, and storage of west nile virus. *Current protocols in microbiology* *31*, 15D 13 11–15D 13 18.
- Calvet, G., Aguiar, R.S., Melo, A.S.O., Sampaio, S.A., de Filippis, I., Fabri, A., Araujo, E.S.M., de Sequeira, P.C., de Mendonça, M.C.L., de Oliveira, L., et al. (2016). Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. *Lancet Infect. Dis.* S1473-3099(16) 00095-5.
- Cao, B., and Mysorekar, I.U. (2014). Intracellular bacteria in placental basal plate localize to extravillous trophoblasts. *Placenta* *35*, 139–142.
- Carteaux, G., Maquart, M., Bedet, A., Contou, D., Brugières, P., Fourati, S., Cleret de Langavant, L., de Broucker, T., Brun-Buisson, C., Leparco-Goffart, I., and Mekontso Dessap, A. (2016). Zika Virus Associated with Meningoencephalitis. *N. Engl. J. Med.* *374*, 1595–1596.
- Cauchemez, S., Besnard, M., Bompard, P., Dub, T., Guillemette-Artur, P., Eyrolle-Guignot, D., Salje, H., Van Kerkhove, M.D., Abadie, V., Garel, C., et al. (2016). Association between Zika virus and microcephaly in French Polynesia, 2013–15: a retrospective study. *The Lancet*, *epub ahead of print*.
- Chaturvedi, U.C., Mathur, A., Chandra, A., Das, S.K., Tandon, H.O., and Singh, U.K. (1980). Transplacental infection with Japanese encephalitis virus. *J. Infect. Dis.* *141*, 712–715.
- Chervenak, F.A., Rosenberg, J., Brightman, R.C., Chitkara, U., and Jeanty, P. (1987). A prospective study of the accuracy of ultrasound in predicting fetal microcephaly. *Obstet. Gynecol.* *69*, 908–910.
- Delorme-Axford, E., Sadovsky, Y., and Coyne, C.B. (2014). The Placenta as a Barrier to Viral Infections. *Annu Rev Virol* *1*, 133–146.
- Edwards, M.S., Popek, E.J., Wise, B., Hatzembuehler, L., Arunachalam, A.R., and Hair, A.B. (2015). Ascending in utero herpes simplex virus infection in an initially healthy-appearing premature infant. *Pediatr. Dev. Pathol.* *18*, 155–158.
- Faria, N.R., Azevedo, Rdo.S., Kraemer, M.U.G., Souza, R., Cunha, M.S., Hill, S.C., Thézé, J., Bonsall, M.B., Bowden, T.A., Rissanen, I., et al. (2016). Zika virus in the Americas: Early epidemiological and genetic findings. *Science* *352*, 345–349.
- Finlay, B.L., and Darlington, R.B. (1995). Linked regularities in the development and evolution of mammalian brains. *Science* *268*, 1578–1584.
- Hwang, S.Y., Hertzog, P.J., Holland, K.A., Sumarsono, S.H., Tymms, M.J., Hamilton, J.A., Whitty, G., Bertoncello, I., and Kola, I. (1995). A null mutation in the gene encoding a type I interferon receptor component eliminates anti-proliferative and antiviral responses to interferons alpha and beta and alters macrophage responses. *Proc. Natl. Acad. Sci. USA* *92*, 11284–11288.
- Lancioti, R.S., Kosoy, O.L., Laven, J.J., Velez, J.O., Lambert, A.J., Johnson, A.J., Stanfield, S.M., and Duffy, M.R. (2008). Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg. Infect. Dis.* *14*, 1232–1239.
- Lazear, H.M., and Diamond, M.S. (2016). Zika Virus: New Clinical Syndromes and Its Emergence in the Western Hemisphere. *J. Virol.* *90*, 4864–4875.
- Lazear, H.M., Govero, J., Smith, A.M., Platt, D.J., Fernandez, E., Miner, J.J., and Diamond, M.S. (2016). A Mouse Model of Zika Virus Pathogenesis. *Cell Host Microbe*, S1931-3128(16)30102-0.
- Martines, R.B., Bhatnagar, J., Keating, M.K., Silva-Flannery, L., Muehlenbachs, A., Gary, J., Goldsmith, C., Hale, G., Ritter, J., Rollin, D., et al. (2016). Notes from the Field: Evidence of Zika Virus Infection in Brain and Placental Tissues from Two Congenitally Infected Newborns and Two Fetal Losses - Brazil, 2015. *MMWR Morb. Mortal. wkly. Rep.* *65*, 159–160.
- Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., Kolenc, M., Resman Rus, K., Vesnaver Vipotnik, T., Fabjan Vodusek, V., et al. (2016). Zika Virus Associated with Microcephaly. *N. Engl. J. Med.* *374*, 951–958.
- Musso, D., Roche, C., Robin, E., Nhan, T., Teissier, A., and Cao-Lormeau, V.M. (2015). Potential sexual transmission of Zika virus. *Emerg. Infect. Dis.* *21*, 359–361.
- Nguyen, Q., Morrow, C., Novick, L., Cambareli, C., Olson, B., Aubry, R., Sne-deker, J., Anand, M., Huang, C., Morse, D., et al.; Centers for Disease Control and Prevention (CDC) (2002). Intrauterine West Nile virus infection—New York, 2002. *MMWR Morb. Mortal. wkly. Rep.* *51*, 1135–1136.
- Oehler, E., Watrin, L., Larre, P., Leparco-Goffart, I., Lastère, S., Valour, F., Baudouin, L., Mallet, H., Musso, D., and Ghawche, F. (2014). Zika virus infection complicated by Guillain-Barre syndrome—case report, French Polynesia, December 2013. *Euro Surveill.* *19*, 20720.
- Oliphant, T., Nybakken, G.E., Engle, M., Xu, Q., Nelson, C.A., Sukupolvi-Petty, S., Marri, A., Lachmi, B.-E., Olshevsky, U., Fremont, D.H., et al. (2006). Antibody recognition and neutralization determinants on domains I and II of West Nile Virus envelope protein. *J. Virol.* *80*, 12149–12159.
- Pinto, A.K., Brien, J.D., Lam, C.-Y.K., Johnson, S., Chiang, C., Hiscott, J., Sarathy, V.V., Barrett, A.D., Shresta, S., and Diamond, M.S. (2015). Defining New Therapeutics Using a More Immunocompetent Mouse Model of Antibody-Enhanced Dengue Virus Infection. *MBio* *6*, e01316–e15.
- Rasmussen, S.A., Jamieson, D.J., Honein, M.A., and Petersen, L.R. (2016). Zika Virus and Birth Defects - Reviewing the Evidence for Causality. *N. Engl. J. Med.* Published online April 13, 2016. <http://dx.doi.org/10.1056/NEJMs1604338>.
- Robbins, J.R., and Bakardjiev, A.I. (2012). Pathogens and the placental fortress. *Curr. Opin. Microbiol.* *15*, 36–43.
- Rossi, S.L., Tesh, R.B., Azar, S.R., Muruato, A.E., Hanley, K.A., Auguste, A.J., Langsjoen, R.M., Paessler, S., Vasilakis, N., and Weaver, S.C. (2016). Characterization of a Novel Murine Model to Study Zika Virus. *Am. J. Trop. Med. Hyg.* *16*-0111.
- Sarathy, V.V., White, M., Li, L., Gorder, S.R., Pyles, R.B., Campbell, G.A., Milligan, G.N., Bourne, N., and Barrett, A.D.T. (2015). A lethal murine infection model for dengue virus 3 in AG129 mice deficient in type I and II interferon receptors leads to systemic disease. *J. Virol.* *89*, 1254–1266.
- Sarno, M., Sacramento, G.A., Khouri, R., do Rosário, M.S., Costa, F., Arch-anjo, G., Santos, L.A., Nery, N., Jr., Vasilakis, N., Ko, A.I., and de Almeida, A.R. (2016). Zika Virus Infection and Stillbirths: A Case of Hydrops Fetalis, Hydranencephaly and Fetal Demise. *PLoS Negl. Trop. Dis.* *10*, e0004517.
- Semple, B.D., Blomgren, K., Gimlin, K., Ferriero, D.M., and Noble-Haesslein, L.J. (2013). Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Prog. Neurobiol.* *106*-107, 1–16.
- Sheehan, K.C., Lai, K.S., Dunn, G.P., Bruce, A.T., Diamond, M.S., Heutel, J.D., Dongo-Arthur, C., Carrero, J.A., White, J.M., Hertzog, P.J., et al. (2006). Blocking monoclonal antibodies specific for mouse IFN-alpha/beta receptor subunit 1 (IFNAR-1) from mice immunized by in vivo hydrodynamic transfection. *Journal of interferon & cytokine research* *26*, 804–819.
- Sheehan, K.C., Lazear, H.M., Diamond, M.S., and Schreiber, R.D. (2015). Selective Blockade of Interferon- $\alpha$  and - $\beta$  Reveals Their Non-Redundant Functions in a Mouse Model of West Nile Virus Infection. *PLoS ONE* *10*, e0128636.

- Simmons, D.G., and Cross, J.C. (2005). Determinants of trophoblast lineage and cell subtype specification in the mouse placenta. *Dev. Biol.* 284, 12–24.
- Staples, J.E., Dziuban, E.J., Fischer, M., Cragan, J.D., Rasmussen, S.A., Cannon, M.J., Frey, M.T., Renquist, C.M., Lanciotti, R.S., Muñoz, J.L., et al. (2016). Interim Guidelines for the Evaluation and Testing of Infants with Possible Congenital Zika Virus Infection - United States, 2016. *MMWR Morb. Mortal. wkly. Rep.* 65, 63–67.
- Tang, H., Hammack, C., Ogden, S.C., Wen, Z., Qian, X., Li, Y., Yao, B., Shin, J., Zhang, F., Lee, E.M., et al. (2016). Zika Virus Infects Human Cortical Neural Progenitors and Attenuates Their Growth. *Cell Stem Cell*. Published online March 4, 2016. <http://dx.doi.org/10.1016/j.stem.2016.02.016>.
- Ventura, C.V., Maia, M., Bravo-Filho, V., Góis, A.L., and Belfort, R., Jr. (2016). Zika virus in Brazil and macular atrophy in a child with microcephaly. *Lancet* 387, 228.
- Watson, E.D., and Cross, J.C. (2005). Development of structures and transport functions in the mouse placenta. *Physiology (Bethesda)* 20, 180–193.
- Zeldovich, V.B., and Bakardjiev, A.I. (2012). Host defense and tolerance: unique challenges in the placenta. *PLoS Pathog.* 8, e1002804.

# Microbial Reconstitution Reverses Maternal Diet-Induced Social and Synaptic Deficits in Offspring

Shelly A. Buffington,<sup>1,2</sup> Gonzalo Viana Di Prisco,<sup>1,2</sup> Thomas A. Auchtung,<sup>3,4</sup> Nadim J. Ajami,<sup>3,4</sup> Joseph F. Petrosino,<sup>3,4</sup> and Mauro Costa-Mattioli<sup>1,2,\*</sup>

<sup>1</sup>Department of Neuroscience

<sup>2</sup>Memory and Brain Research Center

<sup>3</sup>Alkek Center for Metagenomics and Microbiome Research

<sup>4</sup>Department of Molecular Virology and Microbiology  
Baylor College of Medicine, Houston, TX 77030, USA

\*Correspondence: costamat@bcm.edu

<http://dx.doi.org/10.1016/j.cell.2016.06.001>

## SUMMARY

Maternal obesity during pregnancy has been associated with increased risk of neurodevelopmental disorders, including autism spectrum disorder (ASD), in offspring. Here, we report that maternal high-fat diet (MHFD) induces a shift in microbial ecology that negatively impacts offspring social behavior. Social deficits and gut microbiota dysbiosis in MHFD offspring are prevented by co-housing with offspring of mothers on a regular diet (MRD) and transferable to germ-free mice. In addition, social interaction induces synaptic potentiation (LTP) in the ventral tegmental area (VTA) of MRD, but not MHFD offspring. Moreover, MHFD offspring had fewer oxytocin immunoreactive neurons in the hypothalamus. Using metagenomics and precision microbiota reconstitution, we identified a single commensal strain that corrects oxytocin levels, LTP, and social deficits in MHFD offspring. Our findings causally link maternal diet, gut microbial imbalance, VTA plasticity, and behavior and suggest that probiotic treatment may relieve specific behavioral abnormalities associated with neurodevelopmental disorders.

## INTRODUCTION

Recent epidemiological evidence suggests that exposure to maternal obesity in utero increases the risk of neurodevelopmental disorders, such as autism spectrum disorder (ASD) in children (Connolly et al., 2016; Krakowiak et al., 2012; Sullivan et al., 2014). Given the increase in the prevalence of obesity (Skinner and Skelton, 2014), it is important to understand the neurobiological mechanism by which maternal obesity affects offspring behavior and brain function.

The amount and type of dietary macronutrients strongly influence the intestinal microbiota (Tremaroli and Bäckhed, 2012), which consists of a vast bacterial community that resides in

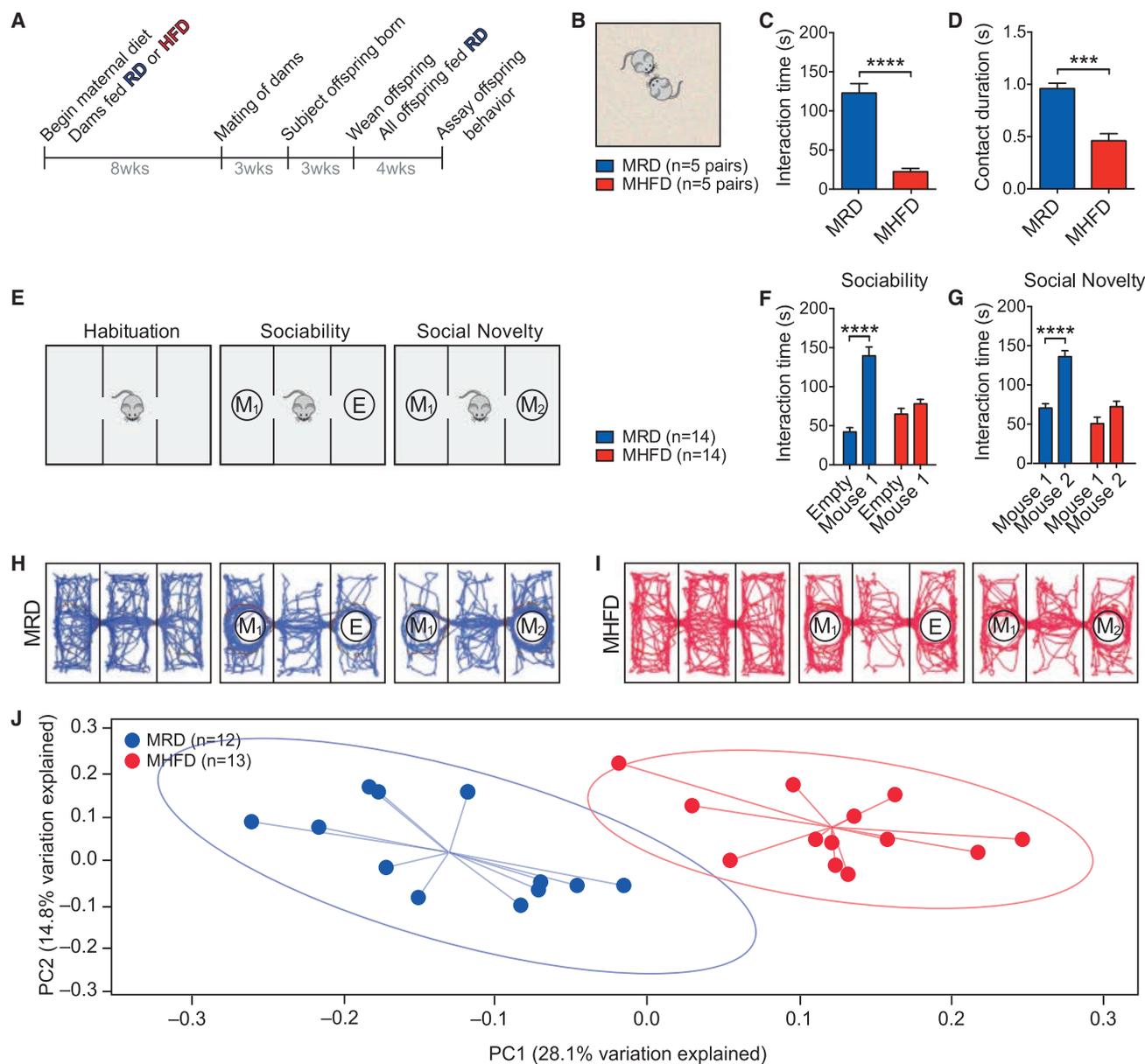
the lower gut and lives in a symbiotic relationship with the host. Maternal obesity has been associated with alterations in the gut microbiome in offspring in both human and non-human primates (Galley et al., 2014; Ma et al., 2014). In addition, some individuals with neurodevelopmental disorders, including ASD, co-present with gastrointestinal problems and dysbiosis of the gut microbiota (Bresnahan et al., 2015; Mayer et al., 2014; Parracho et al., 2005). Given the large body of preclinical literature supporting the notion that a bidirectional communication system between the gut and the brain—known as the gut-brain axis—links gut and brain activities (Cryan and Dinan, 2012; Mayer et al., 2015), it has been speculated that changes in the gut microbiome may be relevant to the development of behavioral symptoms associated with ASD (Hsiao et al., 2013; Mayer et al., 2014). However, how changes in bacteria that inhabit the intestine could influence brain development and function remains unknown.

Here we report that maternal high-fat diet (MHFD)-induced obesity in mice is associated with social behavioral deficits, which are mediated by alterations in the offspring gut microbiome. Notably, we also found that MHFD-induced changes in the offspring gut microbiota block long-lasting neural adaptation in the mesolimbic dopamine reward system (ventral tegmental area [VTA]). Moreover, oral treatment with a single commensal bacterial species corrects oxytocin levels and synaptic dysfunction in the VTA and selectively reverses social deficits in MHFD offspring.

## RESULTS

### Social Behaviors are Impaired in MHFD Offspring

To investigate how maternal diet-induced obesity affects offspring neurodevelopment, female mice were fed either regular diet (RD) or high-fat diet (HFD) for 8 weeks, a standard period required to reach a state of diet-induced obesity in mice (Aye et al., 2015). Females were then paired with males to produce offspring that were given regular diet after weaning (Figure 1A). As expected, MHFD significantly increased maternal weight (Figures S1A–S1C). Consistent with reports of more frequent spontaneous abortion in obese mothers (King, 2006), the litter size



**Figure 1. Social Deficits and Dysbiosis of the Gut Microbiota in MHFD Offspring**

(A) Schematic of the maternal diet regimen and breeding.

(B) Schematic of the reciprocal social interaction task.

(C and D) MHFD offspring showed reduced reciprocal interaction (C,  $p < 0.0001$ ,  $t = 7.90$ ; D,  $p < 0.001$ ,  $t = 5.89$ ).

(E) Schematic of the three-chamber social interaction task.

(F–G) In the sociability test, MRD offspring spent more time interacting with a mouse than with an empty wire cage (F,  $p < 0.0001$ ,  $t = 8.817$ ), whereas MHFD offspring showed no preference for the mouse (F,  $p = 0.48$ ,  $t = 1.19$ ; maternal diet effect  $F_{1,52} = 6.08$ ,  $p < 0.05$ ). In the social novelty test, unlike MRD (G,  $p < 0.0001$ ,  $t = 6.68$ ), MHFD offspring had no preference for interacting with a novel versus a familiar mouse (G,  $p = 0.086$ ,  $t = 2.08$ ; maternal diet effect  $F_{1,52} = 34.96$ ,  $p < 0.0001$ ).

(H and I) Representative exploratory activity of MRD (H) and MHFD (I) offspring in the three-chamber test.

(J) Principal coordinates analysis (PCoA) of unweighted UniFrac distances from the averaged rarefied 16S rRNA gene dataset ( $n = 1,000$  rarefactions; 7,617 reads/sample) showed that MRD samples clustered separately from MHFD samples ( $p < 0.001$ ,  $R^2 = 0.37$ ). Plots show mean  $\pm$  SEM. See also Figures S1 and S2.

was reduced (Figure S1D) and latency to first litter increased in female mice fed HFD (Figure S1E). It is noteworthy that there was no significant difference in offspring weight between maternal diet cohorts at 7–12 weeks of age (Figures S1F and

S1G), the time at which behavioral and electrophysiological experiments were performed.

Given that maternal obesity has been associated with increased risk for neurodevelopmental disorders including ASD in offspring

(Bilder et al., 2013; Krakowiak et al., 2012; Moss and Chugani, 2014) and deficient social interactions are a salient behavioral feature of ASD (Mefford et al., 2012), we studied social behavior in maternal regular diet (MRD) and MHFD offspring. First, we assessed reciprocal social interactions by recording the amount of time a pair of mice, unfamiliar with each other, spent interacting in a neutral arena (Figure 1B). When compared to MRD offspring, MHFD offspring had fewer reciprocal interactions (Figures 1C and 1D, S1H, and S1I). Next, we used the three-chamber test (Silverman et al., 2010) to assess (1) sociability by comparing the time mice spent interacting with an empty wire cage versus one containing a mouse and (2) preference for social novelty by measuring the time mice spent interacting with a familiar versus a stranger mouse (Figure 1E). Consistent with the results from reciprocal social interactions, MHFD offspring had impaired sociability and showed no preference for social novelty (Figures 1F–1I, S1J, and S1K). Taken together these data indicate that MHFD offspring display social deficits.

### Dysbiosis of the Gut Microbiota in MHFD Offspring

A variety of factors could contribute to the etiology of MHFD-induced social behavioral abnormalities. However, maternal obesity has been shown to alter the gut microbiome of offspring (Galley et al., 2014; Ma et al., 2014) and individuals diagnosed with ASD can co-present dysbiosis of the gut microbiota (Bresnahan et al., 2015; Mayer et al., 2014; Parracho et al., 2005). To examine whether MHFD induces alterations in offspring gut microbiota, we analyzed the bacterial composition and community structure in the feces of MRD and MHFD offspring by 16S ribosomal RNA (rRNA) gene sequencing. The microbial communities in both MRD and MHFD offspring were comprised of a typical mouse gut microbiota, dominated by Bacteroidetes and Firmicutes (Figures S2A–S2D). While bacterial diversity computed based on weighted UniFrac distances (the assessment of community structure by considering abundance of operational taxonomic units [OTUs]) did not differ significantly between the offspring from either diet group (Figure S2E), unweighted analyses of UniFrac distances (assessment of community structure by considering only OTU presence/absence) revealed a marked difference between the structures of the bacterial communities (Figure 1J). Moreover, the diversity of microbiota in MHFD offspring was reduced compared to MRD microbiota (Figure S2F).

Consistent with previous reports (Turnbaugh et al., 2006), an HFD regimen in mothers induced a remarkable change in the maternal microbiome composition and diversity (Figures S2G and S2H), which was similar to that observed in their offspring (Figures 1J and S2F).

### Gut Microbiota Mediate MHFD-Induced Social Deficits

While microbial communities vary across individuals (Yatsunen et al., 2012), co-housed family members are known to share their microbiota (Song et al., 2013). Since mice are coprophagic and transfer gut microbiota between each other by the fecal-oral route (Ridaura et al., 2013), we examined whether co-housing MHFD with MRD mice prevents the social deficits in MHFD offspring. To this end, at weaning (3 weeks) an MHFD mouse was co-housed with three MRD mice (Figures 2A and 2B). Control groups consisted of individual cages containing either four MHFD mice or

four MRD mice (Figure 2B). Fecal samples were collected and social behavior in MRD and MHFD offspring was assessed when mice were 7–8 weeks old. Strikingly, MHFD mice co-housed with MRD mice exhibited normal reciprocal social interactions (Figures 2C and 2D and S3A–S3C), as well as normal sociability and preference for social novelty, as determined by the 3-chamber test (Figures 2E, 2F, S3D and S3E). Thus, co-housing with control mice corrects social deficits in MHFD offspring.

We next examined whether co-housing also corrected the changes in the microbiota of MHFD offspring. Indeed, co-housing caused a shift in the bacterial phylogenetic profile of MHFD mice to resemble that of MRD or MRD co-housed mice (Figure 2G), thus correcting the MHFD-induced alterations in the commensal microbiota.

In agreement with the idea that the fecal microbiota of MHFD offspring lacks one or more beneficial bacterial species required for normal social behavior, co-housing one MRD with three MHFD offspring was sufficient to rescue both the social behaviors and microbiota phylogenetic profile of MHFD offspring (Figures S3F–S3L).

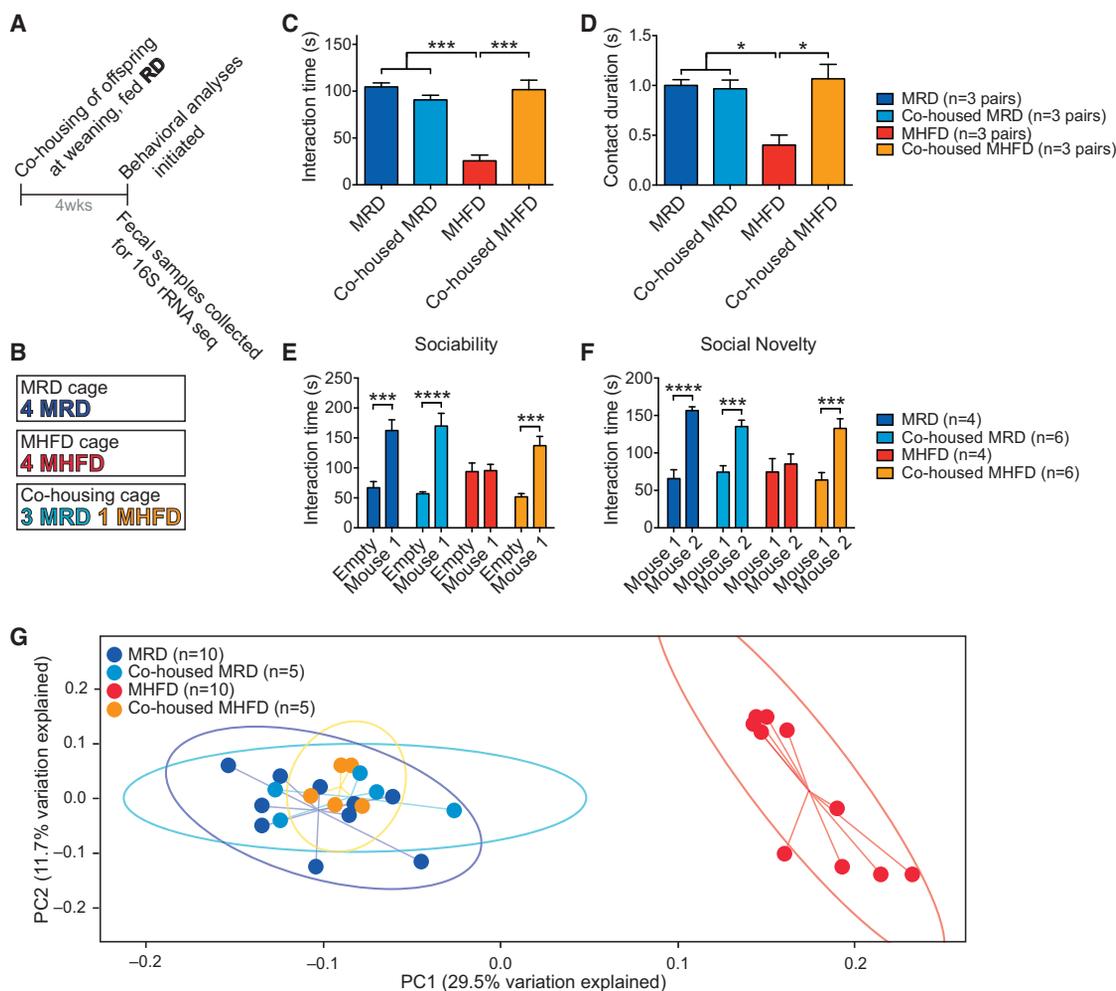
### Colonization of Germ-Free Mice with the Microbiota from MRD, but Not MHFD Mice, Reverses Their Deficient Social Behavior

Studies on germ-free (GF) mice have shown that the intestinal microbiota can influence brain development and function (Cryan and Dinan, 2012). We hypothesized that, if the lack of one or more bacterial species in the microbiota of MHFD offspring is responsible for their defective social behavior, GF mice should also be socially deficient. Confirming this hypothesis and in keeping with recent results (Desbonnet et al., 2014), social behaviors were impaired in GF mice (Figures 3A–3D and S4A–S4D).

To identify functional differences between gut microbial communities and determine whether their role is causal, we transplanted (gavaged) fecal microbiota from adult MRD and MHFD offspring into 4-week- (Figure 3E) and 8-week-old (Figure 3F) GF mice. Interestingly, GF mice that received fecal microbiota from MRD offspring at weaning (4 weeks; Figures 3G, 3H, S4E and S4F), but not at 8 weeks (Figures 3I, 3J, S4G, and S4H), showed normal social behavior. By contrast, GF mice that received fecal microbiota from MHFD offspring remained socially impaired, regardless of the age at which the fecal transfer was performed (Figures 3G–3J and S4E–S4H). Moreover, as in the case of the phylogenetic separation of MRD and MHFD microbiota (Figure 1J), the bacterial communities in GF mice receiving feces from MHFD donor offspring clustered separately from those of GF mice receiving feces from MRD donor offspring, irrespective of whether the fecal transplant was performed at 4 or 8 weeks (Figures 3K, 3L, and S4I–S4N). These data reveal a neurodevelopmental window during which microbial reconstitution effectively improves social behavior.

### MHFD Negatively Impacts a Subset of Bacteria in the Intestinal Gut and Selective Re-introduction of *Lactobacillus (L.) reuteri* Restores Social Deficits in MHFD Offspring

To investigate which bacterial species are absent in the microbiota community of MHFD offspring, we performed metagenomic



**Figure 2. Co-housing MHFD with MRD Offspring Rescues Both Social Dysfunction and the Microbiota Phylogenetic Profile of MHFD Mice**

(A) Schematic of the co-housing experiment.

(B) MRD and MHFD offspring were weaned into one of three cage compositions.

(C–G) Social interaction time (C, MRD versus MHFD  $p < 0.001$ ,  $t = 9.30$ ; MRD versus co-housed MHFD  $p > 0.99$ ,  $t = 0.31$ ; MHFD versus co-housed MHFD  $p < 0.001$ ,  $t = 7.99$ ;  $p < 0.0001$ ,  $F_{3,8} = 30.51$ ) and contact duration (D, MRD versus MHFD  $p < 0.05$ ,  $t = 4.13$ ; MRD versus co-housed MHFD  $p > 0.99$ ,  $t = 0.46$ ; MHFD versus co-housed MHFD  $p < 0.05$ ,  $t = 4.59$ ;  $p < 0.001$ ,  $F_{3,8} = 9.01$ ) in the reciprocal interaction test; social interaction times in the sociability (E, MRD  $p < 0.001$ ,  $t = 4.36$ ; MHFD  $p > 0.99$ ,  $t = 0.078$ ; Co-housed MRD  $p < 0.0001$ ,  $t = 6.33$ ; Co-housed MHFD  $p < 0.001$ ,  $t = 4.78$ ; Maternal diet/Housing/Interaction effect  $F_{3,32} = 6.13$ ,  $p < 0.01$ ) and social novelty tests (F, MRD  $p < 0.0001$ ,  $t = 5.12$ ; MHFD  $p > 0.99$ ,  $t = 0.60$ ; Co-housed MRD  $p < 0.001$ ,  $t = 4.20$ ; Co-housed MHFD  $p < 0.001$ ,  $t = 4.76$ ; maternal diet/housing/interaction effect  $F_{3,32} = 4.37$ ,  $p < 0.01$ ), as well as UniFrac-based phylogenetic clustering (G,  $p < 0.001$ ,  $R^2 = 0.552$ ;  $n = 1,000$  rarefactions; 3,390 reads/sample), are all restored in MHFD offspring co-housed with MRD mice. Plots show mean  $\pm$  SEM. See also Figure S3.

shotgun sequencing of fecal samples from both MHFD and MRD offspring. Our analysis identified several species whose relative abundance was dramatically reduced in the MHFD offspring microbiota (Table 1). Among these, *L. reuteri* was the most drastically reduced (>9-fold) in the MHFD microbiota population, compared to the MRD microbiota (Table 1).

*L. reuteri* has been shown to promote oxytocin levels (Poutahidis et al., 2013), a hormone that plays a crucial role in social behaviors (Donaldson and Young, 2008). We hypothesized that the selective decrease in *L. reuteri* in the microbiota of MHFD offspring was causally related to their social deficits. To test this hypothesis, we introduced *L. reuteri* into the drinking water of MHFD offspring at weaning for 4 weeks, after which behavior

was tested (Figure 4A). Remarkably, treatment with *L. reuteri* significantly improved sociability and preference for social novelty in MHFD offspring (Figures 4B, 4C, 4E, S5A, and S5B). Results from several control experiments underscore the specificity of *L. reuteri*-mediated rescue of social behaviors in MHFD offspring. First, drinking water treated with either resuspension media or heat-killed *L. reuteri* (80°C for 20 min) failed to restore social behavior in MHFD offspring (Figures 4B–4D, S5A, and S5B). Second, drinking water with live *L. reuteri* did not change the social behavior of MRD offspring (Figures 4B, 4C, S5A, and S5B), presumably because their gut microbiota already contains ample *L. reuteri*. Finally, addition of *L. reuteri* to the drinking water had no major effect on bacterial viability and the heat-killing

procedure completely abrogated colony-forming units (Figure S5C). Importantly, the amelioration of the deficient social behavior is specific to *L. reuteri* since similar treatment with another *Lactobacillus* species, *L. johnsonii*, whose abundance is also reduced in the gut microbiota of MHFD offspring (Table 1), failed to rescue social behaviors in MHFD offspring (Figures S5D–S5M).

MHFD offspring also show other behavioral traits that are associated with ASD, like repetitive behaviors and anxiety (Figures S6). Interestingly, while co-housing MHFD with MRD offspring restores social behavior (Figures 2C and 2D), it failed to rescue marble burying (Figure S6B), a behavioral task reflecting repetitive and perseverative behavior (Thomas et al., 2009). Accordingly, GF mice also showed increased marble burying, and fecal microbial transplants from MRD (or MHFD) offspring into GF mice failed to reverse the repetitive behavior (Figure S6B). Thus, these data suggest that repetitive behaviors in MHFD offspring do not depend on changes in the gut microbiome. In addition, *L. reuteri* treatment had no effect on anxiety in MHFD offspring (Figures S6C–S6H). Taken together, these data indicate that *L. reuteri* reconstitution specifically rescues social, but not other behavioral endophenotypes associated with ASD.

### Oxytocin Levels Are Reduced in the Hypothalamus of MHFD Offspring

There is growing evidence that the neuropeptide oxytocin modulates numerous aspects of social behavior (Donaldson and Young, 2008) and is implicated in ASD (Lerer et al., 2008; Wu et al., 2005). *L. reuteri*, which rescued social behaviors in MHFD mice (Figures 4B, 4C, and 4E), has been reported to increase oxytocin levels (Poutahidis et al., 2013). Because oxytocin is primarily synthesized in the paraventricular nuclei (PVN) of the hypothalamus, we decided to compare the number of oxytocin-expressing cells in the PVN of MRD and MHFD offspring. Interestingly, MHFD offspring had significantly fewer oxytocin immunoreactive neurons compared to MRD offspring (Figures 4F–4H, 4K, and 4L). The reduction in oxytocin immunoreactivity was not due to an overall decrease in PVN neurons, since the total number of neurons was unchanged (as measured by NeuN staining; Figures 4G, 4H, 4M, and 4N). Notably, in *L. reuteri*-treated MHFD offspring, the number of oxytocin-expressing cells was higher than in control-treated MHFD offspring (Figures 4I, 4J, 4O, and 4P). Thus, the number of oxytocin immunoreactive neurons in the PVN is reduced in MHFD offspring but can be restored by *L. reuteri* treatment.

### Mesolimbic Dopamine Reward System Function Is Impaired in MHFD Offspring

Brain regions that respond to naturally rewarding stimuli, including the ventral tegmental area (VTA) and the nucleus accumbens (NAc), are crucially involved in social behaviors (Dölen et al., 2013; Gunaydin et al., 2014). In addition, oxytocin-expressing neurons in the PVN project to the VTA (Melis et al., 2007). Oxytocin activates VTA neurons in both mice and humans, influencing the processing of socially relevant cues (Groppe et al., 2013; Tang et al., 2014) and oxytocin receptor blockade in the VTA prevents social attachment in rodents (Pedersen et al., 1994). Given that social stimulation can be particu-

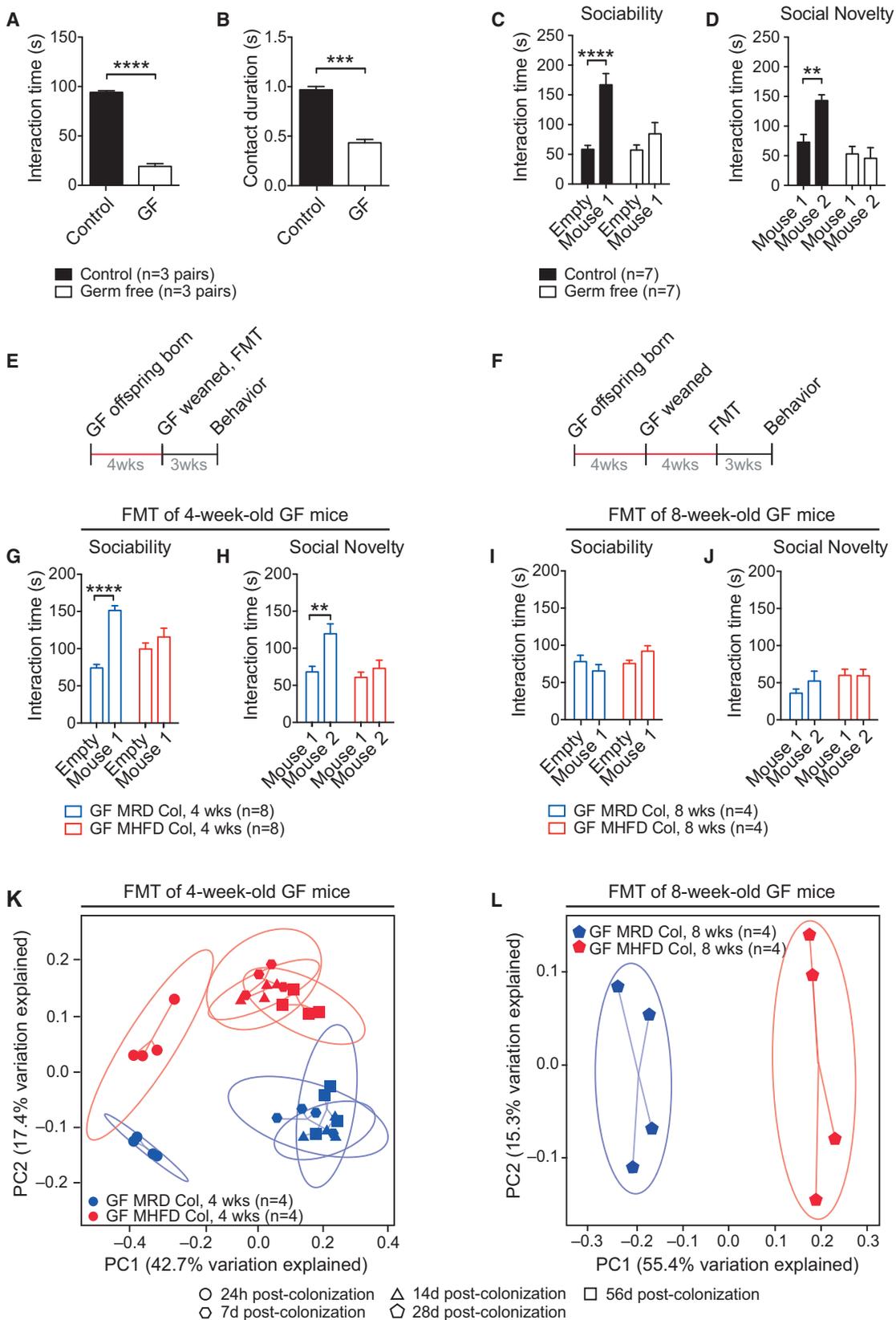
larly rewarding and triggers synaptic potentiation in VTA DA neurons of birds (Huang and Hessler, 2008), we examined whether direct social interaction evokes long-term potentiation (LTP) of synaptic inputs to VTA DA neurons (Figures S7A–S7C). To this end, we recorded AMPAR/NMDAR ratios of glutamatergic excitatory postsynaptic currents (EPSCs) in MRD and MHFD offspring 24 hr following a 10 min reciprocal interaction with either a stranger or a familiar mouse (Figure 5A). In control MRD mice, social interaction with a stranger, but not a familiar mouse, triggered LTP in VTA DA neurons, as determined by an increase in AMPAR/NMDAR ratios (Figures 5B and 5D). By contrast, in MHFD offspring, social interaction with a stranger failed to induce LTP in their VTA DA neurons (Figures 5C and 5E). Input-output curves, paired-pulse ratios and miniature EPSCs (mEPSCs) frequency and amplitude show that the impairment of LTP induced by social interaction in MHFD offspring cannot be attributed to changes in basal synaptic transmission (Figures S7D–S7H).

Mirroring the electrophysiological results, MRD offspring spent significantly more time interacting with a stranger than a familiar mouse, but MHFD offspring did not (Figures 5F and 5G). Thus, social interaction induces a long-lasting increase in the activity of the dopaminergic reward system of MRD, but not in MHFD, offspring.

### Treatment with *L. reuteri* or Oxytocin Reverses Both the LTP in VTA DA Neurons and Social Behavior in MHFD Offspring

We next wondered whether *L. reuteri* treatment, which restores sociability and preference for social novelty in MHFD offspring (Figures 4B, 4C, and 4E), would also rescue reciprocal social interaction and related changes in synaptic strength in the VTA. Live (Figures 5H and 5J), but not heat-killed (Figures 5I and 5K), *L. reuteri* rescued stranger interaction-induced LTP in the VTA as well as reciprocal social interactions in MHFD offspring (Figure 5L). Thus, *L. reuteri* restores social interaction-induced LTP in the VTA of MHFD offspring.

These findings, together with the fact that *L. reuteri* treatment increased oxytocin immunoreactivity in the PVN of MHFD offspring (Figures 4I and 4J, 4O, and 4P), led us to examine whether direct oxytocin application could also reverse the behavioral and electrophysiological deficits characteristic of MHFD offspring. To test this idea, we administered oxytocin intranasally—a preferred method of administering neuropeptides to the brain bypassing more invasive procedures (Peñagarikano et al., 2015)—to MHFD offspring and measured reciprocal social interactions 30 min later. Although either oxytocin alone or social interaction alone failed to rescue social interaction-induced LTP in the VTA, the combination of social interaction and oxytocin treatment restored LTP in the VTA of MHFD offspring (Figures 6A and 6B), supporting prior work implicating a synergistic effect of oxytocin and dopamine in the processing of socially relevant cues (Modi and Young, 2012). Accordingly, oxytocin treatment improved reciprocal social interaction (Figures 6C–6F), as well as sociability and the preference for social novelty (Figures 6G–6J). Thus, oxytocin administration rescues social behavior and related neural adaptations in the VTA of MHFD offspring. Collectively, our data show that MHFD impairs



(legend on next page)

**Table 1. Species Whose Abundance Is Reduced in the Gut Microbiota of MHFD Offspring**

Species of Interest	MRD Representation	MHFD Representation	Fold Change MRD/MHFD
<i>Lactobacillus reuteri</i>	7.49 ± 3.0	0.879 ± 0.21	9.24 ± 0.65
<i>Parabacteroides distasonis</i>	0.00709 ± 0.0055	0.00126 ± 0.0011	5.63 ± 1.17
<i>Helicobacter hepaticus</i>	7.35 ± 2.4	2.58 ± 1.3	2.84 ± 0.61
<i>Bacteroides uniformis</i>	5.49 ± 2.2	2.07 ± 0.78	2.65 ± 0.56
<i>Olsenella unclassified</i>	0.230 ± 0.064	0.121 ± 0.031	1.90 ± 0.38
<i>Collinsella unclassified</i>	0.0866 ± 0.031	0.0494 ± 0.016	1.75 ± 0.48
<i>Bifidobacterium pseudolongum</i>	19.4 ± 3.3	11.3 ± 2.4	1.71 ± 0.27
<i>Lactobacillus johnsonii</i>	24.5 ± 6.2	17.1 ± 5.2	1.43 ± 0.40

oxytocin-mediated synaptic adaptations in the VTA that underlie social behaviors.

## DISCUSSION

Both genetic and environmental factors, and their interactions, play a crucial role in the etiology of neurodevelopmental disorders including ASD (Hallmayer et al., 2011). There is growing epidemiological evidence that maternal obesity heightens the risk of neuropsychiatric disorders in offspring (Krakowiak et al., 2012; Sullivan et al., 2014). Indeed, a recent study reported that mothers with obesity were 1.5 times more likely to have a child with ASD, and the increased risk of children with ASD was two-fold greater for pregnant mothers with both obesity and gestational diabetes (Connolly et al., 2016).

While most of the focus in the field has been on inflammation (Bolton and Bilbo, 2014) or epigenetic changes (Mathers and McKay, 2009), the biological mechanism by which maternal obesity affects offspring neurodevelopment remains to be determined. Here, we show that the behavioral dysfunction associated with MHFD-induced obesity is induced by alterations in the offspring gut microbiota. Several lines of evidence support this idea. First, some individuals diagnosed with ASD present dysbiosis of the gut microbiota and gastrointestinal issues (Bresnahan et al., 2015; Mayer et al., 2014; Parracho et al., 2005). Second, maternal obesity leads to alterations in the offspring's gut microbiome in humans and non-human primates (Galley et al., 2014; Ma et al., 2014). Third, in mice, the gut microbiota of MHFD offspring is altered (Figure 1J) by the reduction in specific bacterial species (Table 1). Fourth, manipulation of the microbiome community by co-housing MHFD with MRD offspring rescues MHFD-induced social deficits and corrects their microbial

phylogenetic profile (Figures 2 and S3). Fifth, GF mice are socially impaired and fecal microbiota transplanted from MRD (but not MHFD) offspring rescues GF social behavior (Figures 3 and S4). Finally, treatment with a single bacterial species, *L. reuteri*, which is dramatically reduced in MHFD offspring (Table 1), selectively restores social behavior in MHFD mice (Figures 4 and S5A–S5C).

We propose a model in which *L. reuteri* improves social behavior by promoting oxytocin-mediated functions. Consistent with this model, *L. reuteri*-treatment enhances oxytocin levels in the PVN of MHFD mice (Figures 4I and 4J) and direct oxytocin-treatment normalizes the social behavior of MHFD offspring (Figure 6). Although the precise mechanism by which *L. reuteri* promotes oxytocin in the brain remains to be determined, we favor the idea that the vagus nerve (Davari et al., 2013) could be the main pathway of communication between the gut/*L. reuteri* and changes in oxytocin in the PVN. It is known that vagal nerve fibers project to the PVN (Sabatier et al., 2013; Uvnäs-Moberg et al., 2014). In addition, neuronal activity in the PVN induced by bacterial colonization is blocked by subdiaphragmatic vagotomy (Wang et al., 2002). Especially relevant are the reports that the *L. reuteri*-mediated increase in oxytocin depends on the vagus nerve (Poutahidis et al., 2013) and that another *Lactobacillus* species, *L. rhamnosus*, reduced stress-induced anxiety in mice in a vagus-dependent manner (Bravo et al., 2011).

Our results provide new insight into the mechanism by which a marked shift in microbial ecology, caused by MHFD, can negatively impact social behaviors and related neuronal changes in offspring. These neuronal adaptations, which underlie social behavior by enhancing the salience and rewarding value of social stimuli, are surprisingly impaired by maternal diet-induced

### Figure 3. Fecal Microbiota from MRD, but not MHFD, Offspring Improves Germ-Free (GF) Recipient Social Behavior

(A–D) GF mice show reduced reciprocal social interaction (A,  $p < 0.0001$ ,  $t = 22.73$ ; B,  $p < 0.001$ ,  $t = 11.31$ ) and deficits in sociability (C, Control  $p < 0.0001$ ,  $t = 5.30$ , GF  $p > 0.99$ ,  $t = 0.39$ ; main group effect  $F_{1,24} = 21.98$ ,  $p < 0.0001$ ) and preference for social novelty (D, control  $p < 0.01$ ,  $t = 3.64$ , GF  $p = 0.39$ ,  $t = 1.33$ ; main group effect  $F_{1,24} = 5.29$ ,  $p < 0.05$ ).

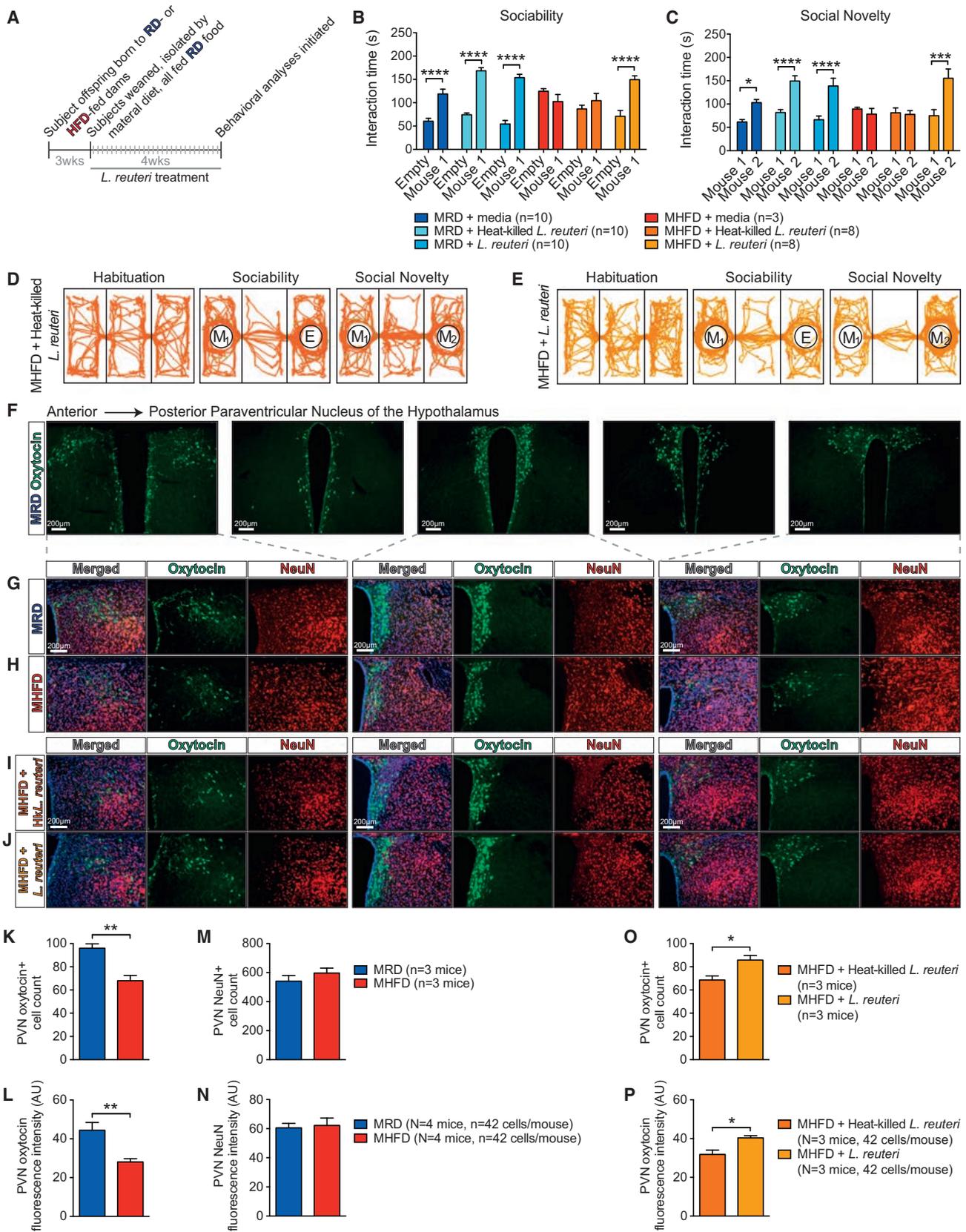
(E and F) Schematic of fecal microbiota transplant (FMT) at 4 (E) and 8 weeks of age (F).

(G and H) FMT from MRD, but not MHFD, offspring at weaning restored both GF sociability (G,  $GF_{MRDCol} p < 0.0001$ ,  $t = 6.66$ ;  $GF_{MHFDCol} p = 0.35$ ,  $t = 1.40$ ; Donor effect  $F_{1,28} = 32.44$ ,  $p < 0.0001$ ) and preference for social novelty (H,  $GF_{MRDCol} p < 0.01$ ,  $t = 3.60$ ;  $GF_{MHFDCol} p = 0.81$ ,  $t = 0.84$ ; Donor effect  $F_{1,28} = 9.86$ ,  $p < 0.01$ ).

(I and J) At 8 weeks, FMT from either MRD or MHFD donors failed to improve sociability (I,  $GF_{MRDCol} p = 0.51$ ,  $t = 1.20$ ;  $GF_{MHFDCol} p = 0.28$ ,  $t = 1.58$ ; Donor effect  $F_{1,12} = 0.07$ ,  $p = 0.79$ ) or preference for social novelty in GF mice (J,  $GF_{MRDCol} p = 0.48$ ,  $t = 1.23$ ;  $GF_{MHFDCol} p > 0.99$ ,  $t = 0.043$ ; Donor effect  $F_{1,12} = 0.71$ ,  $p = 0.42$ ).

(K and L) PCoA of unweighted UniFrac distances based on the 16S rRNA gene sequencing dataset from GF recipients of stools from either MRD or MHFD donors at four (K,  $p = 0.001$ ,  $R^2 = 0.83$ ;  $n = 1,000$  rarefactions; 4,628 reads/sample) or eight (L,  $p < 0.001$ ,  $R^2 = 0.77$ ;  $n = 1,000$  rarefactions; 4,805 reads/sample) weeks of age. Plots show mean ± SEM.

See also Figure S4.



(legend on next page)

changes in the gut microbiome (Figure 5). Interestingly, according to a recent study, probiotic-based restoration of gut permeability in a mouse model of ASD can improve some behavioral abnormalities, but not social behaviors (Hsiao et al., 2013). Given that we identified a different probiotic candidate, *L. reuteri*, that rescues social behavior (Figures 4 and 5), but not other behavioral endophenotypes associated with ASD (Figure S6), in MHFD mice, we propose that a carefully selected combination of probiotics may be useful as a potential non-invasive treatment for patients suffering from neurodevelopmental disorders including ASD.

## EXPERIMENTAL PROCEDURES

### Mice and Maternal Diet

C57Bl6/J mice were obtained from Jackson Laboratories (#000-664) and were kept on a 12 hr light/dark cycle and had access to food and water ad libitum. Females were placed on either a regular diet (RD) consisting of 13.4% kcal from fat, 30% kcal from protein, and 57% kcal from carbohydrates (Lab Diets, #5001) or HFD consisting of 60% kcal from fat, 20% kcal from protein, and 20% kcal from carbohydrates (Research Diets, #D12492). Maternal weight was measured weekly. Maternal total and fat mass were measured using an mq7.5 Minispec NMR body composition analyzer. After 8 weeks on diet, females were paired with C57Bl6/J adult males to produce subject offspring. Resulting offspring were weaned at 3 weeks of age and all placed on RD, regardless of maternal diet (MRD or MHFD). Germ-free mice (C57Bl6/J) were maintained in a flexible isolator fed with HEPA-filtered air and provided with irradiated food and water. Germ-free offspring were weaned at 4 weeks of age. All behavioral tests were performed on 7- to 12-week-old male mice. Animal care and experimental procedures were approved by Baylor College of Medicine's Institutional Animal Care and Use Committee in accordance with all guidelines set forth by the U.S. National Institutes of Health.

### Reciprocal Social Interaction

Mice were placed in a 25 × 25 × 25 cm Plexiglass arena, to which they had not been previously habituated, with either a familiar (cage-mate) or stranger age- and sex-matched conspecific. In all experiments, paired mice were matched for maternal diet, colonization source, and/or treatment. We recorded the time a pair of mice engaged in social interaction (close following, touching, nose-to-nose sniffing, nose-to-anus sniffing, and/or crawling over/under each other). The human observer was blind to maternal diet and/or treatment group. Social behavior was analyzed with AnyMaze automated software.

### Three-Chamber Social Test

Crawley's three-chamber test for sociability and preference for social novelty was performed as described (Silverman et al., 2010). In brief, the mouse first experienced a 10 min period of habituation during which it was allowed to freely explore a 60 × 40 × 23 cm Plexiglass arena divided into three equally sized, interconnected chambers (left, center, right). Sociability was measured

during a second 10-min period in which the subject could interact either with an empty wire cup (empty) or a wire cup containing an age and sex-matched stranger conspecific (mouse 1). Time spent interacting (sniffing, crawling upon) with either the empty cup or the stranger mouse contained in the other cup as well as time spent in each chamber was recorded using the AnyMaze software, by independent observers. Empty cup placement in the left or right chamber during the sociability period was counterbalanced between trials. Finally, preference for social novelty was assayed by introducing a second stranger mouse (mouse 2) into the previously the empty wire cup. Time spent in each chamber as well as time spent interacting with either mouse 1 or mouse 2 was recorded using the automated AnyMaze software by independent observers.

### Marble Burying

Marble burying was performed as previously described (Thomas et al., 2009). Briefly, mice were placed in a standard-sized cage containing 20 regularly-spaced black marbles sitting on fine-wood chipped bedding 5 cm in depth. After 20 min, the mouse was removed and marbles with at least two-thirds of their depth obscured by wood chips were counted as buried.

### Open Field

Mice were placed in an open arena (40 × 40 × 20 cm) and allowed to explore freely for 10 min while their position was continually monitored using tracking software (AnyMaze). Tracking allowed for measurement of distance traveled, speed, and position in the arena throughout the task. Time spent in the center of the arena, defined as the interior 20 × 20 cm, was recorded.

### 16 s rRNA Gene Sequencing

Methods were adapted from protocols developed for the NIH-Human Microbiome Project (Human Microbiome Project Consortium, 2012a, 2012b). For a detailed protocol see Supplemental Experimental Procedures.

### Whole Genome Shotgun Sequencing

Individual libraries constructed from each sample were loaded onto the HiSeq platform (Illumina) and sequenced using the 2 × 100 bp pair-end read protocol. Illumina paired-end libraries were constructed from total genomic DNA isolated from each sample. The DNA was sheared into approximately 400–600 bp fragments followed by ligation of Illumina adaptors containing molecular barcodes for downstream de-multiplexing. These products were then amplified through ligation-mediated PCR (LM-PCR) using KAPA HiFi DNA Polymerase (Kapa Biosystems, Wilmington, MA, USA). Following bead purification with Agencourt AMPure XP (Beckman Coulter, Brea, CA, USA), quantification and size distribution of the LM-PCR product was determined using the LabChip GX electrophoresis system (PerkinElmer, Akron, OH, USA). Libraries were pooled in equimolar amounts at six samples per pool, and prepared for sequencing with TruSeq PE Cluster Generation Kit (Illumina). Each library pool was loaded onto one lane of a HiSeq 2000 flow cell spiked with 1% PhiX control library. Sequencing files were de-multiplexed with CASAVA version 1.8.3 (Illumina).

Quality filtering, trimming and de-multiplexing was carried out by a custom pipeline containing Trim Galore and cutadapt (Martin, 2011) for adaptor and

## Figure 4. Selective Treatment with *Lactobacillus reuteri* Restores Social Deficits and Oxytocin Levels in MHFD Offspring

(A) Schematic of *L. reuteri*-treatment.

(B and C) Unlike resuspension media (B,  $p > 0.99$ ,  $t = 1.03$ ; c,  $p > 0.99$ ,  $t = 0.40$ ) or heat-killed *L. reuteri* (B,  $p > 0.99$ ,  $t = 1.35$ ; c,  $p > 0.99$ ,  $t = 0.21$ ), administration of live *L. reuteri* in the drinking water rescued sociability (B,  $p < 0.0001$ ,  $t = 5.98$ ) and preference for social novelty (C,  $p < 0.001$ ,  $t = 5.01$ ) in MHFD offspring (B, treatment effect  $F_{1,86} = 87.53$ ,  $p < 0.0001$ ; C, treatment effect  $F_{1,86} = 30.24$ ,  $p < 0.0001$ ).

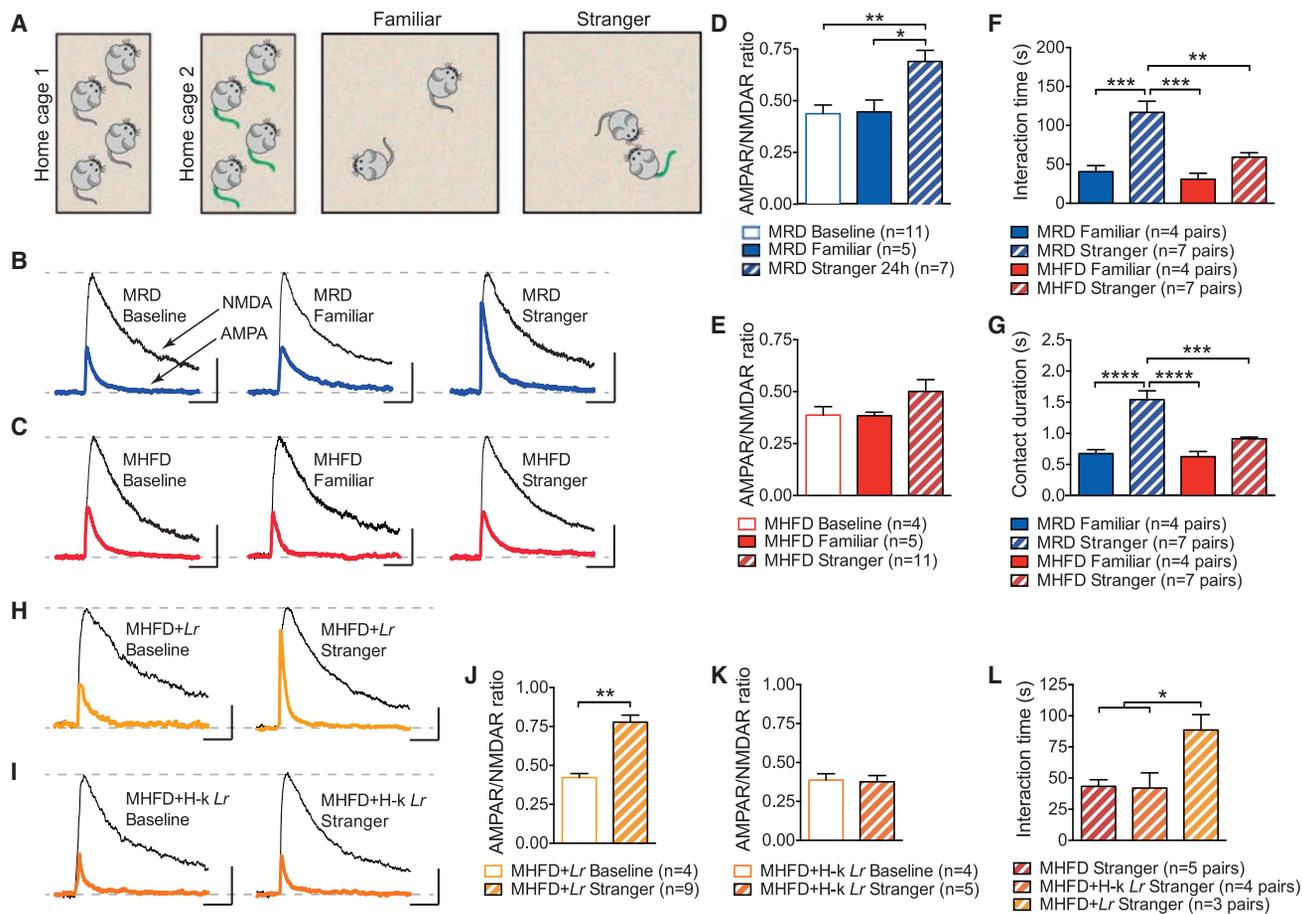
(D and E) Representative track plots showing exploratory activity.

(F) Representative images of control oxytocin immunoreactivity at different anteroposterior levels of the PVN.

(G–J) Oxytocin immunoreactivity in the PVN of MRD (G), MHFD (H), heat-killed *L. reuteri*-treated MHFD (I), and live *L. reuteri*-treated MHFD offspring (J).

(K–N) Oxytocin immunoreactive cell number (K,  $p < 0.01$ ,  $t = 4.76$ ) and oxytocin immunofluorescence intensity (L,  $p < 0.01$ ,  $t = 3.80$ ) were reduced in the PVN of MHFD versus MHFD mice. In the PVN of MRD and MHFD offspring, NeuN cell number immunoreactivity (M,  $p = 0.34$ ,  $t = 1.09$ ) and immunofluorescence intensity (N,  $p = 0.79$ ,  $t = 0.28$ ) were similar.

(O and P) Relative to treatment with heat-killed *L. reuteri*, treatment with live *L. reuteri* significantly increased oxytocin-positive cell number (O,  $p < 0.05$ ,  $t = 2.93$ ) and oxytocin immunofluorescence intensity (P,  $p < 0.05$ ,  $t = 3.09$ ) in the PVN of MHFD offspring. AU: arbitrary units. Plots show mean ± SEM. See also Figures S5 and S6.



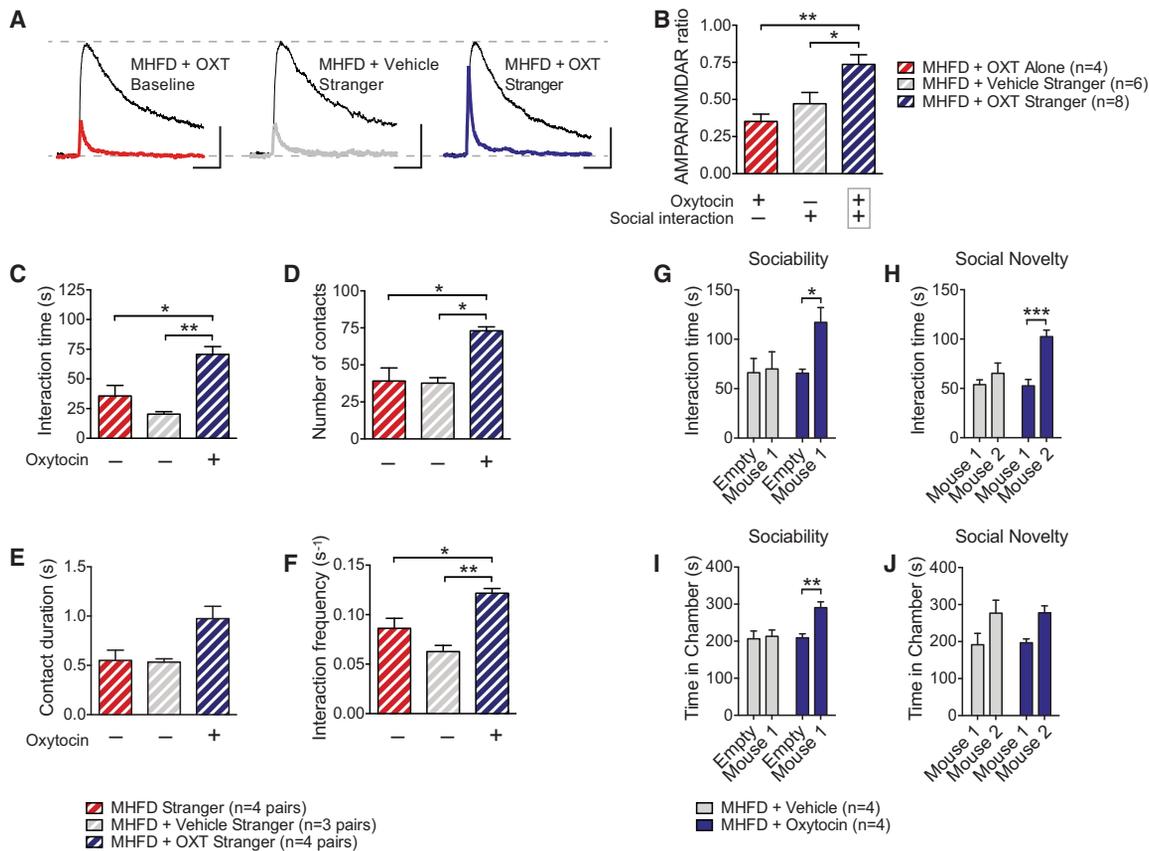
quality trimming, and PRINSEQ (Schmieder and Edwards, 2011) for low-complexity filtering and sequence deduplication. In addition, Bowtie2 v2.2.1 was used to map reads to MetaPhiAn markers for the classification of bacterial species (Segata et al., 2012).

#### Colonization of Germ-Free Mice by Fecal Microbiota Transplant

Fresh fecal samples were collected from donor mice/microbiome cohort and homogenized on ice in sterile PBS under sterile conditions. The resulting slurry was spun at 1,000g for 3 min at 4°C. The supernatants were isolated and diluted to  $5 \times 10^9$  CFU/ml with sterile PBS. Four- or eight-week-old C57Bl6/J germ-free (GF) recipient mice were then immediately colonized by a single gavage with 0.2 mL solution. Fecal samples were collected from the colonized GF mice at 24 hr, 7 days, 14 days, 28 days, and 56 days following colonization. Fecal samples were snap frozen and stored at  $-80^\circ\text{C}$  until prepared for sequencing. Behavioral experiments were initiated at 3 weeks post-transplant.

#### Culture and Treatment with *L. reuteri* and *L. johnsonii*

*Lactobacillus reuteri* MM4-1A (ATCC-PTA-6475) and *Lactobacillus johnsonii* (ATCC 33200) were cultured anaerobically in MRS broth in a 90%  $\text{N}_2$ , 5%  $\text{CO}_2$ , 5%  $\text{H}_2$  environment. *L. reuteri* was heat-killed by keeping the bacteria at  $80^\circ\text{C}$  for 20 min. Bacterial viability was assessed by plating and the efficacy of the heat-kill procedure was confirmed by the absence of colony growth following plating. Cultures were centrifuged, washed, and resuspended in an anaerobic solution (PBS) and frozen at  $-80^\circ\text{C}$  until use. PBS, live or heat-killed *L. reuteri* were added to the drinking water, which was changed daily to minimize dosage variability. Whereas the experimental group received live bacteria, one control group received identically prepared cultures of heat-killed bacteria. A second group of control mice received water treated with PBS alone. Live and heat-killed *L. reuteri* were supplied at a dosage of  $1 \times 10^8$  organisms/mouse/day continuously in drinking water. Mice consumed the treated water ad libitum over the treatment period. The treated drinking water



**Figure 6. Oxytocin Restores Social Interaction-Induced VTA Plasticity and Social Behavioral Deficits in MHFD Offspring**

(A and B) LTP was measured 1–3 hr following a reciprocal interaction. Intranasal oxytocin administration rescued LTP in the VTA of MHFD offspring (B, MHFD+OXT alone versus MHFD+OXT stranger  $p < 0.01$ ,  $t = 3.66$ ; MHFD+Vehicle stranger versus MHFD+OXT stranger  $p < 0.05$ ,  $t = 2.86$ ;  $F_{2,15} = 7.97$ ,  $p < 0.01$ ). (C–F) Oxytocin restored reciprocal social interaction (C, MHFD versus MHFD+Vehicle  $p = 0.55$ ,  $t = 1.46$ ; MHFD versus MHFD+OXT  $p < 0.05$ ,  $t = 3.62$ ; MHFD+Vehicle versus MHFD+OXT  $p < 0.01$ ,  $t = 4.81$ ;  $F_{2,8} = 12.82$ ,  $p < 0.01$ ; D, MHFD versus MHFD+Vehicle  $p > 0.99$ ,  $t = 0.16$ ; MHFD versus MHFD+OXT  $p < 0.05$ ,  $t = 4.075$ ; MHFD+Vehicle versus MHFD+OXT  $p < 0.05$ ,  $t = 3.94$ ;  $F_{2,8} = 10.97$ ,  $p < 0.01$ ; E, MHFD versus MHFD+Vehicle,  $p > 0.99$ ,  $t = 0.11$ ; MHFD versus MHFD+OXT  $p = 0.052$ ,  $t = 2.99$ ; treatment effect  $F_{2,8} = 5.87$ ,  $p < 0.05$ ; F, MHFD versus MHFD+Vehicle  $p = 0.2$ ,  $t = 2.11$ ; MHFD versus MHFD+OXT  $p < 0.05$ ,  $t = 3.43$ ; treatment effect  $F_{2,8} = 14.58$ ,  $p < 0.01$ ), sociability (G, MHFD+Vehicle  $p > 0.99$ ,  $t = 0.44$ ; MHFD+Oxytocin  $p = 0.24$ ,  $t = 1.74$ ; treatment effect  $F_{1,8} = 2.37$ ,  $p = 0.16$ ; H, MHFD+Vehicle  $p > 0.99$ ,  $t = 0.29$ ; MHFD+Oxytocin  $p < 0.01$ ,  $t = 3.50$ ; treatment effect  $F_{1,12} = 5.16$ ,  $p < 0.05$ ) and preference for social novelty in MHFD offspring (I, MHFD+Vehicle  $p = 0.65$ ,  $t = 1.05$ ; MHFD+Oxytocin  $p < 0.05$ ,  $t = 3.54$ ; treatment effect  $F_{1,8} = 10.54$ ,  $p < 0.05$ ; J, MHFD+Vehicle  $p = 0.50$ ,  $t = 1.25$ ; MHFD+Oxytocin  $p = 0.096$ ,  $t = 2.34$ ; treatment effect  $F_{(1,8)} = 6.41$ ,  $p < 0.05$ ). Plots show mean  $\pm$  SEM.

for each group was replaced daily 2 hr prior to the onset of the dark cycle to minimize variation in microbial exposure. Behavioral assays were initiated after 4 weeks of *L. reuteri* or control treatment. The protocol of the *L. johnsonii* preparation and administration matched the *L. reuteri* protocol. Fecal samples for sequencing and tissue used in the immunofluorescence studies were collected at the end of the treatment.

#### Immunofluorescence

Mice were deeply anesthetized by inhalation of isoflurane and perfused transcardially with 10 mL 0.9% phosphate-buffered saline followed by 30 mL 4% paraformaldehyde in 0.1M phosphate buffer (PFA). Brains were post-fixed in 4% PFA at 4°C overnight, then cryoprotected in 30% sucrose 0.1M PB over 3 days. Coronal slices (30  $\mu$ m) thick were obtained from frozen tissue using a sliding blade microtome then transferred to ice cold PBS. Slices were blocked with 5% normal goat serum, 0.3% Triton X-100 0.1M PB (PBTgs) for 1 hr rocking at RT and then incubated in primary antibodies (rabbit anti-oxytocin, ImmunoStar #20068, 1:2,000; mouse anti-NeuN, Millipore, #MAB377, 1:2,000) diluted in PBTgs rocking at 4°C for 24 hr. Slices were then washed three times with 0.3% Triton X-100 0.1M PB. Primary antibodies were visual-

ized using secondary goat anti-rabbit Alexa Fluor 488 (ThermoFisher Scientific, #A-11034) and goat anti-mouse Alexa Fluor 594 (ThermoFisher Scientific, #A-11032) antibodies (1:1,000 dilution). Slices were incubated in secondary antibodies rocking in the dark for 1h at RT. Five minute final washes with each of PBTgs, 0.1M PB, and 0.05M PB preceded mounting onto 2% gelatin (Sigma-Aldrich, #G9391)-coated coverslips. Nuclei were visualized using Vectashield H-1200 with DAPI (Vector Labs, #H-1200).

Fluorescent imaging and data acquisition was performed on a Zeiss AxioImager.Z2 microscope (Carl Zeiss MicroImaging) mounted with an AxioCam digital camera (Carl Zeiss MicroImaging). Images were captured using AxioVision acquisition software (Carl Zeiss MicroImaging). All images within a given dataset were acquired at identical exposure times, within a given channel, to allow comparison of signal intensity. In some images, contrast and brightness were linearly adjusted using Photoshop (Adobe). Image processing was applied uniformly across all images within a given dataset. Fluorescence intensity was measured in ImageJ (NIH) by selecting regions of interest (Oxytocin- and NeuN-positive hypothalamic cell bodies). Hypothalamic oxytocin-expressing neuronal population and NeuN+ cell number was assessed in ImageJ using the following operational sequence: (1) open image

file, (2) subtract background, (3) adjust threshold, (4) convert to mask, (5) watershed, (6) analyze particles. Automatic identification of cell boundaries was validated against the source image by an experimenter blind to group allocation.

### Oxytocin Administration

Oxytocin was obtained from Tocris Bioscience (product 1910) and solubilized in 10% dimethyl sulfoxide (DMSO) in PBS. 10% DMSO in PBS was used as the vehicle control. Mice received oxytocin intranasally (at approximately 200 µg/kg) 30 min prior to behavior. 1.25 µL of oxytocin or vehicle solution was injected into each nostril from P10 pipette. Oxytocin dose was selected according to dosages reported to rescue social behavior in genetic models phenotypically expressing ASD-like behaviors (Peñagarikano et al., 2015).

### Electrophysiology

Recordings were performed as recently described (Huang et al., 2016) and the investigators were kept blind to treatment conditions. Briefly, mice were anesthetized with a mixture of ketamine (100 mg/kg), xylazine (10 mg/kg), and acepromazine (3 mg/kg). Horizontal slices (225–300 µm thick) containing the VTA were cut from the brains of C57BL/6J mice. Mice were transcardially perfused with an ice-cold, oxygenated solution containing (in mM) NaCl, 87; NaHCO<sub>3</sub>, 25; KCl, 2.5; NaH<sub>2</sub>PO<sub>4</sub>, 1.25; MgCl<sub>2</sub>, 7; CaCl<sub>2</sub>, 0.5; dextrose, 25; sucrose, 75. Horizontal slices were cut with a vibrating tissue slicer (VF-100 Compressome, Precisionary Instruments, San Jose, CA, or Leica VT 1000S, Leica Microsystems, Buffalo Grove, IL), incubated at 34°C for 40 min, kept at room temperature for at least 30 min before their transfer to a recording chamber continuously perfused with artificial cerebrospinal fluid (ACSF) at 32°C and a flow rate of 2–3 mL/min. The recording ACSF contained in mM: 120 NaCl, 3.3 KCl, 1.25 NaH<sub>2</sub>PO<sub>4</sub>, 25 NaHCO<sub>3</sub>, 10 Dextrose, 1 MgCl<sub>2</sub> and 2 CaCl<sub>2</sub>. Recording pipettes were made from thin-walled borosilicate glass (TW150F-4, WPI, Sarasota, FL). After filling with intracellular solution (in mM): 117 CsMeSO<sub>3</sub>; 0.4 EGTA; 20 HEPES; 2.8 NaCl, 2.5 ATP-Mg 2.0; 0.25 GTP-Na; 5 TEA-Cl, adjusted to pH 7.3 with CsOH and 290 mosmol/L, they had a resistance of 3–5 MΩ. Data were obtained with a MultiClamp 700B amplifier, digitized at 20 kHz with a Digidata 1440A, recorded by Clampex 10 and analyzed with Clampfit 10 software (Molecular Devices). Recordings were filtered online at 3 kHz with a Bessel low-pass filter. A 2 mV hyperpolarizing pulse was applied before each EPSC to evaluate the input and access resistance (R<sub>a</sub>). Data were discarded when R<sub>a</sub> was either unstable or greater than 25MΩ, holding current was >200 pA, input resistance dropped >20% during the recording, or EPSCs baseline changed by >10%. After establishing a gigaohm seal (>2GΩ) and recording stable spontaneous firing in cell-attached, voltage clamp mode (–70 mV holding potential), cell phenotype was determined by measuring the width of the action potential and the presence of an I<sub>h</sub> current. AMPAR/NMDAR ratios were calculated as previously described (Huang et al., 2016). Briefly, neurons were voltage-clamped at +40 mV until the holding current stabilized (at <200 pA). Monosynaptic EPSCs were evoked at 0.05 Hz with a bipolar stimulating electrode placed 50–150 µm rostral to the lateral VTA. Picrotoxin (100 µM) was added to the recording ACSF to block GABA<sub>A</sub>-mediated IPSCs. After recording the dual-component EPSC, DL-AP5 (100 µM) was bath-applied for 10 min to remove the NMDAR component, which was then obtained by offline subtraction of the remaining AMPAR component from the original EPSC. The peak amplitudes of the isolated components were used to calculate the AMPAR/NMDAR ratios. Picrotoxin and DL-AP5 were purchased from Tocris Bioscience and all other reagents were obtained from Sigma-Aldrich.

### Statistical Analysis

Data are presented as mean ± SEM. For behavioral experiments, statistics were based on the two-sided unpaired Student's *t* tests, one- or two-way ANOVA with Bonferroni post hoc analysis to correct for multiple comparisons, unless otherwise indicated. *P*, *t*, and *F* values are presented in the figure legends, *n* values are provided in the figures. *p* < 0.05 was considered significant. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. GraphPad's *Prism* 6 (La Jolla, CA) software was used to perform statistical analyses and for generating graphical representations of data. For 16S rRNA gene sequencing, analysis and visualization of microbiome communities was conducted in RStudio

0.99.292 (<http://www.R-project.org/> [2014]), utilizing the phyloseq package (McMurdie and Holmes, 2013) to import sample data and calculate alpha- and beta-diversity metrics. 16S rRNA gene sequencing data was analyzed using Silva 115 and 123 (Quast et al., 2013). Analyses were performed on datasets that were rarefied 1,000×, then averaged and rounded. Amplicon sequences were deposited to the NCBI Sequence Read Archive under accession number TBD. Dirichlet Multinomial Mixture modeling was performed in mother (Schloss et al., 2009). Significance of categorical variables were determined using the non-parametric Mann-Whitney test for two category comparisons or the Kruskal-Wallis test when comparing three or more categories. One-way ANOVA followed by Dunnett's multiple comparisons test was performed using GraphPad's *Prism* 6. Correlation between two continuous variables was determined by linear regressions, where *p* values indicate the probability that the slope of the regression line is zero. Principal coordinate plots employed the Monte Carlo permutation test to estimate *p* values. All *p* values were adjusted for multiple comparisons with the FDR algorithm (Benjamini et al., 2001). No animals or data points were excluded from analyses.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and seven figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.06.001>.

A video abstract is available at <http://dx.doi.org/10.1016/j.cell.2016.06.001#mmc3>.

### AUTHOR CONTRIBUTIONS

Conceptualization and Methodology, S.A.B. and M.C.-M.; Investigation, S.A.B. and G.V.D.P.; Analysis, S.A.B., G.V.D.P., T.A.A., N.J.A., J.F.P., and M.C.-M.; Writing – Original Draft, S.A.B. and M.C.-M.; Writing – Review & Editing, S.A.B., G.V.D.P., T.A.A., N.J.A., J.F.P., and M.C.-M.; Visualization, S.A.B.; Funding Acquisition, M.C.-M.; Supervision, M.C.-M.

### ACKNOWLEDGMENTS

We thank members of the Costa-Mattioli laboratory, K. Krnjević, M. Beauchamp, and H. Dierick for comments on the manuscript; A. Swennes, J. Garcia, D. Quach and J. Auchtung for technical advice. This work was supported by funding from the NIH (NIMH 096816, NINDS 076708) to M.C.-M and the Alkek Foundation and Baylor College of Medicine to the Alkek Center for Metagenomics and Microbiome Research to J.F.P.

Received: November 18, 2015

Revised: February 26, 2016

Accepted: May 31, 2016

Published: June 16, 2016

### REFERENCES

- Aye, I.L., Rosario, F.J., Powell, T.L., and Jansson, T. (2015). Adiponectin supplementation in pregnant mice prevents the adverse effects of maternal obesity on placental function and fetal growth. *Proc. Natl. Acad. Sci. USA* *112*, 12858–12863.
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* *125*, 279–284.
- Bilder, D.A., Bakian, A.V., Viskochil, J., Clark, E.A., Botts, E.L., Smith, K.R., Pimentel, R., McMahon, W.M., and Coon, H. (2013). Maternal prenatal weight gain and autism spectrum disorders. *Pediatrics* *132*, e1276–e1283.
- Bolton, J.L., and Bilbo, S.D. (2014). Developmental programming of brain and behavior by perinatal diet: focus on inflammatory mechanisms. *Dialogues Clin. Neurosci.* *16*, 307–320.
- Bravo, J.A., Forsythe, P., Chew, M.V., Escaravage, E., Savignac, H.M., Dinan, T.G., Bienenstock, J., and Cryan, J.F. (2011). Ingestion of *Lactobacillus* strain

- regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc. Natl. Acad. Sci. USA* *108*, 16050–16055.
- Bresnahan, M., Hornig, M., Schultz, A.F., Gunnes, N., Hirtz, D., Lie, K.K., Magnus, P., Reichborn-Kjennerud, T., Roth, C., Schjølberg, S., et al. (2015). Association of maternal report of infant and toddler gastrointestinal symptoms with autism: evidence from a prospective birth cohort. *JAMA Psychiatry* *72*, 466–474.
- Connolly, N., Anixt, J., Manning, P., Ping-I Lin, D., Marsolo, K.A., and Bowers, K. (2016). Maternal metabolic risk factors for autism spectrum disorder—An analysis of electronic medical records and linked birth data. *Autism Res.* Published online January 29, 2016.
- Cryan, J.F., and Dinan, T.G. (2012). Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* *13*, 701–712.
- Davari, S., Talaei, S.A., Alaei, H., and Salami, M. (2013). Probiotics treatment improves diabetes-induced impairment of synaptic activity and cognitive function: behavioral and electrophysiological proofs for microbiome-gut-brain axis. *Neuroscience* *240*, 287–296.
- Desbonnet, L., Clarke, G., Shanahan, F., Dinan, T.G., and Cryan, J.F. (2014). Microbiota is essential for social development in the mouse. *Mol. Psychiatry* *19*, 146–148.
- Dölen, G., Darvishzadeh, A., Huang, K.W., and Malenka, R.C. (2013). Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. *Nature* *501*, 179–184.
- Donaldson, Z.R., and Young, L.J. (2008). Oxytocin, vasopressin, and the neurogenetics of sociality. *Science* *322*, 900–904.
- Galley, J.D., Bailey, M., Kamp Dush, C., Schoppe-Sullivan, S., and Christian, L.M. (2014). Maternal obesity is associated with alterations in the gut microbiome in toddlers. *PLoS ONE* *9*, e113026.
- Groppe, S.E., Gossen, A., Rademacher, L., Hahn, A., Westphal, L., Gründer, G., and Spreckelmeyer, K.N. (2013). Oxytocin influences processing of socially relevant cues in the ventral tegmental area of the human brain. *Biol. Psychiatry* *74*, 172–179.
- Gunaydin, L.A., Grosenick, L., Finkelstein, J.C., Kauvar, I.V., Fenno, L.E., Adhikari, A., Lammel, S., Mirzabekov, J.J., Airan, R.D., Zalocusky, K.A., et al. (2014). Natural neural projection dynamics underlying social behavior. *Cell* *157*, 1535–1551.
- Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J., Fedele, A., Collins, J., Smith, K., et al. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Arch. Gen. Psychiatry* *68*, 1095–1102.
- Hsiao, E.Y., McBride, S.W., Hsien, S., Sharon, G., Hyde, E.R., McCue, T., Codelli, J.A., Chow, J., Reisman, S.E., Petrosino, J.F., et al. (2013). Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell* *155*, 1451–1463.
- Huang, Y.C., and Hessler, N.A. (2008). Social modulation during songbird courtship potentiates midbrain dopaminergic neurons. *PLoS ONE* *3*, e3281.
- Huang, W., Placzek, A.N., Viana Di Prisco, G., Khatiwada, S., Sidrauski, C., Krnjević, K., Walter, P., Dani, J.A., and Costa-Mattioli, M. (2016). Translational control by eIF2 $\alpha$  phosphorylation regulates vulnerability to the synaptic and behavioral effects of cocaine. *eLife* *5*, 5.
- Human Microbiome Project Consortium (2012a). A framework for human microbiome research. *Nature* *486*, 215–221.
- Human Microbiome Project Consortium (2012b). Structure, function and diversity of the healthy human microbiome. *Nature* *486*, 207–214.
- King, J.C. (2006). Maternal obesity, metabolism, and pregnancy outcomes. *Annu. Rev. Nutr.* *26*, 271–291.
- Krakowiak, P., Walker, C.K., Bremer, A.A., Baker, A.S., Ozonoff, S., Hansen, R.L., and Hertz-Picciotto, I. (2012). Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders. *Pediatrics* *129*, e1121–e1128.
- Lerer, E., Levi, S., Salomon, S., Darvasi, A., Yirmiya, N., and Ebstein, R.P. (2008). Association between the oxytocin receptor (OXTR) gene and autism: relationship to Vineland Adaptive Behavior Scales and cognition. *Mol. Psychiatry* *13*, 980–988.
- Ma, J., Prince, A.L., Bader, D., Hu, M., Ganu, R., Baquero, K., Blundell, P., Alan Harris, R., Frias, A.E., Grove, K.L., and Aagaard, K.M. (2014). High-fat maternal diet during pregnancy persistently alters the offspring microbiome in a primate model. *Nat. Commun.* *5*, 3889.
- Martin, M. (2011). Cutadapt removes adaptor sequences from high-throughput sequencing reads. *EMBnetjournal* *17*, 10–12.
- Mathers, J.C., and McKay, J.A. (2009). Epigenetics - potential contribution to fetal programming. *Adv. Exp. Med. Biol.* *646*, 119–123.
- Mayer, E.A., Padua, D., and Tillisch, K. (2014). Altered brain-gut axis in autism: comorbidity or causative mechanisms? *BioEssays* *36*, 933–939.
- Mayer, E.A., Tillisch, K., and Gupta, A. (2015). Gut/brain axis and the microbiota. *J. Clin. Invest.* *125*, 926–938.
- McMurdie, P.J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* *8*, e61217.
- Mefford, H.C., Batshaw, M.L., and Hoffman, E.P. (2012). Genomics, intellectual disability, and autism. *N. Engl. J. Med.* *366*, 733–743.
- Melis, M.R., Melis, T., Cocco, C., Succu, S., Sanna, F., Pillolla, G., Boi, A., Ferri, G.L., and Argiolas, A. (2007). Oxytocin injected into the ventral tegmental area induces penile erection and increases extracellular dopamine in the nucleus accumbens and paraventricular nucleus of the hypothalamus of male rats. *Eur. J. Neurosci.* *26*, 1026–1035.
- Modi, M.E., and Young, L.J. (2012). The oxytocin system in drug discovery for autism: animal models and novel therapeutic strategies. *Horm. Behav.* *61*, 340–350.
- Moss, B.G., and Chugani, D.C. (2014). Increased risk of very low birth weight, rapid postnatal growth, and autism in underweight and obese mothers. *Am. J. Health Promot.* *28*, 181–188.
- Parracho, H.M., Bingham, M.O., Gibson, G.R., and McCartney, A.L. (2005). Differences between the gut microflora of children with autistic spectrum disorders and that of healthy children. *J. Med. Microbiol.* *54*, 987–991.
- Pedersen, C.A., Caldwell, J.D., Walker, C., Ayers, G., and Mason, G.A. (1994). Oxytocin activates the postpartum onset of rat maternal behavior in the ventral tegmental and medial preoptic areas. *Behav. Neurosci.* *108*, 1163–1171.
- Peñagarikano, O., Lázaro, M.T., Lu, X.H., Gordon, A., Dong, H., Lam, H.A., Peles, E., Maidment, N.T., Murphy, N.P., Yang, X.W., et al. (2015). Exogenous and evoked oxytocin restores social behavior in the *Cntnap2* mouse model of autism. *Sci. Transl. Med.* *7*, 271ra8.
- Poutahidis, T., Kearney, S.M., Levkovich, T., Qi, P., Varian, B.J., Lakritz, J.R., Ibrahim, Y.M., Chatzigiagkos, A., Alm, E.J., and Erdman, S.E. (2013). Microbial symbionts accelerate wound healing via the neuropeptide hormone oxytocin. *PLoS ONE* *8*, e78898.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* *41*, D590–D596.
- Ridaura, V.K., Faith, J.J., Rey, F.E., Cheng, J., Duncan, A.E., Kau, A.L., Griffin, N.W., Lombard, V., Henrissat, B., Bain, J.R., et al. (2013). Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* *341*, 1241214.
- Sabatier, N., Leng, G., and Menzies, J. (2013). Oxytocin, feeding, and satiety. *Front. Endocrinol. (Lausanne)* *4*, 35.
- Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* *75*, 7537–7541.
- Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* *27*, 863–864.

- Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., and Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* 9, 811–814.
- Silverman, J.L., Yang, M., Lord, C., and Crawley, J.N. (2010). Behavioural phenotyping assays for mouse models of autism. *Nat. Rev. Neurosci.* 11, 490–502.
- Skinner, A.C., and Skelton, J.A. (2014). Prevalence and trends in obesity and severe obesity among children in the United States, 1999–2012. *JAMA pediatrics* 168, 561–566.
- Song, S.J., Lauber, C., Costello, E.K., Lozupone, C.A., Humphrey, G., Berg-Lyons, D., Caporaso, J.G., Knights, D., Clemente, J.C., Nakielny, S., et al. (2013). Cohabiting family members share microbiota with one another and with their dogs. *eLife* 2, e00458.
- Sullivan, E.L., Nosen, E.K., and Chamlou, K.A. (2014). Maternal high fat diet consumption during the perinatal period programs offspring behavior. *Physiol. Behav.* 123, 236–242.
- Tang, Y., Chen, Z., Tao, H., Li, C., Zhang, X., Tang, A., and Liu, Y. (2014). Oxytocin activation of neurons in ventral tegmental area and interfascicular nucleus of mouse midbrain. *Neuropharmacology* 77, 277–284.
- Thomas, A., Burant, A., Bui, N., Graham, D., Yuva-Paylor, L.A., and Paylor, R. (2009). Marble burying reflects a repetitive and perseverative behavior more than novelty-induced anxiety. *Psychopharmacology (Berl.)* 204, 361–373.
- Tremaroli, V., and Bäckhed, F. (2012). Functional interactions between the gut microbiota and host metabolism. *Nature* 489, 242–249.
- Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444, 1027–1031.
- Uvnäs-Moberg, K., Handlin, L., and Petersson, M. (2014). Self-soothing behaviors with particular reference to oxytocin release induced by non-noxious sensory stimulation. *Front. Psychol.* 5, 1529.
- Wang, X., Wang, B.R., Zhang, X.J., Xu, Z., Ding, Y.Q., and Ju, G. (2002). Evidences for vagus nerve in maintenance of immune balance and transmission of immune information from gut to brain in STM-infected rats. *World J. Gastroenterol.* 8, 540–545.
- Wu, S., Jia, M., Ruan, Y., Liu, J., Guo, Y., Shuang, M., Gong, X., Zhang, Y., Yang, X., and Zhang, D. (2005). Positive association of the oxytocin receptor gene (OXTR) with autism in the Chinese Han population. *Biol. Psychiatry* 58, 74–77.
- Yatsunencko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 222–227.

# Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery

Alan Merk,<sup>1,3</sup> Alberto Bartesaghi,<sup>1,3</sup> Soojay Banerjee,<sup>1,3</sup> Veronica Falconieri,<sup>1</sup> Prashant Rao,<sup>1</sup> Mindy I. Davis,<sup>2</sup> Rajan Pragani,<sup>2</sup> Matthew B. Boxer,<sup>2</sup> Lesley A. Earl,<sup>1</sup> Jacqueline L.S. Milne,<sup>1</sup> and Sriram Subramaniam<sup>1,\*</sup>

<sup>1</sup>Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, MD 20892, USA

<sup>2</sup>National Center for Advancing Translational Sciences, 9800 Medical Center Drive, Rockville, MD 20850, USA

<sup>3</sup>Co-first author

\*Correspondence: ss1@nih.gov

<http://dx.doi.org/10.1016/j.cell.2016.05.040>

## SUMMARY

Recent advances in single-particle cryoelectron microscopy (cryo-EM) are enabling generation of numerous near-atomic resolution structures for well-ordered protein complexes with sizes  $\geq \sim 200$  kDa. Whether cryo-EM methods are equally useful for high-resolution structural analysis of smaller, dynamic protein complexes such as those involved in cellular metabolism remains an important question. Here, we present 3.8 Å resolution cryo-EM structures of the cancer target isocitrate dehydrogenase (93 kDa) and identify the nature of conformational changes induced by binding of the allosteric small-molecule inhibitor ML309. We also report 2.8-Å- and 1.8-Å-resolution structures of lactate dehydrogenase (145 kDa) and glutamate dehydrogenase (334 kDa), respectively. With these results, two perceived barriers in single-particle cryo-EM are overcome: (1) crossing 2 Å resolution and (2) obtaining structures of proteins with sizes  $< 100$  kDa, demonstrating that cryo-EM can be used to investigate a broad spectrum of drug-target interactions and dynamic conformational states.

## INTRODUCTION

Cryo-electron microscopy (cryo-EM) is now firmly established as a central tool in the arsenal of structural biology. The ability to obtain near-atomic resolution structures using cryo-EM was shown initially almost three decades ago in the context of electron crystallographic studies of membrane proteins (Henderson et al., 1990). Continued advances in single-particle cryo-EM over the next two decades enabled high-resolution analysis of non-crystalline samples with high internal symmetry such as icosahedral and helical viruses (Ge and Zhou, 2011; Settembre et al., 2011; Yu et al., 2008; Zhang et al., 2010). Large and relatively stable complexes such as ribosomes also proved especially amenable to analysis using cryo-EM methods, first at medium resolution (Matadeen et al., 1999; Rawat et al., 2003) and more recently at near-atomic resolution (Amunts et al., 2014; Fischer et al., 2015; Jomaa et al., 2016; Wong et al., 2014). These suc-

cesses have now been extended to a wide spectrum of protein complexes, including several integral membrane proteins (Bai et al., 2015b; Du et al., 2015; Liao et al., 2013; Matthies et al., 2016). Structures determined by cryo-EM can now reach resolutions as high as 2.2 Å and 2.3 Å, as exemplified by structures of the 465 kDa  $\beta$ -galactosidase (Bartesaghi et al., 2015) and the 540 kDa AAA ATPase p97 (Banerjee et al., 2016). However, all of the near-atomic-resolution structures reported have been of proteins with sizes in the range of  $\sim 200$  kDa or larger, and an informal opinion in the field is that cryo-EM technology is primarily suited for analysis of relatively stable proteins with sizes  $>150$  kDa (Thompson et al., 2016). The smallest protein for which a cryo-EM structure has been reported using single particle cryo-EM is that of the 135 kDa ABC exporter TmrAB, at  $\sim 10$  Å resolution (Kim et al., 2015), and the challenges in achieving near-atomic resolution for small proteins, even with sizes as large as  $\sim 300$  kDa have been noted (Skiniotis and Southworth, 2016; Cheng, 2015; Belnap, 2015). Crystallographic analyses are generally very powerful for proteins in this size range but usually require that flexible regions are either deleted or altered. Given that small, dynamic protein complexes are implicated in numerous cellular processes, there is considerable interest in determining whether cryo-EM methods can be also applicable for structural analysis of this class of proteins under near-native conditions and at near-atomic resolution.

A principal reason why small proteins such as those with sizes  $<150$  kDa have been considered intractable for analysis by cryo-EM is that the errors in alignment of individual projection images become progressively higher as the size of the scattering entity gets smaller (Henderson et al., 2011). In principle, with a perfect detector that displays minimal falloff in detective quantum efficiency (DQE) even at Nyquist frequency (Henderson, 1995; McMullan et al., 2014), it should be possible to achieve accurate alignment of projection images of smaller proteins, but all currently available detectors still show a significant drop in DQE at higher frequencies. The use of phase plates is an option that may help partially alleviate the problem of image contrast (Danev and Baumeister, 2016), but these developments are still at an early stage. One way to compensate for the falloff in DQE at higher spatial frequencies is to collect data at higher magnification. This strategy, however, lowers contrast and makes the alignment of individual frames collected in movie mode of data collection more challenging. Experimental approaches to optimize specimen preparation provide an alternative route to

**Table 1. Summary of Data Collection and Image-Processing Parameters for the Density Maps of LDH, IDH1, and GDH Reported in this Manuscript**

	Total number of micrographs	Defocus range (microns)	Starting number of particles	Particles used in final map	Frames used in final map	B-factor for sharpening ( $\text{\AA}^2$ )	Final resolution ( $\text{\AA}$ )
LDH-inh	1,707	0.8–2.2	508,402	50,865 (10%)	1–30	–150	2.8
IDH-apo	1,506	0.7–3.0	499,184	49,936 (10%)	1–30	–180	3.8
IDH-inh	820	0.7–2.6	232,343	46,483 (20%)	1–30	–180	3.8
GDH	232	0.8–2.1	45,388	21,818 (48%)	3–9	–90	1.8

improve image contrast: it can be easier to minimize the background scattering from the ice layer for smaller proteins because the lower the molecular weight, the thinner the ice layer that is required to surround the protein with an aqueous layer.

To further test the limits of what is possible with present-day cryo-EM technology, we have analyzed structures of two small, soluble enzymes implicated in cancer metabolism: the 145 kDa lactate dehydrogenase (LDH B, a tetramer composed of four identical  $\sim 36$  kDa subunits) and the 93 kDa isocitrate dehydrogenase (IDH1, a dimer composed of two identical  $\sim 47$  kDa subunits). In both cases, we tested whether structures can be obtained at high enough resolution to localize bound small-molecule ligands and to determine the structures of ligand-bound complexes. We also carried out structural analysis of the conformationally dynamic enzyme glutamate dehydrogenase (GDH). Earlier crystallographic analyses of apo-GDH, a 334 kDa hexameric protein composed of six identical  $\sim 56$  kDa subunits, have been limited to a resolution of 2.7  $\text{\AA}$ , possibly because some regions of the protein may display a continuum of conformational states even in the context of a crystal lattice. This is a feature that may well be shared by a large majority of protein complexes involved in signaling and metabolism. While it is expected that image averaging will limit resolutions that can be achieved for the conformationally flexible components, it is interesting to ascertain whether cryo-EM analyses can be used to obtain atomic resolution for the ordered regions of these types of protein complexes.

## RESULTS

### Structural Studies of Lactate Dehydrogenase

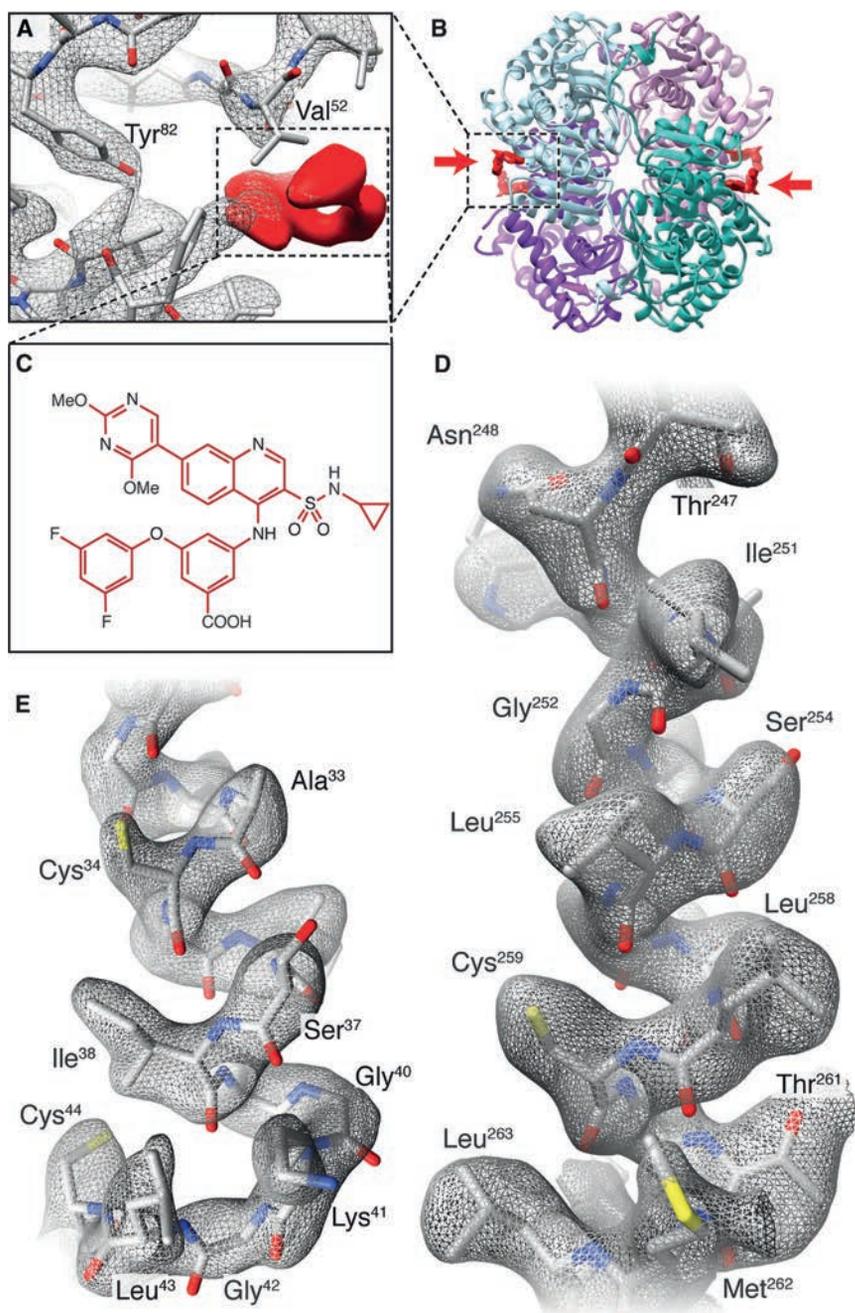
Cancer cells catalyze the production of lactate from pyruvate, the final metabolite generated from glucose during glycolysis. In non-malignant cells, pyruvate enters mitochondria to participate in the Krebs cycle and in oxidative phosphorylation. Under anaerobic conditions, such as that found in muscle cells, tetrameric LDH mediates the formation of lactate from pyruvate using NADH as a cofactor. However, in cancer cells, irrespective of oxygen availability, glycolysis followed by production of lactate through LDH is the preferred pathway, enhancing the production of metabolic precursors required for biosynthesis of cellular macromolecules. The recognition in a growing number of cancers of the central role of both LDH A and B isoforms (Fiume et al., 2014; McClelland et al., 2013; Rodriguez et al., 2003) has driven the identification of small-molecule inhibitors, including mercaptocyclohex-2-enone derivatives that bind away from the NAD-binding pocket and are not competitive inhibitors

(Dragovich et al., 2014), as well as sulfamoylquinoline benzoic acid derivatives, which compete with NAD binding to LDH (Billiard et al., 2013). To evaluate the potential of cryo-EM to determine structures of a small protein complex such as LDH and localize the binding sites of potential small-molecule inhibitors, we carried out cryo-EM analysis of LDH B in complex with GSK2837808A, a 650 Dalton compound in the quinoline 3-sulfonamide series (Billiard et al., 2013). These compounds potently inhibit activity of LDH A and B, as well as isoforms with mixtures of LDH A and B, making them attractive candidates for use in treatment of solid tumors that rely principally on aerobic glycolysis for survival.

The cryo-EM structure of the LDH B-GSK2837808A complex at an overall resolution of 2.8  $\text{\AA}$  (Figure S1A; Table 1) enables delineation of the binding site of the inhibitor (Figures 1A–1C). The location of the compound at the periphery of the protein is in excellent agreement with the placement of inhibitors in the two X-ray structures (4QSM and 4QT0 at resolutions of 3.0  $\text{\AA}$  and 3.2  $\text{\AA}$ , respectively) that are available for LDH A in complex with quinoline 3-sulfonamides (Figures S1B and S1C). The density for the inhibitor is not adequate to model in a specific conformation, which could be a consequence of high conformational flexibility coupled with its location at the outer edge of the protein with a high degree of solvent exposure. Nevertheless, most regions of the polypeptide display density that enables unambiguous construction of an atomic model for the protein (Figures 1D and 1E), including localization of numerous water molecules in the density map (Figures S1D and 1E). The fact that the overall quality and resolution of the structure is comparable to that obtained from X-ray crystallography validates the potential of cryo-EM to obtain near-atomic resolution structures of <150 kDa complexes of this kind, without the need for crystallization.

### Structural Studies of Isocitrate Dehydrogenase

The conversion of isocitrate to  $\alpha$ -ketoglutarate ( $\alpha$ KG) is catalyzed by NADP<sup>+</sup>-bound IDH1. IDH1 proteins with mutations at residue Arg<sup>132</sup> are found in cancer cells such as those in glioblastoma. This residue, most often found mutated to a histidine or a cysteine, is critical for the maintenance of the active site (Dang et al., 2009; Yang et al., 2010). The R132H and R132C variants of IDH1 lose affinity for isocitrate; instead, with the cofactor NADPH, these mutants bind  $\alpha$ KG and convert it to R(–)-2-hydroxyglutarate (2HG), an oncometabolite. ML309 is a potent and selective inhibitor of IDH1 R132H/C, decreasing production of 2HG in glioblastoma cells (Davis et al., 2014). Although a number of structures of the R132H mutant exist, structures of



**Figure 1. Cryo-EM Analysis of LDH B Bound to Inhibitor GSK2837808A**

(A) Cryo-EM density map of complex showing density for GSK2837808A (red) and selected residues in the binding pocket.

(B) Ribbon diagram of refined structure of LDH B showing the location of the bound inhibitor (marked by arrows) on the periphery of the tetramer.

(C) Structure of GSK2837808A.

(D and E) Visualization of side-chains in the cryo-EM density map in an  $\alpha$ -helical segment (D) and loop region (E) of the polypeptide.

See also Figure S1.

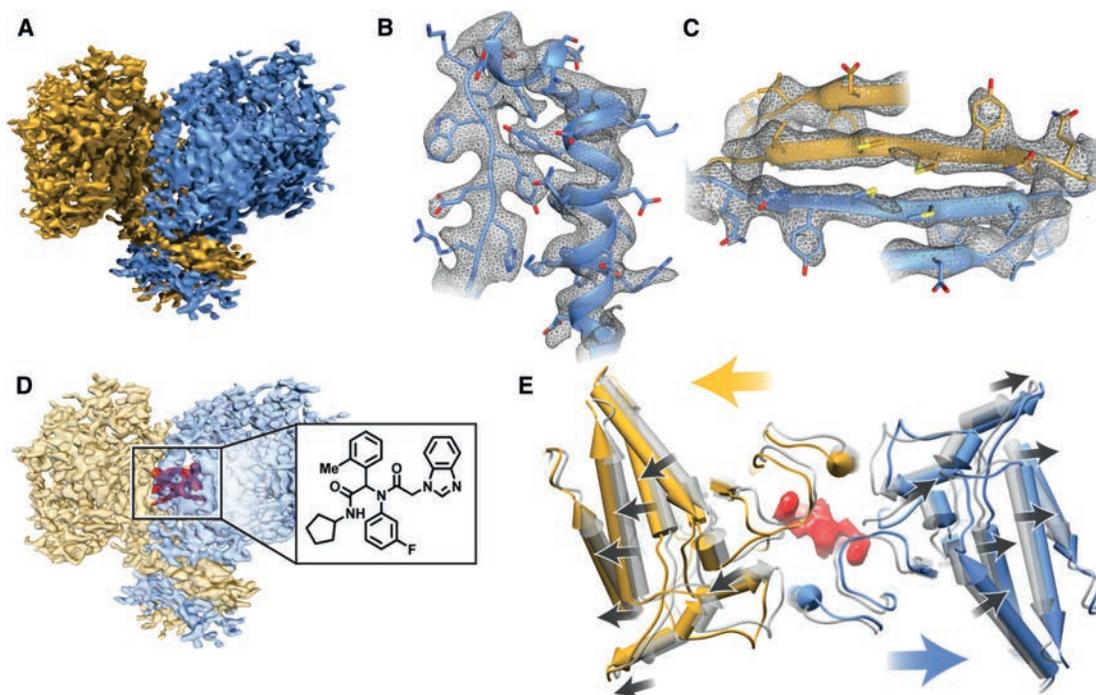
variable, with the expected gradient from regions that are in the well-packed interior at higher resolution, to those on the periphery with greater solvent exposure at lower resolution. Comparison of the structure of apo-IDH1 with that of ML309-bound IDH1 enables localization of ML309 in the structure (Figures 2A and 2D). The location of the inhibitor in the cryo-EM structure (Figure 2D) in a solvent-exposed region close to the 2-fold axis and near the catalytic site is similar to that seen for many IDH1-inhibitor complexes that have been amenable to crystallography (Deng et al., 2015; Okoye-Okafor et al., 2015; Wu et al., 2015; Zheng et al., 2013).

Despite the modest resolution of the maps, the structures provide new insights into structural aspects of IDH1 and its interaction with the inhibitor. Cryo-EM reconstructions carried out without imposition of 2-fold symmetry resulted in maps at a resolution of  $\sim 4.5$  Å, similar to those obtained with the use of 2-fold symmetry (Figure S2B), indicating that the structure of the IDH1 dimer in solution is symmetric at least at this resolution. This is a different result than that obtained from crystal structures of IDH1, which are mostly asymmetric with regions in the two protomers exhibiting different conformations (Xu et al., 2004; Yang et al., 2010). In our structures, both polypeptide chains that form the IDH1 dimer appear to be equivalent, suggesting that the apparent asymmetry is a likely consequence of crystal lattice formation and is not an intrinsic property of the dimer, as confirmed by inspection of lattice packing (Figure S3A). Further, comparison of the apo- and inhibitor-bound structures shows that ML309 binding essentially drives a wedge between the two protomers, leading to an outward movement of the peripheral domains (Figure 2E) that is evident even when comparing reconstructions obtained without imposition of symmetry (Movie S1). This structural change includes displacement

IDH Arg<sup>132</sup> mutants in complex with the ML309 inhibitor have not yet been determined. Structural analysis of IDH1 by cryo-EM is technically challenging given its small size, low symmetry (C<sub>2</sub>), and potential flexibility.

In Figure 2A, we present a cryo-EM structure for the R132C mutant of IDH1 that shows the characteristic butterfly fold of the dimer at a nominal resolution of 3.8 Å (Figure S2A). The polypeptide backbone and densities for the larger side chains are resolved in many regions (Figures 2B and 2C), enabling placement of the polypeptide chain into the density map using prior crystal structures of IDH1 as a guide. The quality of the map is

comparable to that of the apo-IDH1 structure (Figure S2A), with the expected gradient from regions that are in the well-packed interior at higher resolution, to those on the periphery with greater solvent exposure at lower resolution. Comparison of the structure of apo-IDH1 with that of ML309-bound IDH1 enables localization of ML309 in the structure (Figures 2A and 2D). The location of the inhibitor in the cryo-EM structure (Figure 2D) in a solvent-exposed region close to the 2-fold axis and near the catalytic site is similar to that seen for many IDH1-inhibitor complexes that have been amenable to crystallography (Deng et al., 2015; Okoye-Okafor et al., 2015; Wu et al., 2015; Zheng et al., 2013).



**Figure 2. Cryo-EM Analysis of IDH1 in the Absence and Presence of the Inhibitor ML309**

(A) Cryo-EM map of the apo-IDH1 (isocitrate dehydrogenase) dimer, colored by subunit.

(B and C) Selected regions of the IDH1 map demonstrating density for side-chains in an  $\alpha$ -helical region (B) and a  $\beta$  sheet region (C).

(D) Cryo-EM map of IDH1 in complex with ML309 showing density (red) for the inhibitor (inset) close to the dimer interface.

(E) Superposition of the apo- (gray) and ML309-bound (yellow/blue) IDH1 structures shows the outward movement of the subunits with ML309 binding (density shown in red). The black arrows indicate the direction of the changes in tertiary structure while the yellow and blue arrows show the overall movements at the level of quaternary structure.

See also Figures S2 and S3, and Movie S1.

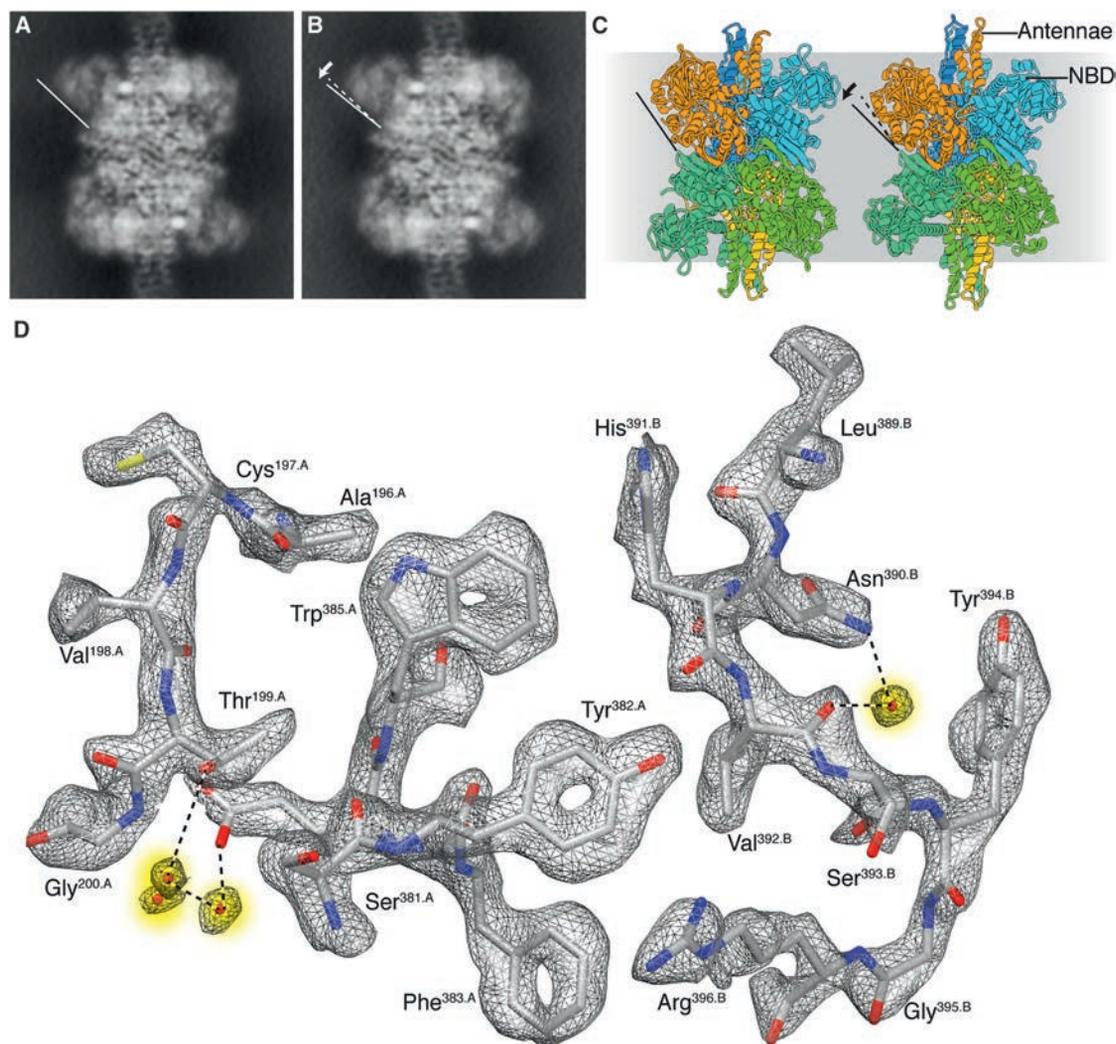
of a stretch of polypeptides spanning residues 270–281. The density for bound NADPH and the surrounding residues in the pocket is adequate for placement of the cofactor in the pocket (Figure S3B), but the low 3.8 Å resolution of the maps precludes identification of the detailed interactions within the binding pocket. Inspection of the asymmetrically reconstructed density map also enables the unambiguous identification of additional density consistent with the binding of a single molecule of ML309 at the dimer interface (Figure S3C), although its precise orientation cannot be determined. Nevertheless, the location of the catalytic site between the NADPH binding site and the site of bound ML309 potentially explains how ML309 binding can cause a profound change in the catalytic activity of the enzyme. These results show that cryo-EM density maps from small, <100 kDa dynamic proteins can not only help localize binding sites for small molecule inhibitors, but also potentially predict changes in activity that can complement information obtained from crystallographic studies.

#### Cryo-EM Analysis of Glutamate Dehydrogenase

Smaller size, lower symmetry, and conformational flexibility may all be contributing factors that limit resolution of the IDH1 structures obtained using cryo-EM methods. Intrinsic flexibility of proteins also poses a challenge for determining high-resolution structures even with X-ray crystallography. In many cases,

crystal contacts can freeze specific protein conformations, enabling structure determination at resolutions of 2 Å or better, but in other instances, local flexibility can be high enough to limit overall resolution. The cancer target and hexameric enzyme GDH is a good example of this problem. GDH displays closed and open conformations involving large movements of the nucleotide binding domain (NBD), but even when trapped in closed or open conformations, the highest resolutions reported so far from crystallographic studies and cryo-EM analyses of the open conformation are  $\sim 2.7$  Å (Smith et al., 2002) and  $\sim 3.3$  Å (Borgnia et al., 2016), respectively. We therefore used GDH to evaluate whether we could further optimize cryo-EM methods to achieve resolutions of  $\sim 2$  Å or better for the more ordered regions, despite the presence of significant conformational flexibility in other areas of the protein.

Using cryo-EM projection images that were selected for reduced beam-induced drift, we determined a cryo-EM structure for GDH in the open state using <22,000 molecular images (Figures 3, S4, S5; Table 1; Movie S3). Irrespective of whether the maps were divided into one, two, or more classes, the density maps displayed a clear difference between well-ordered central regions covering the central half of the complex and “fuzzy” outer regions for the rest, as illustrated by slices through projection images of the density map (Figures 3A and 3B). Based on comparison of the extent of movement determined by classifying



### Figure 3. Cryo-EM Analysis of GDH

(A and B) Projection views of two 3D classes from cryo-EM analysis of GDH (glutamate dehydrogenase). In both classes, there is well-defined density at the core, but it is weakly defined at the peripheral nucleotide binding domain (NBD). The two classes display similar structures in the interior but differ slightly in the peripheral NBD. The two classes are likely to be subsets of a continuum of states with varying orientations for the outer domain relative to the core.

(C) Ribbon diagrams of the open and closed structures (Borgnia et al., 2016) demonstrating the more extensive NBD movement associated with substrate binding and catalytic cleft closure.

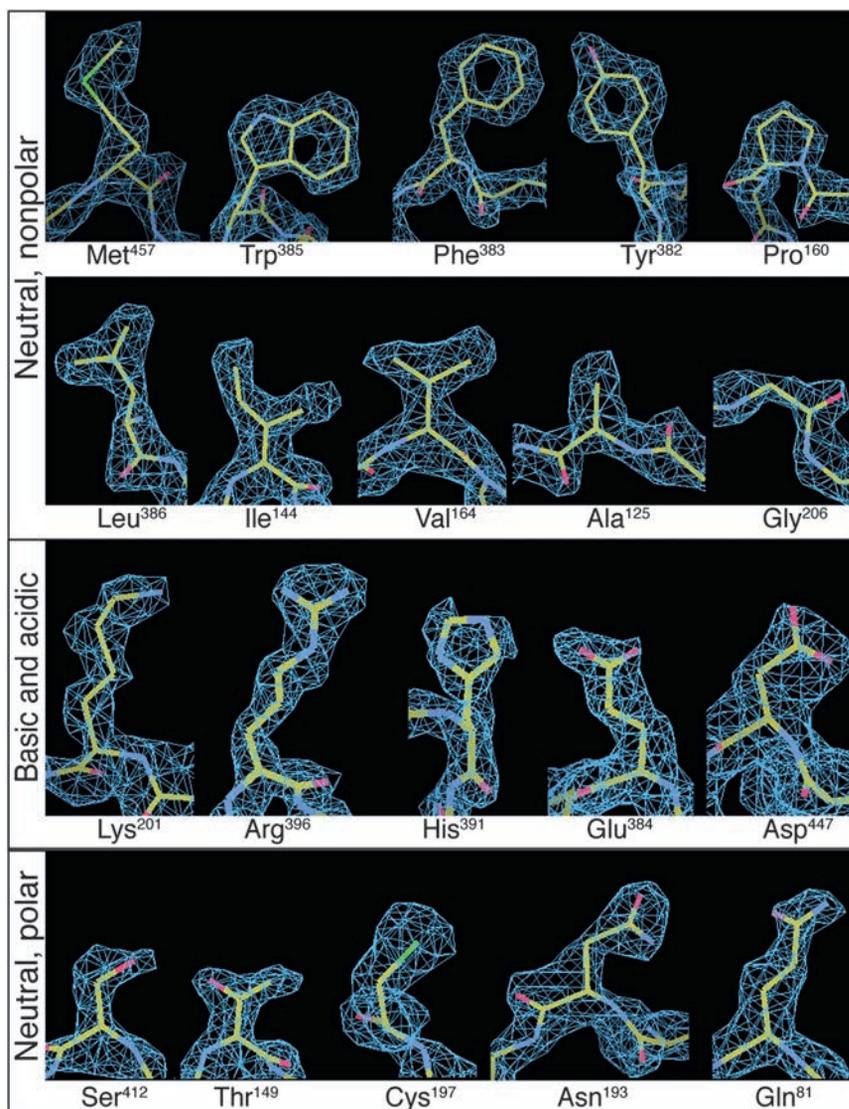
(D) A selected region of the cryo-EM map of the GDH structure, highlighting high-resolution features such as “holes” in the aromatic rings of Tyr<sup>382</sup>, Phe<sup>383</sup>, and Trp<sup>385</sup>, water molecules (shaded yellow), and well-resolved densities for carbonyl bonds.

See also Figures 4 and S4 and Movie S2.

the data to produce two discrete 3D structures, the range of motion in the peripheral NBD (Figures 3A and 3B; Movie S2) is seen to be small in comparison to the large domain movements observed for transition from open to closed conformations (Figure 3C) (Borgnia et al., 2016).

The binding of GTP (an inhibitor) and ADP (an activator) to two allosteric sites in each protomer modulates the transition between “closed” and “open” states of GDH. These allosteric modulators tightly control GDH function *in vivo*. The two states we present in Figures 3A and 3B are both in the “open” conformation with differences likely due to small perturbations arising from the unhindered sampling of conformations due to the

twisting motions of the pivot helix in the absence of either ADP or GTP in solution. Despite the presence of the flexible outer domains (Figures 3A and 3B), most of the interior of GDH can nevertheless be visualized at resolutions better than 2 Å, well beyond the 2.7 Å resolution obtained with X-ray crystallographic analysis, in which the flexible regions presumably limit the order of the crystals obtained. In comparison to our earlier work with  $\beta$ -galactosidase (Bartesaghi et al., 2015), in which we reported an average resolution of 2.2 Å, the density map for GDH at 1.8 Å shows better-defined and more uniform density over most of the ordered central core, spanning the subunit interface (Figure 3D), as well as clear density for carbonyl bonds,



**Figure 4. Density Representations for Each of the 20 Amino Acid Types from the 1.8 Å Resolution Cryo-EM Structure of apo-GDH**

Features such as holes in aromatic rings, as well as the “zigzag” structure of extended side-chains such as Arg and Lys, are visible in the density maps. See also Figure S5 and Movie S3.

fraction of the protein that is conformationally mobile may preclude achieving atomic resolution for the more structured or stable regions of the complex. The fact that we can visualize densities for each of the 20 amino acids in the GDH cryo-EM map (Figure 4), without 3D masking during refinement, addresses this question definitively while also enabling the crossing of the 2 Å resolution mark in single particle cryo-EM. These cryo-EM densities are easily comparable to those obtained by X-ray crystallography at a similar resolution (Figure S5).

The maps we have obtained for LDH display 2.8 Å resolution in the protein component and densities for numerous water molecules (Figures S1D and S1E), but density for the bound ligand is at lower resolution. One reason that the density for the bound ligand is less well-defined than that of the polypeptide may be because small molecule ligands bound to the protein periphery are likely to sample far more conformations in aqueous solution than what is possible in the context of a crystal lattice. Most solvents used for protein crystallization invariably include additives such as PEG that dehydrate the crystal by removing

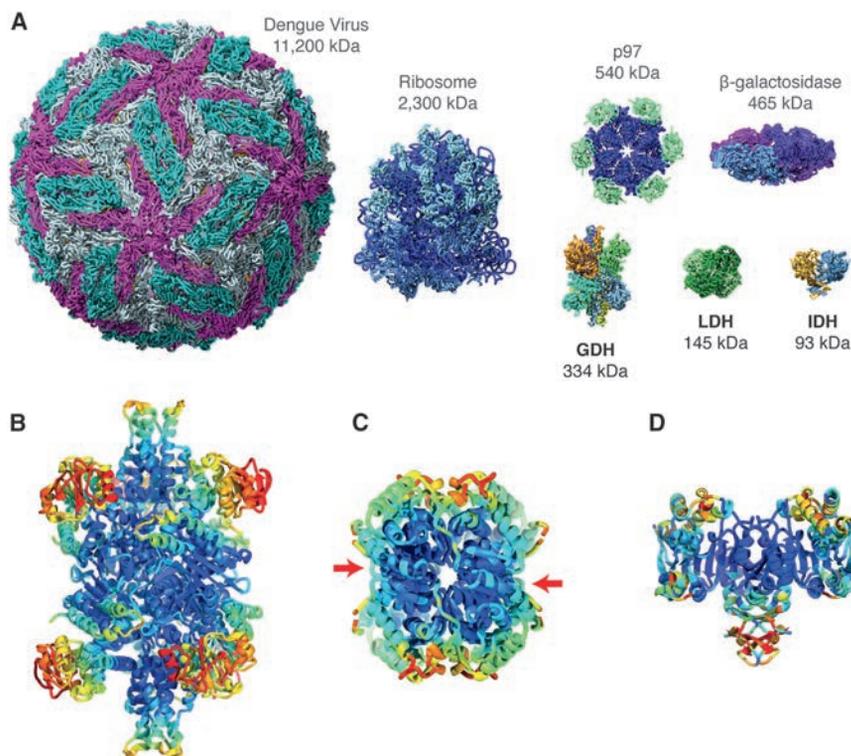
several H-bonded water molecules and well-defined density for extended Lys side chains (Figures S4B–S4E).

## DISCUSSION

We have previously shown that it is possible to achieve atomic resolution information from an inhibitor-bound protein complex such as the 465 kDa  $\beta$ -galactosidase (Banerjee et al., 2016; Bartesaghi et al., 2015). The inhibitor-bound state of  $\beta$ -galactosidase is relatively stable, with only a small gradient in temperature factor between the central and peripheral regions, making it possible to achieve high resolution across the breadth of the protein. However, many proteins can have flexible domains, as in the case of GDH, even when they are expected to be in a single conformational state. While cryo-EM can be used to effectively mask out flexible or disordered areas of a protein during steps in image processing (Amunts et al., 2014; Bai et al., 2015a), it has remained an open question whether the presence of a large

excess solvent and tightening the packing of protein molecules while reducing the size of the solvent channels around the protein. As a result, these additives improve crystal order and diffraction resolution, but at the same time, also affect ligand-protein interactions, as dehydration reduces the number of water molecules in the binding pocket and may minimize mobility of the ligand. Thus, cryo-EM studies in aqueous dispersions and X-ray crystallographic studies in ordered 3D crystals provide potentially complementary snapshots of ligand-protein interactions that are likely to occur in cellular environments.

Finding the correct orientation of each projection image is central for the success of the single particle cryo-EM reconstruction process. The practical size limitation for high-resolution structure determination is currently a question of considerable interest within the cryo-EM field. Figure 5 provides a comparison of the relative sizes of some complexes whose structures have been studied at near-atomic resolution, including those reported in the present study. Given the present state of detector and



**Figure 5. Relative Sizes of Small Metabolic Complexes Compared to Other Structures Solved by Cryo-EM and Variations in Flexibility across Different Regions**

(A) Relative sizes of GDH, LDH B, and IDH1 in comparison to a representative set of other structures that have been solved by cryo-EM to high resolution including icosahedral viruses, ribosomes, and protein complexes such as  $\beta$ -galactosidase and the AAA ATPase p97.

(B–D) Ribbon diagram of hexameric GDH (B), tetrameric LDH B (C), and dimeric IDH1 (D), colored to show B-factor variation from blue (lowest) to red (highest), respectively.

See also Movie S3.

microscope technology, a figure of 150 kDa is thought to be a practical lower limit even for medium resolutions (Thompson et al., 2016), even though theoretical arguments have been advanced for why much smaller sizes should be accessible (Glaeser and Hall, 2011; Henderson, 1995). As instrumentation and methods for data analysis have improved, our ability to overcome the handicaps of small size, low symmetry, and high flexibility have likewise improved, allowing the reconstruction of not just large, highly symmetric complexes such as icosahedral viruses, but now also of smaller, lower symmetry complexes (Figure 5A). In this study, we have shown that not only is it possible to determine the structure of a protein smaller than 150kDa to  $\sim 3$  Å resolution, it is also possible to determine the structure of a protein  $<100$  kDa to near-atomic ( $<4$  Å) resolution, coupled with the visualization of the location of bound small-molecule inhibitors in both cases. From an analysis of all of the parameters that influenced final map quality, our conclusion is that selecting for specimens with thin ice, retaining only those images that display the lowest beam-induced motion, and using a sufficiently small pixel size and the improved XP sensor with higher DQE performance were the key factors that enabled us to obtain high resolution with these smaller protein complexes.

As more examples of cryo-EM structures are reported, there are beginning to be noticeable patterns in the differences between the visual appearance of cryo-EM density maps and those obtained using X-ray crystallography. In visualizing cryo-EM density maps, different regions of the map are typically explored by implementing B-factor sharpening specific to the different segments of the map, but deposited maps and figures generally show density maps sharpened with a single value for B-factor for

the entire protein, which poses a challenge for capturing all of the information in a single view. Nevertheless, it is clear that in all three of the protein complexes that we have analyzed, the outer regions are significantly more flexible than the core of the complex (Figures 5B–5D).

Since cryo-EM analyses report on the structure of the proteins in solution, they can provide information on quaternary structures complementary to that obtained from crystal structures, which

may be influenced by contacts within the crystal lattice. A common feature of the cryo-EM density maps we have observed in structures of complexes of ligands with diverse proteins such as LDH, IDH1,  $\beta$ -galactosidase, and p97 is that the density for bound ligands, especially in the peripheral regions, is generally somewhat weaker and at lower resolution than that of residues in the binding pocket. As noted earlier in the discussion, this is perhaps not surprising given the presence of numerous single bonds around which rotations are possible, coupled with poor steric and electrostatic fits between ligand and protein. Even in regions where the density map of the polypeptide is poorly defined in small proteins such as IDH1, crystallographic information can provide an invaluable framework to interpret the cryo-EM structures, enabling identification of sites of inhibitor binding.

Size, symmetry, and intrinsic flexibility are all parameters that play a role in determining the resolutions that can be achieved using cryo-EM. LDH and  $\beta$ -galactosidase are both tetramers with the same symmetry (D<sub>2</sub>), but LDH is about a third of the size of  $\beta$ -galactosidase. The lower resolutions currently achieved with LDH are most likely from the reduced accuracy of alignment of projection images, and this comparison is thus a good way to assess effect of size on resolution with presently available technology. The further drop in resolution observed for IDH1 could be due to reduction in size, lower symmetry, and potentially greater flexibility of the dimeric enzyme. Reconstruction of IDH1 without application of symmetry resulted in only a slightly lower resolution ( $\sim 4.5$  Å), indicating that the trends we observe are likely general in nature. It is possible that, for these small enzymes, improvements in resolution could be obtained by the use of one or two bound Fab fragments (Wu et al., 2012). Fab fragments not only

increase the size of a complex, but also assist with orientation assignment and can stabilize certain conformations of highly flexible complexes (Kim et al., 2015; Lyumkis et al., 2013b). Since Fab fragments are ~50 kDa in size, an important consequence of our results is that the presence of one or two bound Fab fragments may bring almost any protein into a size range suitable for analysis by cryo-EM. Although we have not yet achieved ~2 Å resolution with small complexes such as IDH and LDH, it is likely that further advances in detector technology will contribute toward achieving this goal. Ongoing hardware developments such as phase plates (Danev and Baumeister, 2016), as well as improvements in software for image alignment and reconstruction of heterogeneous complexes, will also undoubtedly be necessary to advance cryo-EM frontiers to determine high resolution structures of even smaller and more challenging complexes.

## EXPERIMENTAL PROCEDURES

### Proteins and Grid Preparation

Chicken heart LDH B (catalog #59747, Sigma-Aldrich, St. Louis) lyophilized sample was suspended in 550  $\mu$ l 1 $\times$  PBS buffer and centrifuged at 20,000  $\times$  *g* for 30 min. Thereafter, it was subjected to gel filtration on a Superdex-200 size-exclusion chromatography column connected to a ÄKTA FPLC apparatus (GE Healthcare Bio-Sciences, Piscataway) with an elution buffer comprised of 1 $\times$  PBS. The ligand GSK 2837808A (catalog #5189 Tocris Bioscience) was dissolved in 100% DMSO and added to LDH (~1.5 mg/ml) to a final inhibitor concentration of 75  $\mu$ M and 0.5% DMSO.

Human IDH1 R132C mutant protein was purified following the published protocol (Davis et al., 2014). ML309 was dissolved in 100% DMSO and added to IDH1 (2.8 mg/ml) to a final inhibitor concentration of 50  $\mu$ M and 0.5% DMSO.

Bovine GDH (catalog #G2626, Sigma-Aldrich, St. Louis) was dialyzed overnight against gel-filtration buffer (100 mM potassium phosphate [pH 6.8]) prior to purification by size-exclusion chromatography (SEC) using a Superdex-200 size-exclusion chromatography column connected to a ÄKTA FPLC apparatus (GE Healthcare Bio-Sciences, Piscataway). Final protein concentration was ~3 mg/ml.

Prior to preparation of frozen-hydrated specimens, 3–4  $\mu$ l of a freshly thawed aliquot of purified proteins or inhibitor-protein complexes were deposited on Quantifoil R1.2/1/3 grids and plunge frozen in liquid ethane cooled by liquid nitrogen using either a Leica plunge-freeze device (for the LDH and IDH1 samples) or a Vitrobot (for the GDH samples). For experiments with added inhibitor, the protein was pre-incubated for 30 min at room temperature under the buffer conditions indicated above for each kind of protein complex.

### Data Acquisition

The grids were imaged using a Titan Krios transmission electron microscope (FEI Company, Hillsboro) aligned for parallel illumination and operated at 300 kV, with the specimen maintained at liquid nitrogen temperatures. Images were recorded on a K2 Summit camera equipped with the XP sensor (Gatan, Pleasanton) operated in super-resolution counting mode, placed at the end of a GIF Quantum energy filter (Gatan), functioning in zero-energy-loss mode with a slit width of 20 eV as described previously (Bartesaghi et al., 2015). Images were typically collected with a defocus range between –0.7 and –3.0  $\mu$ m. For LDH and IDH, a dose rate of ~5  $e^-/\text{Å}^2/\text{s}$  (at the specimen plane) was used with a total exposure time of 12 s. Intermediate frames were recorded every 0.2 s giving an accumulated dose of ~60  $e^-/\text{Å}^2$  and a total of 60 frames per image. Pixel size was set to 0.495 Å. For GDH, a dose rate of ~2.6  $e^-/\text{Å}^2/\text{s}$  was used and a total exposure time of 15.2 s. Intermediate frames were recorded every 0.4 s giving an accumulated dose of ~40  $e^-/\text{Å}^2$  and a total of 38 frames per image with a pixel size of 0.637 Å.

### Image Processing

Movie frame alignment and contrast transfer function (CTF) estimation for each micrograph were carried out as described previously (Bartesaghi et al., 2014).

For the LDH B and IDH1 reconstructions, particles were picked from the original micrographs using a Gaussian disk of 42 Å in radius as the search template. Particles were extracted using a binning factor of 8 and a box size of 96 pixels and subsequently subjected to 12 rounds of refinement with the program Frealign (Lyumkis et al., 2013a). The numbers of particle images used for each final reconstruction are shown in Table 1. The subset of particles that contributed to the reconstruction in the last iteration was re-extracted from the original micrographs using a binning factor of 4 and a box size of 192 pixels followed by eight additional iterations of local refinement in Frealign. In a similar way, particles were re-extracted and further refined two additional times using a binning factor of 2 and a box size of 768 pixels, first using the full exposure and then using only the first half of the exposure (frames 1–30), corresponding to ~30  $e^-/\text{Å}^2$ . B-factor sharpening was applied to the final maps for purposes of visualization (Table 1).

Images of GDH were processed in a way similar to that done for LDH B and IDH1, this time using a Gaussian disk of 75 Å in radius as the search template for particle picking. The selection of particles that contributed to the reconstruction at each iteration was done according to the bi-modal distribution of Frealign scores (Banerjee et al., 2016). Re-extraction of selected particles from the original micrographs using progressively less binning followed by local refinement was done as described above for the LDH B and IDH1 datasets. In addition, particles with a binning factor of 2 and a box size of 768 pixels were re-extracted one final time using frames 3–9 and used to produce the final reconstruction by applying the latest set of alignments (derived using the first half of the exposure) without further refinement. A B-factor of –90 Å<sup>2</sup> was applied to the final map for purposes of visualization.

To estimate resolution, we used two approaches, both of which provided mutually consistent results. In one approach, we used the strategy implemented in Frealign to carry out refinement using only low-resolution information (to prevent over-fitting) and then measured the value at which the FSC value drops to 0.143. The highest resolution we used for refining particle orientations was 6 Å for the IDH1 datasets and 4 Å for the LDH B and GDH datasets. We also computed the FSC curve between the experimentally derived map and the refined model, and determined the value at which the correlation drops to 0.5.

### Atomic Model Refinement

The starting model used for the refinement of the LDH coordinates was the X-ray structure 110Z. Changes were made to the initial model in order to incorporate the differences between the human and chicken LDH amino acid sequences. The structure of the apo LDH B was initially refined by rigid-body refinement followed by real-space refinement using the program Phenix (Adams et al., 2010). Manual adjustments were made using the program COOT (Emsley et al., 2010). Solvent molecules were added manually followed by a round of real-space refinement.

The starting model used for the refinement of NADPH-bound IDH1 was from the coordinates from the structure 3MAP. Prior to refinement, the residue Arg<sup>132</sup> was mutated to a cysteine and the isocitrate molecule was removed from the model. The structure of the NADPH-bound IDH1 was initially refined by rigid-body refinement followed by real-space refinement using the program Phenix. Manual adjustments were made by refinement in COOT following another round of real-space refinement. For the ML-309-bound structure, the starting model was the NADPH-bound IDH1 obtained using the cryo-EM data. The refinement protocol was identical to that used for LDH B.

The PDB coordinates 3JCZ served as a starting point for the refinement of apo-GDH. Initially, rigid-body fitting followed by real-space refinement using the program Phenix was used to refine the model. Manual adjustments were made in COOT. Solvent molecules were added automatically using stringent thresholds followed by manual inspection. Thereafter, the model was refined again using real-space refinement as implemented in Phenix.

### ACCESSION NUMBERS

The accession numbers for the density maps reported in this paper are Electron Microscopy Data Bank: 8191, 8192, 8193, and 8194 for LDHB, apo- and

ML309-bound IDH1, and GDH, respectively. Each entry includes uncorrected and B-factor sharpened maps, as well as the asymmetric reconstructions where applicable. The accession numbers for the corresponding refined atomic models reported in this paper are PDB: 5K0Z, 5K10, 5K11, and 5K12.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and three movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.05.040>.

## AUTHOR CONTRIBUTIONS

Specimen preparation was carried out by A.M. and P.R., data collection was carried out by A.M., image processing was carried out by A.B., data analysis, interpretation, and presentation was carried out by S.B., A.B., A.M., V.F., L.A.E., J.L.S.M., and S.S., and samples of purified IDH1 and ML309 were provided by M.I.D., R.P., and M.B.B. Final figure preparation was carried out primarily by V.F. with help from A.M. and A.B. Project conception, organization, and supervision was provided by S.S., who was also primarily responsible for writing the manuscript, with help from S.B., A.B., A.M., L.A.E., V.F., and J.L.S.M.

## ACKNOWLEDGMENTS

This research was supported by funds from the Center for Cancer Research, National Cancer Institute, NIH in Bethesda. We thank Kieran Moynihan, Robert Mueller, and Joe Cometa for technical assistance with electron microscopy; Fred Ulmer, Paul Mooney, and Chris Booth for advice and assistance with optimizing K2 detector performance; and Rishub Jain, Jierui Fang, and Amy Jin for assistance with data analysis. This work utilized the computational resources of the NIH HPC Biowulf cluster (<http://hpc.nih.gov>).

Received: April 29, 2016

Revised: May 9, 2016

Accepted: May 9, 2016

Published: May 26, 2016

## REFERENCES

- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221.
- Amunts, A., Brown, A., Bai, X.C., Llácer, J.L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S.H., and Ramakrishnan, V. (2014). Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485–1489.
- Bai, X.C., Rajendra, E., Yang, G., Shi, Y., and Scheres, S.H. (2015a). Sampling the conformational space of the catalytic subunit of human  $\gamma$ -secretase. *eLife* **4**, e11182.
- Bai, X.C., Yan, C., Yang, G., Lu, P., Ma, D., Sun, L., Zhou, R., Scheres, S.H., and Shi, Y. (2015b). An atomic structure of human  $\gamma$ -secretase. *Nature* **525**, 212–217.
- Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S.L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R.J., Wipf, P., et al. (2016). 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science* **351**, 871–875.
- Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A., and Subramaniam, S. (2014). Structure of  $\beta$ -galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci. USA* **111**, 11709–11714.
- Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J.L., and Subramaniam, S. (2015). 2.2 Å resolution cryo-EM structure of  $\beta$ -galactosidase in complex with a cell-permeant inhibitor. *Science* **348**, 1147–1151.
- Belnap, D.M. (2015). Electron Microscopy and Image Processing: Essential Tools for Structural Analysis of Macromolecules. *Curr. Protoc. Protein Sci.* **82**, 17.2.1–17.2.61.
- Billiard, J., Dennison, J.B., Briand, J., Annan, R.S., Chai, D., Colón, M., Dodson, C.S., Gilbert, S.A., Greshock, J., Jing, J., et al. (2013). Quinoline 3-sulfonamides inhibit lactate dehydrogenase A and reverse aerobic glycolysis in cancer cells. *Cancer Metab.* **1**, 19.
- Borgnia, M.J., Banerjee, S., Merk, A., Matthies, D., Bartesaghi, A., Rao, P., Pierson, J., Earl, L.A., Falconieri, V., Subramaniam, S., and Milne, J.L. (2016). Using Cryo-EM to Map Small Ligands on Dynamic Metabolic Enzymes: Studies with Glutamate Dehydrogenase. *Mol. Pharmacol.* **89**, 645–651.
- Cheng, Y. (2015). Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **161**, 450–457.
- Danev, R., and Baumeister, W. (2016). Cryo-EM single particle analysis with the Volta phase plate. *eLife* **5**, e13046.
- Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., et al. (2009). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744.
- Davis, M.I., Gross, S., Shen, M., Straley, K.S., Pragani, R., Lea, W.A., Popovici-Muller, J., DeLaBarre, B., Artin, E., Thorne, N., et al. (2014). Biochemical, cellular, and biophysical characterization of a potent inhibitor of mutant isocitrate dehydrogenase IDH1. *J. Biol. Chem.* **289**, 13717–13725.
- Deng, G., Shen, J., Yin, M., McManus, J., Mathieu, M., Gee, P., He, T., Shi, C., Bedel, O., McLean, L.R., et al. (2015). Selective inhibition of mutant isocitrate dehydrogenase 1 (IDH1) via disruption of a metal binding network by an allosteric small molecule. *J. Biol. Chem.* **290**, 762–774.
- Dragovich, P.S., Fauber, B.P., Boggs, J., Chen, J., Corson, L.B., Ding, C.Z., Eigenbrot, C., Ge, H., Giannetti, A.M., Hunsaker, T., et al. (2014). Identification of substituted 3-hydroxy-2-mercaptocyclohex-2-enones as potent inhibitors of human lactate dehydrogenase. *Bioorg. Med. Chem. Lett.* **24**, 3764–3771.
- Du, J., Lü, W., Wu, S., Cheng, Y., and Gouaux, E. (2015). Glycine receptor mechanism elucidated by electron cryo-microscopy. *Nature* **526**, 224–229.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501.
- Fischer, N., Neumann, P., Konevega, A.L., Bock, L.V., Ficner, R., Rodnina, M.V., and Stark, H. (2015). Structure of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* **520**, 567–570.
- Fiume, L., Manerba, M., Vettraino, M., and Di Stefano, G. (2014). Inhibition of lactate dehydrogenase activity as an approach to cancer therapy. *Future Med. Chem.* **6**, 429–445.
- Ge, P., and Zhou, Z.H. (2011). Hydrogen-bonding networks and RNA bases revealed by cryo electron microscopy suggest a triggering mechanism for calcium switches. *Proc. Natl. Acad. Sci. USA* **108**, 9637–9642.
- Glaeser, R.M., and Hall, R.J. (2011). Reaching the information limit in cryo-EM of biological macromolecules: experimental aspects. *Biophys. J.* **100**, 2331–2337.
- Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28**, 171–193.
- Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E., and Downing, K.H. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J. Mol. Biol.* **213**, 899–929.
- Henderson, R., Chen, S., Chen, J.Z., Grigorieff, N., Passmore, L.A., Ciccarelli, L., Rubinstein, J.L., Crowther, R.A., Stewart, P.L., and Rosenthal, P.B. (2011). Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J. Mol. Biol.* **413**, 1028–1046.
- Jomaa, A., Boehringer, D., Leibundgut, M., and Ban, N. (2016). Structures of the E. coli translating ribosome with SRP and its receptor and with the translocon. *Nat. Commun.* **7**, 10471.
- Kim, J., Wu, S., Tomasiak, T.M., Mergel, C., Winter, M.B., Stiller, S.B., Robles-Colmanares, Y., Stroud, R.M., Tampé, R., Craik, C.S., and Cheng, Y. (2015). Subnanometre-resolution electron cryomicroscopy structure of a heterodimeric ABC exporter. *Nature* **517**, 396–400.
- Liao, M., Cao, E., Julius, D., and Cheng, Y. (2013). Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107–112.

- Lyumkis, D., Brilot, A.F., Theobald, D.L., and Grigorieff, N. (2013a). Likelihood-based classification of cryo-EM images using FREALIGN. *J. Struct. Biol.* *183*, 377–388.
- Lyumkis, D., Julien, J.P., de Val, N., Cupo, A., Potter, C.S., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., Carragher, B., et al. (2013b). Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* *342*, 1484–1490.
- Matadeen, R., Patwardhan, A., Gowen, B., Orlova, E.V., Pape, T., Cuff, M., Mueller, F., Brimacombe, R., and van Heel, M. (1999). The Escherichia coli large ribosomal subunit at 7.5 Å resolution. *Structure* *7*, 1575–1583.
- Matthies, D., Dalmás, O., Borgnia, M.J., Dominik, P.K., Merk, A., Rao, P., Reddy, B.G., Islam, S., Bartesaghi, A., Perozo, E., and Subramaniam, S. (2016). Cryo-EM Structures of the Magnesium Channel CorA Reveal Symmetry Break upon Gating. *Cell* *164*, 747–756.
- McClelland, M.L., Adler, A.S., Deming, L., Cosino, E., Lee, L., Blackwood, E.M., Solon, M., Tao, J., Li, L., Shames, D., et al. (2013). Lactate dehydrogenase B is required for the growth of KRAS-dependent lung adenocarcinomas. *Clin. Cancer Res.* *19*, 773–784.
- McMullan, G., Faruqi, A.R., Clare, D., and Henderson, R. (2014). Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy* *147*, 156–163.
- Okoye-Okafor, U.C., Bartholdy, B., Cartier, J., Gao, E.N., Pietrak, B., Rendina, A.R., Rominger, C., Quinn, C., Smallwood, A., Wiggall, K.J., et al. (2015). New IDH1 mutant inhibitors for treatment of acute myeloid leukemia. *Nat. Chem. Biol.* *11*, 878–886.
- Rawat, U.B., Zavialov, A.V., Sengupta, J., Valle, M., Grassucci, R.A., Linde, J., Vestergaard, B., Ehrenberg, M., and Frank, J. (2003). A cryo-electron microscopic study of ribosome-bound termination factor RF2. *Nature* *421*, 87–90.
- Rodriguez, S., Jafer, O., Goker, H., Summersgill, B.M., Zafarana, G., Gillis, A.J., van Gurp, R.J., Oosterhuis, J.W., Lu, Y.J., Huddart, R., et al. (2003). Expression profile of genes from 12p in testicular germ cell tumors of adolescents and adults associated with i(12p) and amplification at 12p11.2-p12.1. *Oncogene* *22*, 1880–1891.
- Settembre, E.C., Chen, J.Z., Dormitzer, P.R., Grigorieff, N., and Harrison, S.C. (2011). Atomic model of an infectious rotavirus particle. *EMBO J.* *30*, 408–416.
- Skiniotis, G., and Southworth, D.R. (2016). Single-particle cryo-electron microscopy of macromolecular complexes. *Microscopy (Oxf.)* *65*, 9–22.
- Smith, T.J., Schmidt, T., Fang, J., Wu, J., Siuzdak, G., and Stanley, C.A. (2002). The structure of apo human glutamate dehydrogenase details subunit communication and allostery. *J. Mol. Biol.* *318*, 765–777.
- Thompson, R.F., Walker, M., Siebert, C.A., Muench, S.P., and Ranson, N.A. (2016). An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods* *100*, 3–15.
- Wong, W., Bai, X.C., Brown, A., Fernandez, I.S., Hanssen, E., Condrón, M., Tan, Y.H., Baum, J., and Scheres, S.H. (2014). Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. *eLife* *3*, e03080.
- Wu, S., Avila-Sakar, A., Kim, J., Booth, D.S., Greenberg, C.H., Rossi, A., Liao, M., Li, X., Alian, A., Griner, S.L., et al. (2012). Fabs enable single particle cryoEM studies of small proteins. *Structure* *20*, 582–592.
- Wu, F., Jiang, H., Zheng, B., Kogiso, M., Yao, Y., Zhou, C., Li, X.N., and Song, Y. (2015). Inhibition of Cancer-Associated Mutant Isocitrate Dehydrogenases by 2-Thiohydantoin Compounds. *J. Med. Chem.* *58*, 6899–6908.
- Xu, X., Zhao, J., Xu, Z., Peng, B., Huang, Q., Arnold, E., and Ding, J. (2004). Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *J. Biol. Chem.* *279*, 33946–33957.
- Yang, B., Zhong, C., Peng, Y., Lai, Z., and Ding, J. (2010). Molecular mechanisms of “off-on switch” of activities of human IDH1 by tumor-associated mutation R132H. *Cell Res.* *20*, 1188–1200.
- Yu, X., Jin, L., and Zhou, Z.H. (2008). 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature* *453*, 415–419.
- Zhang, X., Jin, L., Fang, Q., Hui, W.H., and Zhou, Z.H. (2010). 3.3 Å cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell* *141*, 472–482.
- Zheng, B., Yao, Y., Liu, Z., Deng, L., Anglin, J.L., Jiang, H., Prasad, B.V., and Song, Y. (2013). Crystallographic Investigation and Selective Inhibition of Mutant Isocitrate Dehydrogenase. *ACS Med. Chem. Lett.* *4*, 542–546.

# Programmable RNA Tracking in Live Cells with CRISPR/Cas9

David A. Nelles,<sup>1,2</sup> Mark Y. Fang,<sup>1</sup> Mitchell R. O'Connell,<sup>4</sup> Jia L. Xu,<sup>1</sup> Sebastian J. Markmiller,<sup>1</sup> Jennifer A. Doudna,<sup>4,5,6</sup> and Gene W. Yeo<sup>1,2,3,\*</sup>

<sup>1</sup>Department of Cellular and Molecular Medicine and Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA 92037, USA

<sup>2</sup>Materials Science and Engineering Graduate Program, University of California, San Diego, La Jolla, CA 92093, USA

<sup>3</sup>Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 1190777, Singapore

<sup>4</sup>Department of Molecular and Cell Biology and Center for RNA Systems Biology, University of California, Berkeley, Berkeley, CA 94720, USA

<sup>5</sup>Department of Chemistry, Innovative Genomics Initiative and Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA 94720, USA

<sup>6</sup>Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, Berkeley, CA 94720, USA

\*Correspondence: geneyeo@ucsd.edu

<http://dx.doi.org/10.1016/j.cell.2016.02.054>

## SUMMARY

RNA-programmed genome editing using CRISPR/Cas9 from *Streptococcus pyogenes* has enabled rapid and accessible alteration of specific genomic loci in many organisms. A flexible means to target RNA would allow alteration and imaging of endogenous RNA transcripts analogous to CRISPR/Cas-based genomic tools, but most RNA targeting methods rely on incorporation of exogenous tags. Here, we demonstrate that nuclease-inactive *S. pyogenes* CRISPR/Cas9 can bind RNA in a nucleic-acid-programmed manner and allow endogenous RNA tracking in living cells. We show that nuclear-localized RNA-targeting Cas9 (RCas9) is exported to the cytoplasm only in the presence of sgRNAs targeting mRNA and observe accumulation of *ACTB*, *CCNA2*, and *TFRC* mRNAs in RNA granules that correlate with fluorescence in situ hybridization. We also demonstrate time-resolved measurements of *ACTB* mRNA trafficking to stress granules. Our results establish RCas9 as a means to track RNA in living cells in a programmable manner without genetically encoded tags.

## INTRODUCTION

Clustered regularly-interspaced short palindromic repeats (CRISPRs) form the basis of adaptive immune systems in bacteria and archaea by encoding CRISPR RNAs that guide CRISPR-associated (Cas) nucleases to invading genetic material (Wiedenheft et al., 2012). Cas9 from the type II CRISPR system of *S. pyogenes* has been repurposed for genome engineering in eukaryotic organisms (Hwang et al., 2013; Li et al., 2013a; Mali et al., 2013; Nakayama et al., 2013; Sander and Joung, 2014; Yang et al., 2014) and is rapidly proving to be an efficient means

of DNA targeting for other applications such as gene expression modulation (Qi et al., 2013) and imaging (Chen et al., 2013). Cas9 and its associated single-guide RNA (sgRNA) require two critical features to target DNA: a short DNA sequence of the form 5'-NGG-3' (where "N" = any nucleotide) known as the protospacer adjacent motif (PAM) and an adjacent sequence on the opposite DNA strand that is antisense to the sgRNA. By supporting DNA recognition with specificity determined entirely by a short spacer sequence within the sgRNA, CRISPR/Cas9 provides uniquely flexible and accessible manipulation of the genome. Manipulating cellular RNA content, in contrast, remains problematic. Whereas there exist robust means of attenuating gene expression via RNAi and antisense oligonucleotides, other critical aspects of post-transcriptional gene expression regulation such as subcellular trafficking, alternative splicing or polyadenylation, and spatiotemporally restricted translation are difficult to measure in living cells and are largely intractable.

Analogous to the assembly of zinc finger nucleases (Urnov et al., 2010) and transcription activator-like effector nucleases (TALEN) to recognize specific DNA sequences, efforts to recognize specific RNA sequences have focused on engineered RNA-binding domains. Pumilio and FBF homology (PUF) proteins carry well-defined modules capable of recognizing a single base each and have supported successful targeting of a handful of transcripts for imaging and other manipulations (Filipovska et al., 2011; Ozawa et al., 2007; Wang et al., 2009). PUF proteins can be fused to arbitrary effector domains to alter or tag target RNAs, but PUFs must be redesigned and validated for each RNA target and can only recognize eight contiguous bases, which does not allow unique discrimination in the transcriptome. Molecular beacons are self-quenched synthetic oligonucleotides that fluoresce upon binding to target RNAs and allow RNA detection without construction of a target-specific protein (Sokol et al., 1998). But molecular beacons must be microinjected to avoid the generation of excessive background signal associated with endosome-trapped probes and are limited to imaging applications. An alternative approach to recognition of RNA substrates is to introduce RNA aptamers into target RNAs, enabling specific and strong association of cognate aptamer-binding proteins

such as the MS2 coat protein (Fouts et al., 1997). This approach has enabled tracking of RNA localization in living cells over time with high sensitivity (Bertrand et al., 1998) but relies upon laborious genetic manipulation of the target RNA and is not suitable for recognition of arbitrary RNA sequences. Furthermore, insertion of exogenous aptamer sequence has the potential to interfere with endogenous RNA functions. Analogous to CRISPR/Cas9-based recognition of DNA, programmable RNA recognition based on nucleic acid specificity alone without the need for genetic manipulation or libraries of RNA-binding proteins would greatly expand researchers' ability to modify the mammalian transcriptome and enable transcriptome engineering.

Although the CRISPR/Cas9 system has evolved to recognize double-stranded DNA, recent *in vitro* work has demonstrated that programmable targeting of RNAs with Cas9 is possible by providing the PAM as part of an oligonucleotide (PAMmer) that hybridizes to the target RNA (O'Connell et al., 2014). By taking advantage of the Cas9 target search mechanism that relies on PAM sequences (Sternberg et al., 2014), a mismatched PAM sequence in the PAMmer/RNA hybrid allows exclusive targeting of RNA and not the encoding DNA. The high affinity and specificity of RNA recognition by Cas9 in cell-free extracts and the success of genome targeting with Cas9 indicate the potential of CRISPR/Cas9 to support programmable RNA targeting in living cells.

To assess the potential of Cas9 as a programmable RNA-binding protein in live cells, we used a modified sgRNA scaffold with improved expression and Cas9 association (Chen et al., 2013) with a stabilized PAMmer oligonucleotide that does not form a substrate for RNase H. We measured the degree of nuclear export of a nuclear localization signal-tagged nuclease-deficient Cas9-GFP fusion and demonstrate that the sgRNA alone is sufficient to promote nuclear export of Cas9 without influencing the abundance of the targeted mRNA or encoded protein. In order to evaluate whether RNA-targeting Cas9 (RCas9) signal patterns correspond with an established untagged RNA-labeling method, we compared distributions of RCas9 and fluorescence *in situ* hybridization (FISH) targeting *ACTB* mRNA. We observed high correlation among FISH and RCas9 signal that was dependent on the presence of a PAMmer, indicating the importance of the PAM for efficient RNA targeting. RNA trafficking and subcellular localization are critical to gene expression regulation and reaction to stimuli such as cellular stress. To address whether RCas9 allows tracking of RNA to oxidative stress-induced RNA/protein accumulations called stress granules, we measured *ACTB*, *TFRC*, and *CCNA2* mRNA association with stress granules in cells subjected to sodium arsenite. Finally, we demonstrated the ability of RCas9 to track trafficking of *ACTB* mRNA to stress granules over time in living cells. This work establishes the ability of RCas9 to bind RNA in live cells and sets the foundation for manipulation of the transcriptome in addition to the genome by CRISPR/Cas9.

## RESULTS

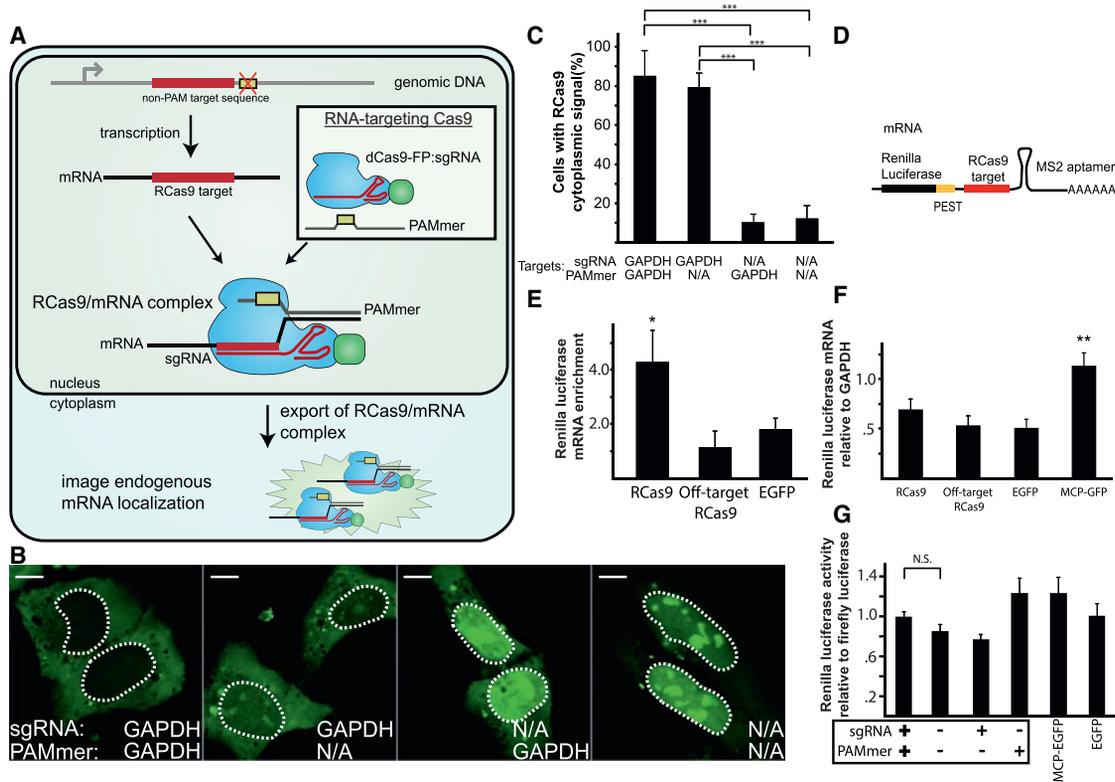
### RNA-Targeting Cas9 Export from Nucleus in Presence of sgRNA Targeting *GAPDH* mRNA

We initially assessed the ability of RCas9 to recognize specific mRNA substrates in human cells by evaluating the degree of

nuclear export of a nuclear-localized RCas9 system. Specifically, we tested whether mCherry-tagged Cas9 containing a nuclear localization signal (NLS) can be co-exported from the nucleus with an mRNA in the presence of a cognate sgRNA and PAMmer designed to recognize that mRNA (Figure 1A). Nuclease null Cas9 (dCas9) was fused to two SV40 NLS sequences at the C terminus with the coding sequence for mCherry and cloned into a mammalian expression vector (dCas9-2xNLS-mCherry, abbreviated as dCas9-mCherry). In a separate expression vector, a modified sgRNA scaffold with an extended stem-loop structure that improves association with Cas9 and mutations that eliminate a partial transcription termination sequence (Chen et al., 2013) was driven by the U6 small nuclear RNA (snRNA) polymerase III promoter. The PAMmer was synthesized as a mixed DNA and 2'-O-methyl (2'OMe) RNA oligonucleotide using standard phosphoramidite chemistry and purified using high-performance liquid chromatography (HPLC) (see Tables S1 and S2 for target mRNA, sgRNA, and PAMmer sequences). As a proof of concept, we designed a sgRNA-PAMmer pair to target the 3' UTR of *GAPDH* mRNA (Figure 1B). As a negative control, we designed a sgRNA-PAMmer pair targeting a sequence in the  $\lambda$  bacteriophage that is absent in human cells ("N/A" sgRNA and PAMmer and " $\lambda$ 2" sgRNA and PAMmer sequences in Tables S1 and S2). We observed that transiently transfected dCas9-mCherry co-transfected with the negative control sgRNA and PAMmer is almost exclusively nuclear with only 12% of cells containing greater mCherry signal in the cytoplasm compared to the nucleus (Figures 1B and 1C). When the negative control PAMmer was replaced with the *GAPDH*-targeting PAMmer, the results were identical. Importantly, upon co-transfection of *GAPDH*-targeting sgRNA plasmid, we observed that 84% of cells had greater mCherry signal in the cytoplasm. Interestingly, even with a non-targeting PAMmer, 79% of cells had predominantly cytoplasmic mCherry signal, suggesting that the sgRNA is the primary determinant of RNA substrate recognition. Indeed, the sgRNA targeting *GAPDH* resulted in a significant increase in the fraction of cells with cytoplasmic mCherry signal compared to a non-targeting sgRNA (Figure 1C). Overall, these results are consistent with previous *in vitro* RNA pull-down experiments showing that RNA binding by Cas9:sgRNA is independent of but strengthened by the PAMmer (O'Connell et al., 2014). Thus, we demonstrate that RCas9 RNA recognition is programmable with a sgRNA and a PAMmer designed to target a specific, abundant mRNA in live cells.

### Recognition of an mRNA with RNA-Targeting Cas9 Does Not Alter RNA Abundance or Amount of Translated Protein

To further characterize the interaction between RCas9 and a target mRNA, we directed RCas9 to the 3' UTR of *Renilla* luciferase carrying a commonly used RNA tag for RNA tracking from the MS2 bacteriophage (Fouts et al., 1997) and a sequence targeted by a previously validated sgRNA:PAMmer pair (" $\lambda$ 2"; see Table S2; O'Connell et al., 2014; Figure 1D). RNA immunoprecipitation with an antibody recognizing EGFP revealed a 4-fold greater association of luciferase mRNA to dCas9-EGFP in the presence of a cognate sgRNA and PAMmer compared to non-targeting sgRNA (sense to the  $\lambda$ 2 RNA sequence) with a



**Figure 1. Targeting mRNA in Living Cells with RCas9**

(A) Components required for RNA-targeting Cas9 (RCas9) recognition of mRNA include a nuclear localization signal-tagged nuclease-inactive Cas9 fused to a fluorescent protein such as GFP, a modified sgRNA with expression driven by the U6 polymerase III promoter, and a PAMmer composed of DNA and 2'-O-methyl RNA bases with a phosphodiester backbone. The sgRNA and PAMmer are antisense to adjacent regions of the target mRNA whose encoding DNA does not carry a PAM sequence. After formation of the RCas9/mRNA complex in the nucleus, the complex is exported to the cytoplasm.

(B) RCas9 nuclear co-export with *GAPDH* mRNA. The RCas9 system was delivered to U2OS cells with a sgRNA and PAMmer targeting the 3' UTR of *GAPDH* or sgRNA and PAMmer targeting a sequence from  $\lambda$  bacteriophage that should not be present in human cells ("N/A"). Cellular nuclei are outlined with a dashed white line. Scale bars represent 5 microns.

(C) Fraction of cells with cytoplasmic RCas9 signal. Mean values  $\pm$  SD (n = 50).

(D) A plasmid carrying the *Renilla* luciferase open reading frame with a  $\beta$ -globin 3' UTR containing a target site for RCas9 and MS2 aptamer. A PEST protein degradation signal was appended to luciferase to reveal any translational effects of RCas9 binding to the mRNA.

(E) RNA immunoprecipitation of EGFP after transient transfection of the RCas9 system in HEK293T cells targeting the luciferase mRNA compared to non-targeting sgRNA and PAMmer or EGFP alone. Mean values  $\pm$  SD (n = 3).

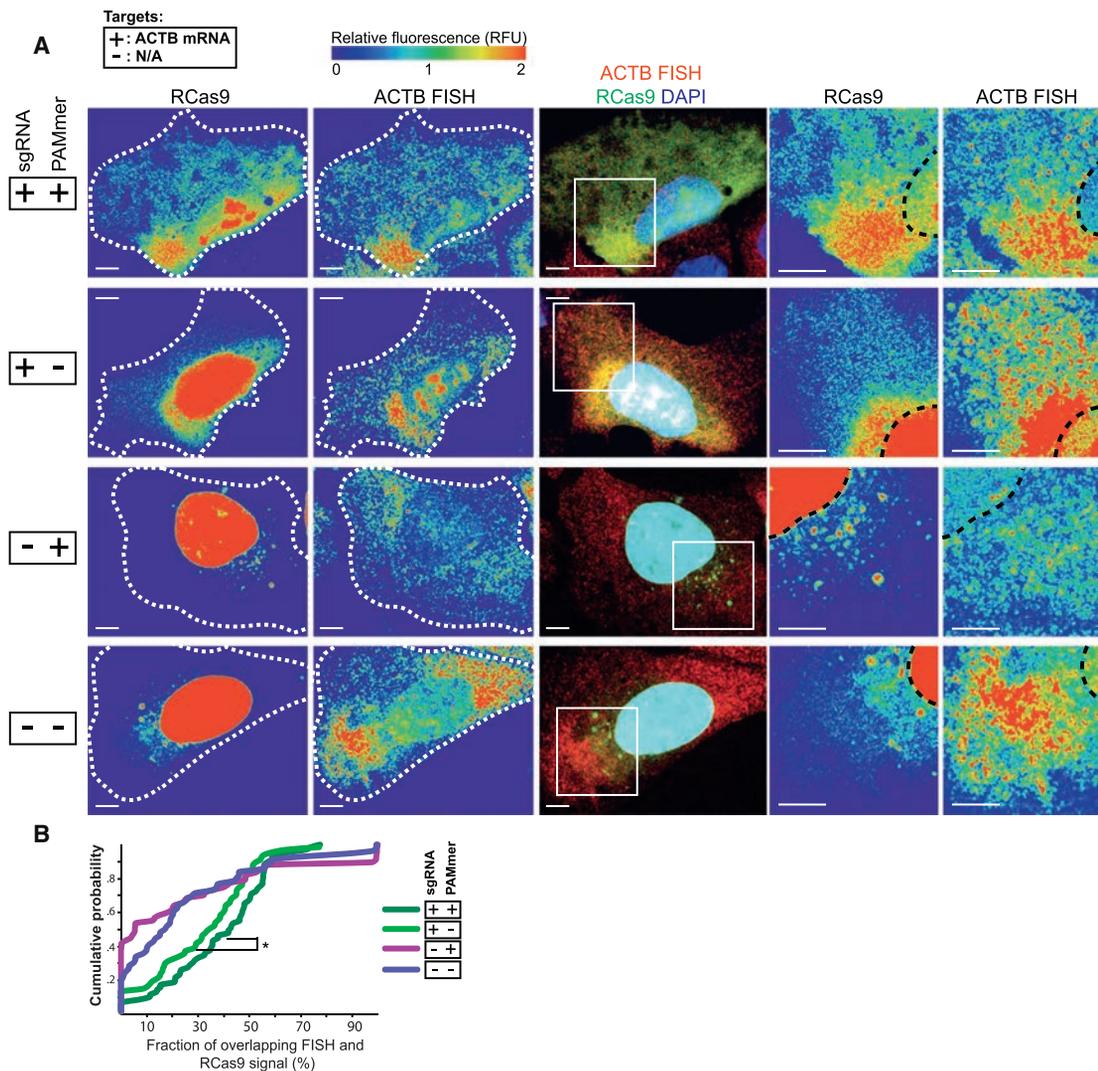
(F and G) *Renilla* luciferase mRNA (F) and protein (G) abundances were compared among the targeting and non-targeting conditions. Mean values  $\pm$  SD (n = 4). p values are calculated by Student's t test, and one, two, and three asterisks represent p values less than 0.05, 0.01, and 0.001, respectively. See also Figure S1.

scrambled PAMmer or to EGFP protein alone (Figure 1E). We next measured the effect of RCas9 targeting on luciferase mRNA abundance by qRT-PCR. We observed no significant difference in the abundance of MS2-tagged luciferase mRNA in the presence of the targeting or non-targeting RCas9 system or EGFP alone. In contrast, co-expression of EGFP fused to the MS2 coat protein (MCP) recognizing the MS2 aptamer had a significant stabilizing effect (Figure 1F). We also considered potential effects of RCas9 targeting on translation of luciferase (Figure 1G) and observed that the presence of the targeting sgRNA and PAMmer caused no significant changes in protein levels compared to non-targeting RCas9. To confirm that the RCas9/sgRNA/PAMmer complex does not perturb its endogenous RNA substrates, we evaluated the influence of RCas9 targeting on *ACTB* and *GAPDH* mRNAs. We transiently transfected the RCas9 system and isolated transfected cells using fluores-

cence-activated cell sorting (FACS) gated on cells positive for both Cas9 and sgRNA transfection. We observed no significant differences in *GAPDH* and *ACTB* mRNA or protein levels as measured by western blot and qRT-PCR among samples with RCas9-targeting *ACTB* mRNA, *GAPDH* mRNA, cells transfected with GFP instead of dCas9-GFP, and RCas9-targeting sequence from  $\lambda$  bacteriophage (Figure S1; Tables S1 and S2). Our results demonstrate that RCas9 recognition of RNA with a sgRNA and PAMmer does not perturb RNA and encoded protein levels.

### Correlation of RNA-Targeting Cas9 Signal Distributions with an Established Untagged RNA Localization Measurement

To assess whether RCas9 signal distributions correlate with an orthogonal method to measure RNA localization, we targeted the 3' UTR of *ACTB* ("+" sgRNA and "+" PAMmer) and compared



**Figure 2. *ACTB* mRNA Localization with RCas9 Compared to FISH**

(A) The RCas9 system was delivered to U2OS cells, and the cells were subjected to FISH for *ACTB* mRNA. RCas9 with sgRNA- and PAMmer-targeting *ACTB* mRNA was compared to non-targeting sgRNA and PAMmer antisense to a sequence from  $\lambda$  bacteriophage (“–” sgRNA and “–” PAMmer). White dotted lines delineate the cellular boundaries, and black dotted lines delineate cellular nuclei. Scale bars represent 5 microns. Insets (on right) are delineated by white boxes. (B) Pixel-by-pixel analysis of RCas9 and FISH colocalization using the Manders’ overlap coefficient is summarized with a cumulative distribution of the percent of cytoplasmic area with overlapping signal in 60–80 cells in each condition. The presence of the PAMmer produces a significantly greater colocalization among RCas9 and FISH in the presence of the sgRNA-targeting *ACTB* mRNA ( $p = 0.035$ ; two-tailed Mann-Whitney U test). See also Figure S2.

dCas9-GFP signal to RNA FISH for *ACTB* mRNA (Figure 2A) and non-targeting sgRNA and PAMmer (“–” sgRNA and “–” PAMmer with sequences corresponding to  $\lambda$  bacteriophage). By comparing the Manders’ overlap coefficients that describe pixel-by-pixel overlap among FISH and RCas9 (Manders et al., 1992; Figures 2B and S2), we determined that the sgRNA primarily accounts for colocalization among FISH and RCas9 with maximal overlap in the presence of both sgRNA- and PAMmer-targeting *ACTB* mRNA. A non-targeting PAMmer results in significantly less overlap (Figure 2B;  $p = 0.035$ ; Mann-Whitney U test) and produces a diffuse pattern of RCas9 signal in the cytoplasm that contrasts with the highly localized pattern revealed by FISH

(Figure 2A). This result is consistent with weaker binding of RCas9 with a non-targeting PAMmer observed in cell-free systems (O’Connell et al., 2014). A non-targeting sgRNA results in largely nuclear retention of RCas9 signal with low correlation between cytoplasmic RCas9 signal and FISH (Figures 2A and 2B). We conclude that the localization patterns of single EGFP-fused RCas9 in live cells correlates surprisingly well with localization obtained by FISH using tens of probes per mRNA in fixed cells.

#### Tracking RNA Trafficking to Stress Granules over Time

Many polyadenylated RNAs including *ACTB* mRNA are known to localize to stress granules (Unsworth et al., 2010) during

oxidative stress. We simultaneously tracked *ACTB* mRNA using RCas9 and RFP fused to the Ras GTPase-activating protein-binding protein 1 (G3BP1) protein, a well-described marker for stress granules (Tourrière et al., 2003). Application of sodium arsenite to induce cellular stress results in an accumulation of mRNAs, including *ACTB* mRNA, into G3BP1-positive granules. We observed accumulation of RCas9 signal to G3BP1-positive foci in the presence of the RCas9 system targeting *ACTB* mRNA. To verify that this accretion of RCas9 signal required specific RNA recognition, we utilized three different sgRNAs and PAMmers targeting sequences from  $\lambda$  bacteriophage (NTC1–3; Figures 3A and 3B). Indeed, in these negative controls, increased signal in G3BP1-positive foci was not observed. We also evaluated the ability of RCas9 to track less-abundant transcripts *CCNA2* and *TFRC* mRNA to stress granules. Notably, we observed statistically significantly higher fractions of stress granules that accumulate RCas9 signal (39% and 23%, respectively) compared to non-targeting controls (Figures 3A and 3C). Next, we tracked RCas9 signal in stressed, live cells over time (Figure 3B). We observed accumulation of RCas9 signal in G3BP1-positive foci in a manner dependent on the presence of sgRNA- and PAMmer-targeting *ACTB* mRNA. We also observed that the rate and degree of RCas9 signal accumulation in stress granules is dependent on dosage of the stressor sodium arsenite (Figure 3D). These results indicate the potential of RCas9 as a means to generate time-resolved RNA localization measurements.

## DISCUSSION

Effective RNA recognition by Cas9 in living cells while avoiding perturbation of the target transcript relies on careful design of the PAMmer and delivery of Cas9 and its cognate guide RNA to the appropriate cellular compartments. Binding of Cas9 to nucleic acids requires two critical features: a PAM DNA sequence and an adjacent spacer sequence antisense to the Cas9-associated sgRNA. By separating the PAM and sgRNA target among two molecules (the PAMmer oligonucleotide and the target mRNA), RCas9 allows recognition of RNA while avoiding the encoding DNA. To avoid unwanted degradation of the target RNA, the PAMmer is composed of a mixed 2'OMe RNA and DNA that does not form a substrate for RNase H. Further, the sgRNA features a modified scaffold that removes partial transcription termination sequences and a modified structure that promotes association with Cas9 (Chen et al., 2013). Other CRISPR/Cas systems have demonstrated RNA binding in bacteria (Hale et al., 2009; Sampson et al., 2013) or eukaryotes (Price et al., 2015), although these systems cannot discriminate RNA from DNA targets, feature RNA-targeting rules that remain unclear, or rely on large protein complexes that may be difficult to reconstitute in mammalian cells.

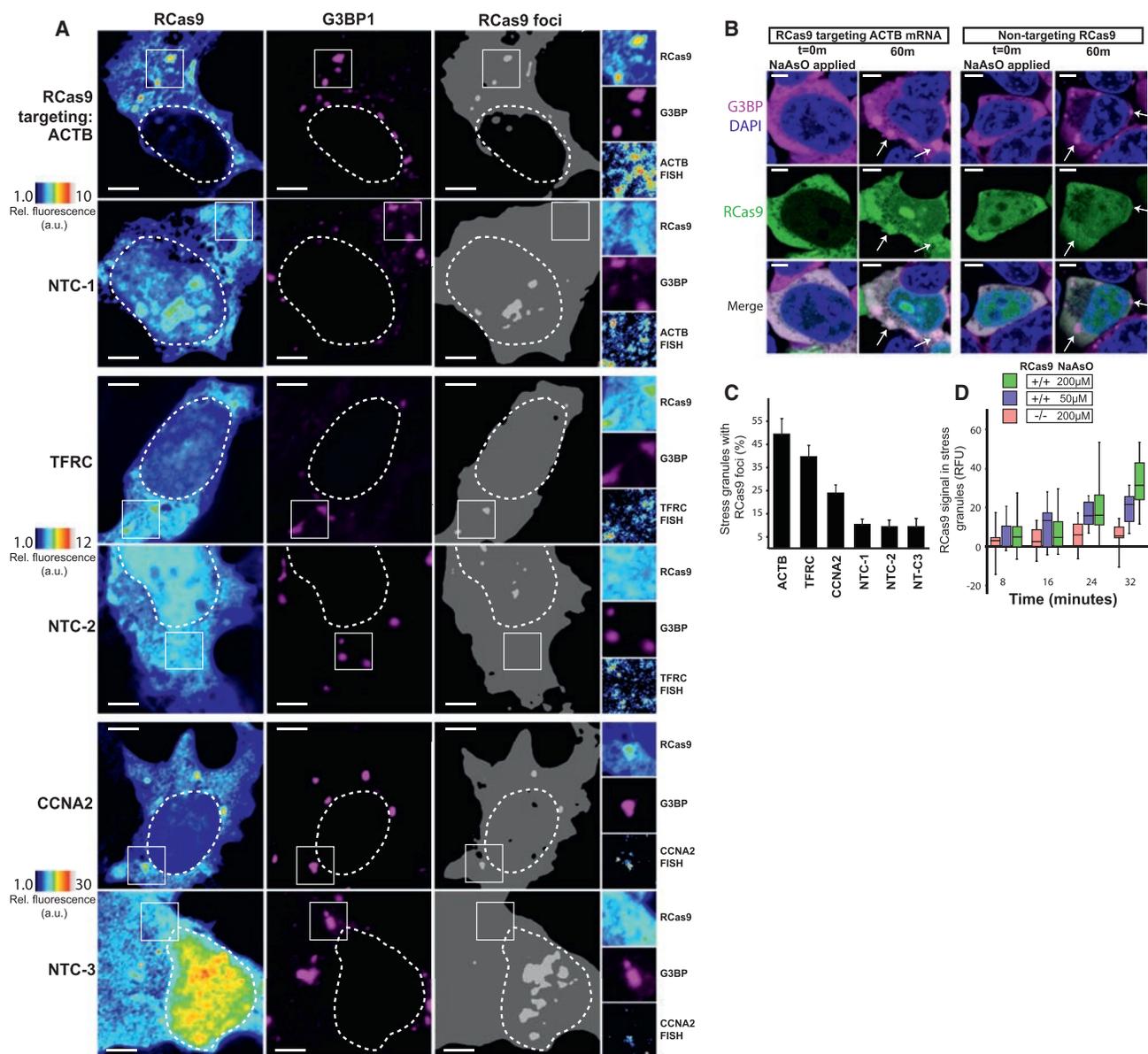
In this work, we demonstrate RCas9-based recognition of *GAPDH*, *ACTB*, *CCNA2*, and *TFRC* mRNAs in live cells. Because the U6-driven sgRNA is largely restricted to the nucleus, the NLS-tagged dCas9 allows association with its sgRNA and subsequent interaction with the target mRNA before nuclear co-export with the target mRNA. As an initial experiment, we

evaluated the potential of RNA recognition with Cas9 by targeting *GAPDH* mRNA and evaluating degree of nuclear export of dCas9-mCherry (Figure 1B). Robust cytoplasmic localization of dCas9-mCherry in the presence of a sgRNA-targeting *GAPDH* mRNA compared to nuclear retention in the presence of a non-targeting sgRNA indicated that Cas9 association with the mRNA was sufficiently stable to support co-export from the nucleus.

RCas9 as an RNA-imaging reagent requires that RNA recognition by RCas9 does not interfere with normal RNA metabolism. Here, we show that RCas9 binding within the 3' UTR of *Renilla* luciferase does not affect its mRNA abundance and translation (Figures 1F and 1G). The utility of RCas9 for imaging and other applications hinges on the recognition of endogenous transcripts, so we evaluated the influence of RCas9 targeting on *GAPDH* and *ACTB* mRNAs and observed no significant differences among the mRNA and protein abundances by western blot analysis and qRT-PCR (Figure S1). These results indicate that RCas9 targeting these 3' UTRs does not perturb the levels of mRNA or encoded protein.

We also evaluated the ability of RCas9 to reveal RNA localization by comparing RCas9 signal patterns to FISH. We utilized a FISH probe set composed of tens of singly labeled probes targeting *ACTB* mRNA and compared FISH signal distributions to a single dCas9-GFP/sgRNA/PAMmer that recognizes the *ACTB* mRNA. Our findings indicate that the sgRNA primarily determines the degree of overlap among the FISH and RCas9 signals whereas the PAMmer plays a significant but secondary role. Importantly, in contrast to other untagged RNA localization determination methods such as FISH and molecular beacons, RCas9 is compatible with tracking untagged RNA localization in living cells and can be delivered rapidly to cells using established transfection methods. We also note that the distribution of *ACTB* mRNA was visualized using a single EGFP tag per transcript, and higher-sensitivity RNA tracking or single endogenous RNA molecule visualization may be possible in the future with RCas9 targeting multiple sites in a transcript or with a multiply tagged dCas9 protein.

Stress granules are translationally silent mRNA and protein accumulations that form in response to cellular stress and are increasingly thought to be involved with neurodegeneration (Li et al., 2013b). There are limited means that can track the movement of endogenous RNA to these structures in live cells (Bertrand et al., 1998). In addition to *ACTB* mRNA, we demonstrate that RCas9 is capable of measuring association of *CCNA2* and *TFRC* mRNA trafficking to stress granules (Figure 3A). Upon stress induction with sodium arsenite, we observed that 50%, 39%, and 23% of stress granules featured overlapping RCas9 foci when targeting *ACTB*, *TFRC*, and *CCNA2* mRNAs, respectively (Figure 3C). This result correlates with the expression levels of these transcripts (Figure S3) as *ACTB* is expressed about 8 and 11 times more highly than *CCNA2* and *TFRC*, respectively. We also observed that RCas9 is capable of tracking RNA localization over time as *ACTB* mRNA is trafficked to stress granules over a period of 30 min (Figure 3B). We noted a dependence of RCas9 signal accumulation in stress granules on stressor concentration (Figure 3D). This approach for live-cell RNA tracking stands in contrast to molecular beacons and aptamer-based RNA-tracking



**Figure 3. Tracking of mRNA Trafficking to Stress Granules with RCas9**

(A) The RCas9 system targeting *ACTB*, *TFRC*, or *CCNA2* mRNAs or one of three non-targeting controls (NTCs) was delivered to HEK293T cells expressing G3BP1, a protein known to be trafficked to stress granules, fused to RFP. Cells were treated with sodium arsenite; fixed; subjected to FISH for *ACTB*, *TFRC*, or *CCNA2* mRNA; and imaged.

(B) RNA trafficking to stress granules was imaged in real time using cells harboring RCas9-targeting *ACTB* mRNA. At time zero, cells were imaged and sodium arsenite applied. Sixty minutes later, cells were imaged again, and a comparison of RCas9 and G3BP1-positive stress granules revealed close correlation of foci only in the presence of sgRNA- and PAMmer-targeting *ACTB* mRNA.

(C) The fraction of stress granules with RCas9 foci when targeting three mRNAs (*ACTB*, *TFRC*, and *CCNA2*) compared to three non-targeting controls. Error bars  $\pm$  SD calculated from 50 cells from each of three biological replicates. *p* values among the RCas9 system targeting *ACTB*, *TFRC*, and *CCNA2* mRNA are  $<0.001$  when compared to each of the NTC conditions. *p* values were calculated with Student's *t* test.

(D) In a similar experiment, RCas9-targeting *ACTB* mRNA signal accumulation in stress granules was tracked over time. Eight to eleven stress granules were tracked in each condition with time points every 8 min for 32 min.

Scale bars represent 5 microns. See also Figure S3.

methods, which suffer from delivery issues and/or require alteration of the target RNA sequence via incorporation of RNA tags.

Future applications of RCas9 could allow the measurement or alteration of RNA splicing via recruitment of split fluores-

cent proteins or splicing factors adjacent to alternatively spliced exons. Further, the nucleic-acid-programmable nature of RCas9 lends itself to multiplexed targeting (Cong et al., 2013) and the use of Cas9 proteins that bind orthogonal sgRNAs

(Esvelt et al., 2013) could support distinct activities on multiple target RNAs simultaneously. It is possible that the simple RNA targeting afforded by RCas9 could support the development of sensors that recognize specific healthy or disease-related gene expression patterns and reprogram cell behavior via alteration of gene expression or concatenation of enzymes on a target RNA (Delebecque et al., 2011; Sachdeva et al., 2014). Efforts toward Cas9 delivery *in vivo* are underway (Dow et al., 2015; Swiech et al., 2015; Zuris et al., 2015), and these efforts combined with existing oligonucleotide chemistries (Bennett and Swayze, 2010) could support *in vivo* delivery of the RCas9 system for targeted modulation of many features of RNA processing in living organisms.

RNA is subject to processing steps that include alternative splicing, nuclear export, subcellular transport, and base or backbone modifications that work in concert to regulate gene expression. The development of a programmable means of RNA recognition in order to measure and manipulate these processes has been sought after in biotechnology for decades. This work is, to our knowledge, the first demonstration of nucleic-acid-programmed RNA recognition in living cells with CRISPR/Cas9. By relying upon a sgRNA and PAMmer to determine target specificity, RCas9 supports versatile and unambiguous RNA recognition analogous to DNA recognition afforded by CRISPR/Cas9. The diverse applications supported by DNA-targeted CRISPR/Cas9 range from directed cleavage, imaging, transcription modulation, and targeted methylation, indicating the utility of both the native nucleolytic activity of Cas9 as well as the range of activities supported by Cas9-fused effectors. In addition to providing a flexible means to track this RNA in live cells, future developments of RCas9 could include effectors that modulate a variety of RNA-processing steps with applications in synthetic biology and disease modeling or treatment.

## EXPERIMENTAL PROCEDURES

### Plasmid Construction, PAMmer Synthesis, and Target Site Choice

The dCas9-2xNLS sequence was amplified from pHR-SFFV-dCas9-BFP-KRAB (a gift from Stanley Qi and Jonathan Weissman; Addgene plasmid no. 46911), tagged with two SV40 NLSs on the C terminus, and fused to EGFP or mCherry in pCDNA 3.1 (Life Technologies) using Gibson assembly. To construct the sgRNA scaffold construct, the human U6 polymerase III promoter with the modified sgRNA scaffold (Chen et al., 2013) was purchased as a gBlock from IDT with two *BbsI* restriction sites at the 5' end of the sgRNA scaffold (see sequence in Table S2) and cloned into the multiple cloning site of pBlueScript II SK (+) (Agilent) using Gibson assembly. Phosphorylated oligonucleotides encoding the sgRNA sequences (with overhangs 5'CACC on the RNA antisense strand and 5'AAAC on the sense strand) were ligated into *BbsI*-digested sgRNA scaffold construct to produce sgRNAs targeting the 3' UTR of *GAPDH*, *ACTB*, and *Renilla* luciferase mRNAs (see Tables S1 and S2). The luciferase-PEST construct for pull-down and abundance experiments was modified from plasmid pRLuc (gift from Jens Lykke-Andersen, UCSD). pCMV-*Renilla* luciferase is a version of the same construct lacking MS2 and RCas9 target sites.

RCas9 target sites were chosen with a combination of the IDT antisense oligonucleotide design tool and the microarray probe design tools Picky (Chou et al., 2004) and OligoWiz (Wernersson and Nielsen, 2005). We designed PAMmers against high-confidence sites with eight bases on the 5' end beyond the PAM sequence. PAMmers were composed of mixed 2'OMe RNA and DNA bases and purified by HPLC (Integrated DNA Technologies).

### Cell Lines

U2OS and HEK293T cells were grown in DMEM supplemented with 10% fetal bovine serum, Glutamax, penicillin/streptomycin, and non-essential amino acids (Life Technologies). Cells were passaged every 3 or 4 days with TrypLE EXPRESS (Life Technologies) using standard methods and maintained in a humidified incubator at 37°C with 5% CO<sub>2</sub>.

### GAPDH and ACTB mRNA Targeting with RCas9

U2OS cells cultured as described above were passaged at ~80% confluency. Glass-bottom 96-well plates or chamber slides were coated with 20 µg/ml fibronectin in PBS for 2 hr at 37°C and then the fibronectin solution was aspirated and 20,000 cells were plated in each well. Sixteen hours later, cells were transfected with the sgRNA and dCas9-2xNLS-EGFP plasmids using Lipofectamine 3000 (Life Technologies) according to the manufacturer's instructions. pCMV-*Renilla* luciferase was co-transfected in these experiments so that total transfected protein load was the similar among various dosages of sgRNA and dCas9. The mass ratio of sgRNA and dCas9-EGFP plasmids was 5:1 in experiments described in this work where the amount of dCas9-EGFP was fixed at 10% of total transfected material. Immediately after plasmid transfection, PAMmers were transfected using Lipofectamine RNAiMax (Life Technologies) according to manufacturer's instructions. Twenty-four hours after transfection, cells were washed with PBS and fixed with 3.7% paraformaldehyde in PBS, permeabilized with 70% ethanol at 4°C for 1 hr, and mounted using Prolong Gold Antifade mounting medium with DAPI (Life Technologies). Confocal microscopy was conducted using a Zeiss LSM 810 confocal microscope.

Nuclear export of RCas9 in the presence of sgRNA and PAMmer targeting the 3' UTR of *GAPDH* mRNA was analyzed by measuring the average signal in the nuclei and cytoplasm of individual cells. Cells with average cytoplasmic signal greater than the average nuclear signal were considered positive and counted toward the fraction of cells with cytoplasmic RCas9 signal.

### RNA Immunoprecipitation

HEK293T cells cultured as described above were passaged at 80% confluency, and 600,000 cells were seeded in each well of 6-well tissue culture plates coated with poly-L-lysine. Sixteen hours later, cells were co-transfected with the RCas9 system as described above or plasmids encoding MS2-EGFP or EGFP along with a plasmid encoding the model *Renilla* luciferase mRNA driven by a cytomegalovirus (CMV) promoter (Figure 1D). Twenty-four hours later, the growth media was aspirated and the cells were washed with PBS. 1% paraformaldehyde in PBS was applied to the cells, incubated for 10 min at room temperature, and then the solution was aspirated and the cells washed twice with cold PBS. Next, the cells were scraped from the wells in cold PBS and the cell suspension was centrifuged at 800 × g for 4 min to pellet the cells. The cells were washed once more and then resuspended in radioimmunoprecipitation assay (RIPA) buffer with protease inhibitors (Roche) and sonicated for 5 min in a Bioruptor sonicator (50% duty cycle; 1 min period). Insoluble material was pelleted after a high-speed centrifugation, and the supernatant was applied to protein G DynaBeads (Life Technologies) coated with mouse anti-GFP antibody (Roche). After overnight incubation at 4°C, the bead supernatant was retained and beads washed three times with RIPA buffer containing 0.02% Tween-20 and once with DNase buffer (350 mM Tris-HCl [pH 6.5]; 50 mM MgCl<sub>2</sub>; 5 mM DTT). The beads were resuspended in DNase buffer and TURBO DNase (Life Technologies) was added to 0.08 units/µl. The beads were incubated at 37°C for 30 min, and then proteinase K (NEB) was added to 0.1 U/µl and incubated with shaking at 37°C for 30 min. Next, urea was added to 2.5 M and the beads were incubated with shaking at 37°C for 30 min. The bead supernatant was collected and subjected to two sequential phenol:chloroform:isoamyl alcohol (25:24:1) extractions followed by three chloroform extractions. The RNA was precipitated and reverse transcribed using SuperScript III (Life Technologies) using random hexamer primers, and relative abundance of *Renilla* luciferase RNA on the beads was compared to the supernatant using RT-PCR (see Table S3 for primer sequences).

### Measurements of Influence of RCas9 on RNA Stability and Translation

HEK293T cells were cultured as described above, passaged and plated in 96- or 12-well tissue culture plates, and co-transfected 24 hr later with the

RCas9 system as described above and the *Renilla* luciferase construct carrying MS2 and RCas9 binding sites in the 3' UTR. In the protein abundance measurements, a small amount of CMV-driven firefly luciferase vector (5% of total transfected plasmid) was co-transfected as a transfection control. For RNA stability measurements, RNA was isolated 24 hr after transfection, DNase treated, and reverse transcribed with Superscript III (Life Technologies) using dT(20) primers according to the manufacturer's instructions. The amount of *Renilla* luciferase cDNA relative to *GAPDH* was then measured using RT-PCR (see Table S3 for primer sequences). For the translation studies, *Renilla* and firefly luciferase protein were measured with the Dual Luciferase Kit (Promega) according to the manufacturer's instructions.

For measuring the influence of RCas9 on *GAPDH* and *ACTB* mRNA and protein abundance, the RCas9 system targeting the 3' UTR of each transcript, GFP alone, or RCas9 with the  $\lambda$ 2-targeting sgRNA and PAMmer was transfected. Modified sgRNA vectors carrying pGK-driven blue fluorescent protein (BFP) were used in this experiment, and transfected cells were isolated using FACS gated on cells positive for GFP and BFP expression. Transfected cells were lysed and subjected to western blotting and RT-PCR for *ACTB*, *GAPDH*, and *TUBA1A* transcripts.

### FISH

Stellaris FISH Probes recognizing human *ACTB*, *CCNA2*, and *TFRC* mRNAs and labeled with Quasar 670 (VSMF-2003-5; Biosearch Technologies) were hybridized to cells 24 hr after transfection with the RCas9 system. Hybridization was conducted according to the manufacturer's instructions. Confocal microscopy was conducted using an Olympus FV1000 confocal microscope or Zeiss LSM 810 confocal microscope.

### Overlap Analysis for FISH and RCas9

Colocalization analysis among FISH and RCas9-targeting *ACTB* mRNA was conducted using the Coloc 2 plugin from the image analysis software FIJI (Schindelin et al., 2012). The cytoplasm of individual cells with similar dCas9-EGFP transfection levels was selected, and the Coloc 2 analysis was conducted using default parameters. The Manders' overlap coefficient describing degree of overlap of the FISH signal with RCas9 for more than 60 cells in each condition was compiled, and p values were calculated with the two-tailed Mann-Whitney U test.

### Tracking ACTB mRNA Trafficking to Stress Granules

A HEK293T cell line was genetically modified with a fusion of RFP to the C terminus of Ras G3BP1 using CRISPR/Cas9. Briefly, a donor plasmid was constructed consisting of the RFP ORF, a puromycin selection cassette, and flanking 1.5-kb homology arms directed at the G3BP1 locus. An sgRNA sequence targeting the C terminus of G3BP1 was cloned into pSpCas9(BB)-2A-GFP (pX458) (gift from Feng Zhang; Addgene plasmid no. 48138) and co-transfected with the donor plasmid using Fugene HD (Roche) following the manufacturer's instructions. Forty-eight hours after transfection, cells were selected with 1  $\mu$ g/ml puromycin in growth medium for 14 days and RFP-positive clones were selected and screened by PCR.

A clone with at least one modified allele was plated on glass chamber slides coated with fibronectin and transfected with the RCas9 system targeting the 3' UTR of *ACTB*, *CCNA2*, or *TFRC* mRNA or one of three non-targeting sgRNA and PAMmer combinations ( $\lambda$ 2- $\lambda$ 4) as described above (see Table S2 for sgRNA and PAMmer sequences). Twenty-four hours after transfection, cells were subjected to 200  $\mu$ M sodium arsenite for 30 min, fixed, and subjected to FISH and imaged with a Zeiss LSM 810 confocal microscope. In a separate experiment, cells were stressed with varying concentrations of sodium arsenite and imaged with a Zeiss LSM 810 confocal microscope with a stage incubator over 32 min. Cells were maintained at 37°C in a humidified atmosphere and 5% CO<sub>2</sub> and imaged every 8 min.

### Analysis of ACTB mRNA Trafficking to Stress Granules

Stress granules with average RCas9 signal at least 2-fold greater than surrounding cytoplasm were determined to be positive for RCas9. The average RCas9 signal in stress granules was calculated using the "measure" function in Fiji and compared to the average cytoplasmic RCas9 signal in the surrounding cytoplasm. RCas9-positive foci were visualized by dividing the value

of each pixel by the average cytoplasmic signal, which reduced cytoplasmic signal to a value of "1" and RCas9-positive foci to a value of "2" or more. Next, a Gaussian blur was applied to improve foci definition. These images are reported in Figure 3A. The accumulation of RCas9 signal over time in stress granules was calculated by recording average signal intensity in the RCas9 channel in areas with overlapping G3BP1-RFP foci at each time point. Average signal intensity in the RCas9 channel surrounding G3BP1-RFP foci was recorded as background and subtracted from the previous value to produce the background-adjusted RCas9 signal in stress granules.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.02.054>.

A video abstract is available at <http://dx.doi.org/10.1016/j.cell.2016.02.054#mmc3>.

### AUTHOR CONTRIBUTIONS

Conceptualization, D.A.N., M.R.O., J.A.D., and G.W.Y.; Methodology, D.A.N. and M.R.O.; Investigation, D.A.N., M.Y.F., J.L.X., and S.J.M.; Writing – Original Draft, D.A.N. and G.W.Y.; Writing – Review & Editing, D.A.N., G.W.Y., and M.R.O.; Funding Acquisition, G.W.Y.; Supervision, G.W.Y.

### ACKNOWLEDGMENTS

We acknowledge members of the G.W.Y. lab, particularly Ashleigh Schaffer, Stefan Aigner, and Ron Batra for critical comments. This work was supported by grants from the NIH (HG004659 and NS075449) and from the California Institute of Regenerative Medicine (RB3-05009 and RB4-06045) to G.W.Y. D.A.N. is supported by the National Science Foundation Graduate Research Fellowship (DGE-1144086) and Powell-Focht graduate fellowship. M.R.O. is supported by a CJ Martin fellowship from the National Health and Medical Research Council (Australia). J.A.D. is a Howard Hughes Medical Institute Investigator. G.W.Y. is an Alfred P. Sloan Research Fellow. J.A.D. is a co-founder of Editas Medicine, Intellia Therapeutics, and Caribou Biosciences and a scientific advisor to Caribou, Intellia, eFFECTOR Therapeutics, and Driver.

Received: November 15, 2015

Revised: February 5, 2016

Accepted: February 24, 2016

Published: March 17, 2016

### REFERENCES

- Bennett, C.F., and Swayze, E.E. (2010). RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu. Rev. Pharmacol. Toxicol.* *50*, 259–293.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* *2*, 437–445.
- Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., and Huang, B. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* *155*, 1479–1491.
- Chou, H.H., Hsia, A.P., Mooney, D.L., and Schnable, P.S. (2004). Picky: oligo microarray design for large genomes. *Bioinformatics* *20*, 2893–2902.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.
- Delebecque, C.J., Lindner, A.B., Silver, P.A., and Aldaye, F.A. (2011). Organization of intracellular reactions with rationally designed RNA assemblies. *Science* *333*, 470–474.

- Dow, L.E., Fisher, J., O'Rourke, K.P., Muley, A., Kastenhuber, E.R., Livshits, G., Tschaharganeh, D.F., Socci, N.D., and Lowe, S.W. (2015). Inducible in vivo genome editing with CRISPR-Cas9. *Nat. Biotechnol.* **33**, 390–394.
- Esvelt, K.M., Mali, P., Braff, J.L., Moosburner, M., Yaung, S.J., and Church, G.M. (2013). Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* **10**, 1116–1121.
- Filipovska, A., Razif, M.F., Nygård, K.K., and Rackham, O. (2011). A universal code for RNA recognition by PUF proteins. *Nat. Chem. Biol.* **7**, 425–427.
- Fouts, D.E., True, H.L., and Celander, D.W. (1997). Functional recognition of fragmented operator sites by R17/MS2 coat protein, a translational repressor. *Nucleic Acids Res.* **25**, 4464–4473.
- Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., Terns, R.M., and Terns, M.P. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945–956.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229.
- Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li, Y., Gao, N., Wang, L., Lu, X., et al. (2013a). Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 681–683.
- Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013b). Stress granules as crucibles of ALS pathogenesis. *J. Cell Biol.* **207**, 361–372.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826.
- Manders, E.M., Stap, J., Brakenhoff, G.J., van Driel, R., and Aten, J.A. (1992). Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. *J. Cell Sci.* **103**, 857–862.
- Nakayama, T., Fish, M.B., Fisher, M., Oomen-Hajagos, J., Thomsen, G.H., and Grainger, R.M. (2013). Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in *Xenopus tropicalis*. *Genesis* **51**, 835–843.
- O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516**, 263–266.
- Ozawa, T., Natori, Y., Sato, M., and Umezawa, Y. (2007). Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat. Methods* **4**, 413–419.
- Price, A.A., Sampson, T.R., Ratner, H.K., Grakoui, A., and Weiss, D.S. (2015). Cas9-mediated targeting of viral RNA in eukaryotic cells. *Proc. Natl. Acad. Sci. USA* **112**, 6164–6169.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183.
- Sachdeva, G., Garg, A., Godding, D., Way, J.C., and Silver, P.A. (2014). In vivo co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner. *Nucleic Acids Res.* **42**, 9493–9503.
- Sampson, T.R., Saroj, S.D., Llewellyn, A.C., Tzeng, Y.L., and Weiss, D.S. (2013). A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* **497**, 254–257.
- Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.
- Sokol, D.L., Zhang, X., Lu, P., and Gewirtz, A.M. (1998). Real time detection of DNA:RNA hybridization in living cells. *Proc. Natl. Acad. Sci. USA* **95**, 11538–11543.
- Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67.
- Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M., and Zhang, F. (2015). In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* **33**, 102–106.
- Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E., and Tazi, J. (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J. Cell Biol.* **160**, 823–831.
- Unsworth, H., Raguz, S., Edwards, H.J., Higgins, C.F., and Yagüe, E. (2010). mRNA escape from stress granule sequestration is dictated by localization to the endoplasmic reticulum. *FASEB J.* **24**, 3370–3380.
- Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646.
- Wang, Y., Cheong, C.G., Hall, T.M., and Wang, Z. (2009). Engineering splicing factors with designed specificities. *Nat. Methods* **6**, 825–830.
- Wernersson, R., and Nielsen, H.B. (2005). OligoWiz 2.0—integrating sequence feature annotation into the design of microarray probes. *Nucleic Acids Res.* **33**, W611–W615.
- Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338.
- Yang, D., Xu, J., Zhu, T., Fan, J., Lai, L., Zhang, J., and Chen, Y.E. (2014). Effective gene targeting in rabbits using RNA-guided Cas9 nucleases. *J. Mol. Cell Biol.* **6**, 97–99.
- Zuris, J.A., Thompson, D.B., Shu, Y., Guilinger, J.P., Bessen, J.L., Hu, J.H., Maeder, M.L., Joung, J.K., Chen, Z.Y., and Liu, D.R. (2015). Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat. Biotechnol.* **33**, 73–80.

# High-Throughput, High-Resolution Mapping of Protein Localization in Mammalian Brain by In Vivo Genome Editing

Takayasu Mikuni,<sup>1,4</sup> Jun Nishiyama,<sup>1,4,\*</sup> Ye Sun,<sup>1,2</sup> Naomi Kamasawa,<sup>3</sup> and Ryohei Yasuda<sup>1,\*</sup>

<sup>1</sup>Neuronal Signal Transduction Group, Max Planck Florida Institute for Neuroscience, Jupiter, FL 33458, USA

<sup>2</sup>Integrative Program in Biology and Neuroscience, Florida Atlantic University, Jupiter, FL 33458, USA

<sup>3</sup>Electron Microscopy Core Facility, Max Planck Florida Institute for Neuroscience, Jupiter, FL 33458, USA

<sup>4</sup>Co-first author

\*Correspondence: jun.nishiyama@mpfi.org (J.N.), ryohei.yasuda@mpfi.org (R.Y.)

<http://dx.doi.org/10.1016/j.cell.2016.04.044>

## SUMMARY

A scalable and high-throughput method to identify precise subcellular localization of endogenous proteins is essential for integrative understanding of a cell at the molecular level. Here, we developed a simple and generalizable technique to image endogenous proteins with high specificity, resolution, and contrast in single cells in mammalian brain tissue. The technique, single-cell labeling of endogenous proteins by clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9-mediated homology-directed repair (SLENDR), uses in vivo genome editing to insert a sequence encoding an epitope tag or a fluorescent protein to a gene of interest by CRISPR-Cas9-mediated homology-directed repair (HDR). Single-cell, HDR-mediated genome editing was achieved by delivering the editing machinery to dividing neuronal progenitors through in utero electroporation. We demonstrate that SLENDR allows rapid determination of the localization and dynamics of many endogenous proteins in various cell types, regions, and ages in the brain. Thus, SLENDR provides a high-throughput platform to map the subcellular localization of endogenous proteins with the resolution of micro- to nanometers in the brain.

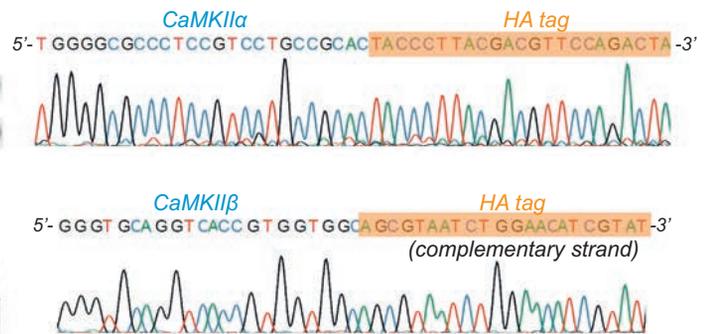
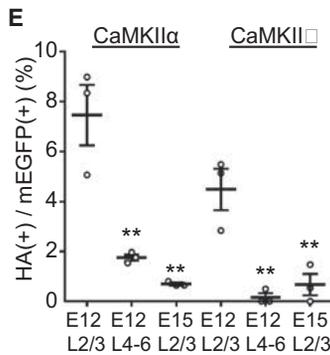
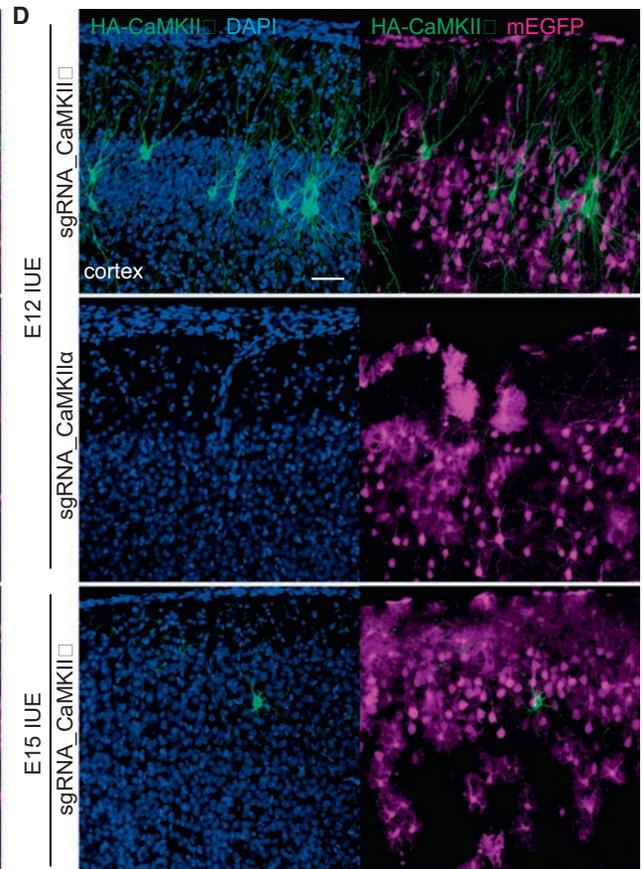
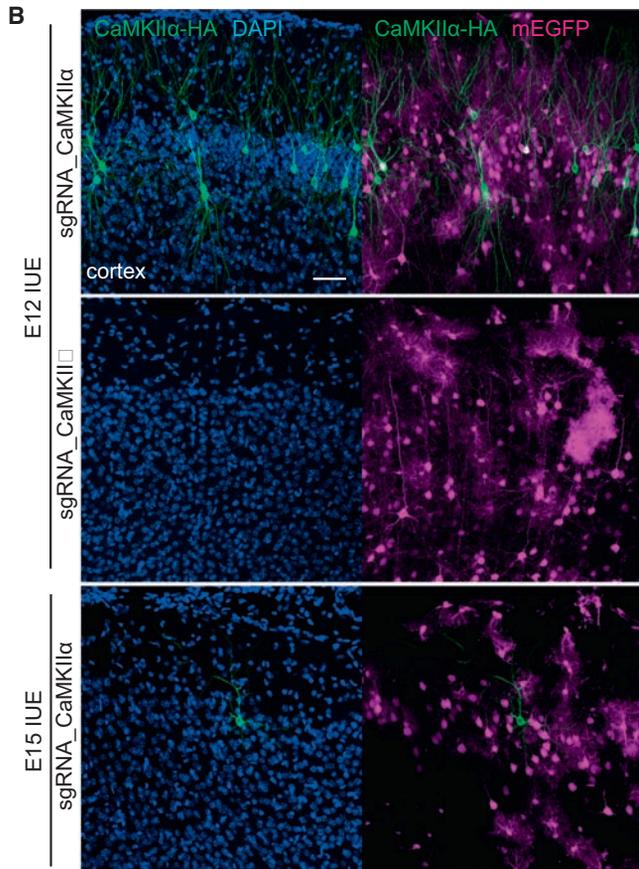
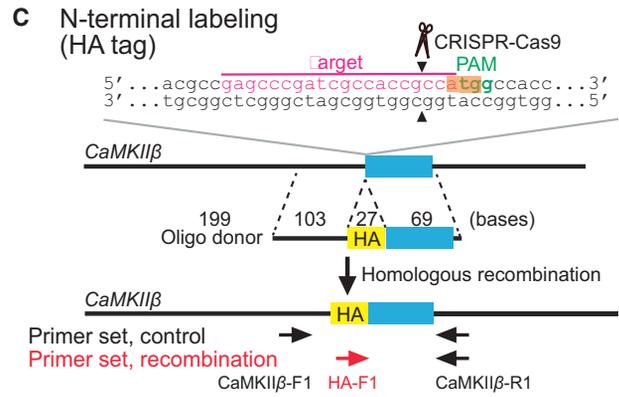
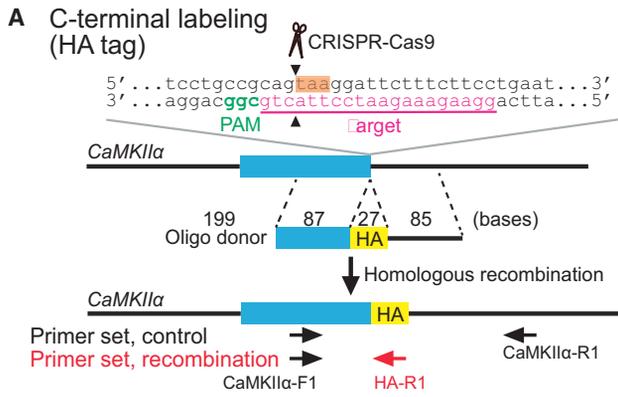
## INTRODUCTION

Precise mapping of a large number of proteins with subcellular resolution is essential to understand cellular processes. Thus, it is critical to develop a rapid and scalable method to determine the localization of proteins with high specificity, resolution, and contrast. Conventionally, either immunostaining of endogenous proteins or overexpression of proteins fused with epitope tags or fluorescent proteins have been used to determine protein localization. These methods, however, have significant problems: immunostaining often suffers from the lack of specific antibodies against a protein of interest and the cross-reaction of

antibodies with non-targeted proteins; overexpression often causes mistargeting of the expressed protein and potential changes in cell function. To address some of these issues, knockin mice in which a specific protein is tagged with an epitope tag or fluorescent protein can be used (Yang et al., 2009). However, in dense tissue, such as mammalian brain, it is difficult to obtain images with high contrast in subcellular processes when all cells are labeled. To overcome these problems, several methods have been recently developed for single-cell labeling of endogenous proteins by using recombinant antibody-like proteins or a conditional tag knockin strategy (Fortin et al., 2014; Gross et al., 2013). However, none of these techniques provides rapid, scalable, and high-throughput readouts for the localization of endogenous proteins.

Direct, single-cell manipulation of the genome in vivo to insert a tag sequence to a gene of interest would overcome these limitations, providing rapid, specific, and sparse labeling of the gene product. Genome editing based on the clustered regularly interspaced short palindromic repeats (CRISPR)-associated endonuclease Cas9 enables rapid and efficient modification of the genome (Cong et al., 2013; Doudna and Charpentier, 2014; Hsu et al., 2014; Jinek et al., 2012; Sander and Joung, 2014; Wang et al., 2013; Yang et al., 2013). CRISPR-Cas9 induces targeted DNA double-strand breaks in the genome, which are then repaired through either non-homologous end-joining (NHEJ) or homology-directed repair (HDR) pathways (Cox et al., 2015; Doudna and Charpentier, 2014; Hsu et al., 2014; Sander and Joung, 2014; Yang et al., 2013). Although introducing frame-shift knockout mutations through NHEJ at the single-cell level has been established (Straub et al., 2014; Swiech et al., 2015), targeted insertion of a sequence through HDR has not been possible in the mammalian brain in vivo (Heidenreich and Zhang, 2016; Platt et al., 2014; Xue et al., 2014; Yin et al., 2014). This is due to the lack of homologous recombination activity in postmitotic cells and the inefficient delivery of HDR machinery to target cells (Chu et al., 2015; Cox et al., 2015; Heidenreich and Zhang, 2016; Hsu et al., 2014; Maruyama et al., 2015; Saleh-Gohari and Helleday, 2004).

Here, we developed single-cell labeling of endogenous proteins by CRISPR-Cas9-mediated homology-directed repair (SLENDR), a technique that allows HDR-mediated genome editing in the mammalian brain in vivo. Since HDR is known to



(legend on next page)

predominantly occur in the S and G2 phases of the cell cycle (Chu et al., 2015; Heidenreich and Zhang, 2016; Hsu et al., 2014; Maruyama et al., 2015; Saleh-Gohari and Helleday, 2004), we targeted mitotic neuronal progenitors, which presumably have homologous recombination activity. We introduced CRISPR-Cas9-based HDR machinery into progenitor cells in the embryonic mouse brain several days before their final neurogenic divisions using in utero electroporation (IUE) (Nishiyama et al., 2012; Tabata and Nakajima, 2001). We demonstrate that a tag sequence for a short epitope or a longer fluorescent protein can be rapidly and precisely inserted into an endogenous gene of interest in vivo. This method is scalable to many species of proteins in diverse cell types, and permits high-resolution imaging with light and electron microscopy both in fixed and live tissue. Thus, SLENDR allows researchers to rapidly and precisely determine the localization and dynamics of endogenous proteins with the resolution of micro- to nanometers in various cell types, regions, and ages of the brain, providing a powerful tool suitable for large-scale analysis on a broad spectrum of proteins.

## RESULTS

### In Vivo Single-Cell Labeling of Endogenous Proteins by Homology-Directed Repair

In order to label a specific protein with a tag in non-dividing brain cells by HDR-mediated gene editing, it is necessary to introduce HDR machinery into progenitor cells before their final cell division. To test whether this strategy, termed SLENDR, can provide the efficiency sufficient for imaging subcellular localization of a protein of interest, we first aimed to insert the human influenza hemagglutinin (HA) tag into endogenous CaMKII $\alpha$  and CaMKII $\beta$ , two major subunits of the dodecameric Ca<sup>2+</sup>/CaM-dependent kinase II (CaMKII) necessary for neuronal plasticity (Kim et al., 2015; Lee et al., 2009; Lisman et al., 2012).

We designed specific single-guide RNAs (sgRNAs) targeting the vicinity of the stop and start codon of CaMKII $\alpha$  and CaMKII $\beta$ , respectively. We also designed corresponding single-stranded oligodeoxynucleotides (ssODNs) (~200 bases) to integrate the HA tag sequence into the genome just upstream of the stop codon of CaMKII $\alpha$  and downstream of the start codon of CaMKII $\beta$  (Figures 1A and 1C). To minimize the possibility of gene knockdown in cells where CRISPR-Cas9-mediated DNA

double-strand breaks were repaired through the NHEJ pathway, we selected the target sequences throughout this study so that the CRISPR-Cas9 cleavage sites were located either in the non-coding region upstream of the start codon or within 10 bp from the stop codon. We also selected the target sequences and designed ssODNs so that Cas9 could not recognize the loci after HDR was completed (see Supplemental Information). We introduced these constructs (*Streptococcus pyogenes* Cas9 or SpCas9 and sgRNA expressing vectors and ssODNs) together with hyperactive *piggyBac* transposase and *piggyBac* transposon vectors expressing monomeric EGFP (mEGFP) as a marker of transfection (Chen and LoTurco, 2012; Yusa et al., 2011) to neuro-progenitor cells using IUE to target pyramidal neurons in the cerebral cortex. The transposon system induces genomic integration of transgenes, preventing the dilution of mEGFP during cell divisions. Following IUE at embryonic day 12 (E12), we performed immunostaining of brain slices at post-natal days 14–48 (P14–P48) using anti-HA antibody together with anti-NeuN antibody to label neurons.

In the stained slices, HA signals were observed in a sparse subset of neurons, suggesting that HDR was successfully induced in these cells. Immunofluorescence signal was localized mostly in cytosol and excluded from the nucleus, consistent with previously reported distribution of CaMKII (Lee et al., 2009). Most of HA-positive neurons were found in layer 2/3 (Figures 1B, 1D, and 1E). Among neurons in transfected area in layer 2/3, about a half of the population was mEGFP-positive (mEGFP/NeuN: CaMKII $\alpha$ , 48.2%  $\pm$  6.7%; CaMKII $\beta$ , 40.0%  $\pm$  2.9%). Among these mEGFP-positive cells, a small population of the cells were found to be HA-positive (HA/mEGFP: CaMKII $\alpha$ , 7.5%  $\pm$  1.2%; CaMKII $\beta$ , 4.5%  $\pm$  0.8%), providing a few percent of overall knockin efficiency (HA/NeuN: CaMKII $\alpha$ , 3.4%  $\pm$  0.2%; CaMKII $\beta$ , 1.8%  $\pm$  0.5%). We also found a smaller population of HA-positive neurons in layers 4–6 (HA/mEGFP: CaMKII $\alpha$ , 1.8%  $\pm$  0.1%; CaMKII $\beta$ , 0.2%  $\pm$  0.1%) (Figures 1B, 1D, and 1E). When IUE was performed at E15, a smaller population of layer 2/3 neurons were labeled with the HA tag (HA/mEGFP: CaMKII $\alpha$ , 0.7%  $\pm$  0.1%; CaMKII $\beta$ , 0.7%  $\pm$  0.4%, HA/NeuN: CaMKII $\alpha$ , 0.1%  $\pm$  0.0%; CaMKII $\beta$ , 0.2%  $\pm$  0.1%) (Figures 1B, 1D, and 1E; Tables S1 and S2). Given that the superficial cortical layers (layer 2/3) are populated with later-born neurons (Chen and LoTurco, 2012), these findings suggest that SLENDR is more efficient

### Figure 1. In Vivo Single-Cell Labeling of Endogenous Proteins by Homology-Directed Repair

(A and C) Graphical representation of the mouse genomic loci of CaMKII $\alpha$  (A) and CaMKII $\beta$  (C) showing the target sites for Cas9, sgRNA, and ssODNs. The sgRNA targeting sequences are labeled in magenta. The protospacer-adjacent motif (PAM) sequences are labeled in green. The stop and start codons of CaMKII $\alpha$  (A) and CaMKII $\beta$  (C), respectively, are marked in orange. The Cas9 cleavage sites are indicated by the black arrowheads. PCR primer sets (control and recombination) for PCR genotyping (E and F) are indicated by the arrows.

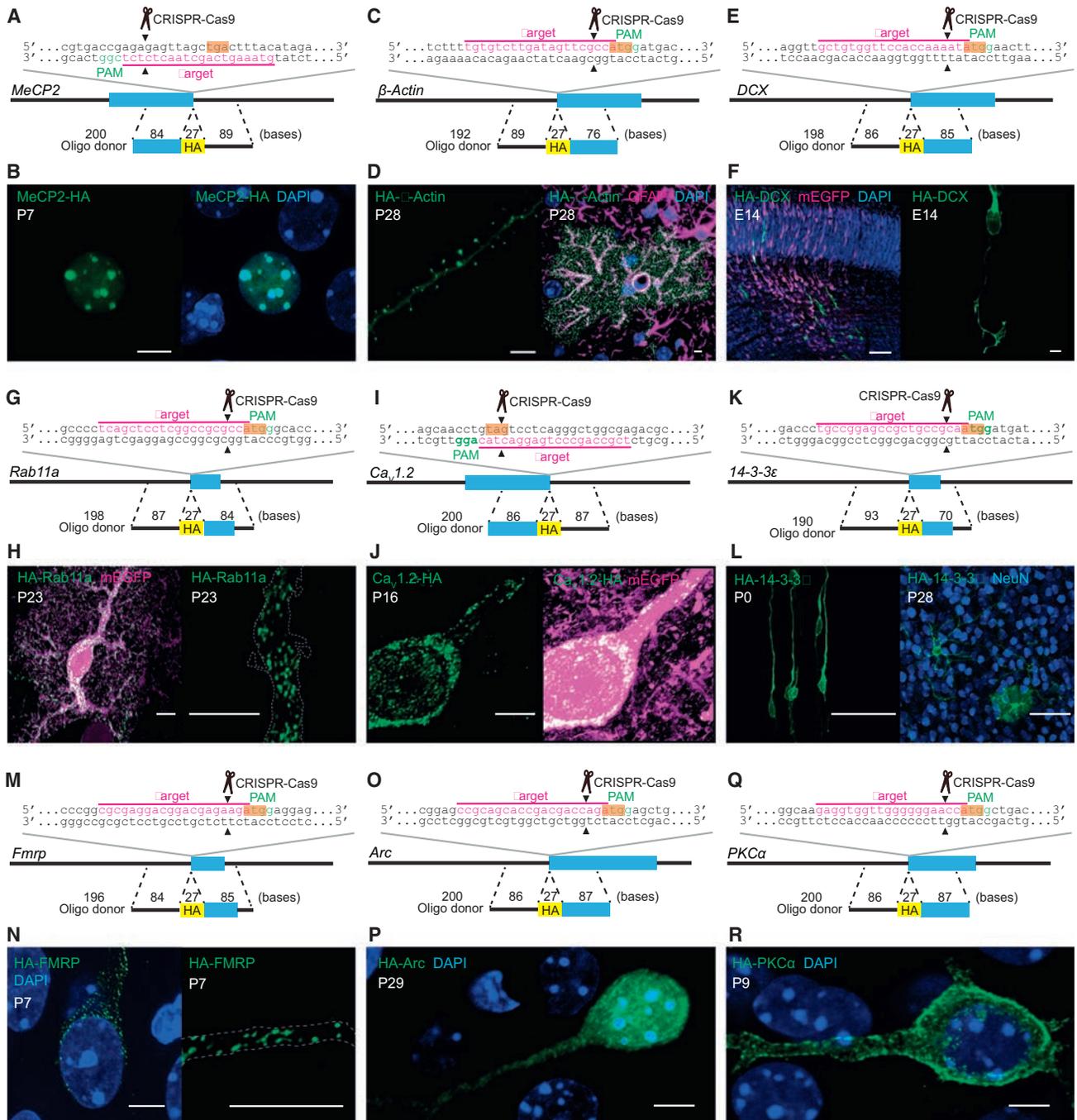
(B and D) Confocal microscopic images of the cerebral cortex electroporated at E12 (top and middle) and E15 (bottom) showing the DAPI signal (blue) and immunoreactivities for mEGFP (magenta) and the HA tag (green) fused to the C terminus of endogenous CaMKII $\alpha$  (B) and N terminus of endogenous CaMKII $\beta$  (D). Middle panels show negative control experiments in which the sgRNA for CaMKII $\alpha$  was paired with the ssODNs for CaMKII $\beta$  (B) and vice versa (D).

(E) The efficiency of SLENDR for CaMKII $\alpha$  and CaMKII $\beta$  (the ratio of the number of HA/mEGFP double-positive neurons to that of mEGFP-positive neurons). CaMKII $\alpha$ , E12 layer (L) 2/3, n = 545 neurons/3 mice; E12 L4–L6, n = 275/3; E15 L2/3, n = 713/3. CaMKII $\beta$ , E12 L2/3, n = 716/3; E12 L4–L6, n = 379/3; E15 L2/3, n = 367/3. \*\*p < 0.01, Dunnett's test, in comparison with E12 L2/3.

(F and G) Left: PCR genotyping using genomic DNA extracted from the electroporated brain. Recombination primer sets (top: F, CaMKII $\alpha$ -F1 and HA-R1; G, HA-F1 and CaMKII $\beta$ -R1) and control primer sets (bottom: F, CaMKII $\alpha$ -F1 and CaMKII $\alpha$ -R1; G, CaMKII $\beta$ -F1 and CaMKII $\beta$ -R1) were used for PCR. Right: DNA sequencing analysis of the PCR products for CaMKII $\alpha$ -HA (F) and HA-CaMKII $\beta$  (G). The HA tag sequence is marked in orange.

Data are represented as mean  $\pm$  SEM. Scale bars, 50  $\mu$ m.

See also Figure S1 and Tables S1, S2, and S3.



**Figure 2. SLENDR Is Scalable to Various Endogenous Proteins**

(A, C, E, G, I, K, M, O, and Q) Graphical representation of the mouse genomic loci of *MeCP2* (A),  $\beta$ -Actin (C), *DCX* (E), *Rab11a* (G), *Ca<sub>v</sub>1.2* (I), *14-3-3 $\epsilon$*  (K), *Fmrp* (M), *Arc* (O), and *PKC $\alpha$*  (Q) showing the target sites for Cas9, sgRNA, and ssODNs. The sgRNA targeting sequences are labeled in magenta. The PAM sequences are labeled in green. The stop (A and I) and start (C, E, G, K, M, O, and Q) codons are marked in orange. The Cas9 cleavage sites are indicated by the black arrowheads.

(B, F, H, J, N, P, and R) Confocal microscopic images of the cerebral cortex showing the DAPI signal (blue) and immunoreactivities for mEGFP (magenta) and the HA tag (green) fused to the C terminus of *MeCP2* (B) and *Ca<sub>v</sub>1.2* (J) and the N terminus of *DCX* (F), *Rab11a* (H), *FMRP* (N), *Arc* (P), and *PKC $\alpha$*  (R). The dashed line represents the shape of the dendrite (H and N).

(D) Images of the cerebral cortex showing the DAPI signal (blue) and immunoreactivities for GFAP (magenta, an astrocyte marker) and the HA tag (green) fused to the N terminus of  $\beta$ -Actin.

(legend continued on next page)

when IUE is performed several days before the final neurogenic divisions of targeted cells.

To test the specificity of the construct, we used incorrect ssODNs-sgRNA pairs, like *CaMKII $\alpha$*  ssODNs-*CaMKII $\beta$*  sgRNA or vice versa. Under these conditions, no fluorescence signal was detected in brain slices stained with anti-HA antibody (Figures 1B and 1D). These results indicate that the genome editing is specific to the sequence of the sgRNA. We further confirmed that the expected genome editing occurred at the DNA level by performing PCR amplification of the targeted locus. Electrophoretic analysis revealed the presence of recombined PCR product at a size consistent with the recombined allele in brains transfected with the constructs necessary for HDR (Figures 1F and 1G). In contrast, neither untransfected control brains nor brains transfected with incorrect ssODNs-sgRNA pairs showed the corresponding PCR amplification (Figures 1F and 1G). Furthermore, DNA sequencing of the amplified PCR products indicated that the HA tag sequence was integrated as expected (Figures 1F and 1G). In addition, sgRNAs targeting different sequences in the respective genes produced a similar pattern of HA staining (Figure S1), demonstrating flexibility of construct design. Thus, SLENDR enables specific single-cell labeling of endogenous proteins by either N- or C-terminal epitope tagging.

### SLENDR Is Scalable to Various Endogenous Proteins

To validate the applicability of SLENDR to visualize the subcellular localization of a broad spectrum of endogenous proteins in brain tissue, we performed SLENDR to fuse the HA tag to either the N or C terminus of a variety of proteins including nuclear, cytoskeletal, vesicular, cytosolic, and membrane proteins (Figures 2A–2R, S2, S3, S4, and S5; Table S1).

We first fused the HA tag to MeCP2, a chromatin-associated protein that regulates gene transcription (Chen et al., 2001). MeCP2 was selected for the following reasons. First, the endogenous subcellular localization can be imaged by traditional immunostaining approaches using a specific, characterized antibody. Second, MeCP2 is known to accumulate in heterochromatin in the nucleus, allowing for the contrast needed for staining of the endogenous protein in tissue. Therefore, double-immunostaining for the SLENDR inserted HA tag and endogenous MeCP2 enabled further validation of the SLENDR approach. As expected, we found that MeCP2-HA and HA-MeCP2 were distributed exclusively in the nucleus at P7 (Figures 2A, 2B, S4B, and S4C). In addition, MeCP2-HA was accumulated in the heterochromatin regions (labeled with DAPI) (Figure 2B) and well co-localized with endogenous MeCP2 (Figure S4A), confirming the specificity of SLENDR. We also evaluated potential effects of SLENDR in transfected cells that did not undergo HDR. Since many mEGFP-positive cells were expected to undergo the error-prone, NHEJ-mediated repair of Cas9-induced DNA double-strand breaks, we designed the targeting sgRNA to place the cleavage site in the 5'-UTR of *MeCP2* gene (Figure S4B). This strategy was designed to minimize potential effects on expres-

sion of the targeted gene in these cells. Indeed, the intensities of endogenous MeCP2 detected by an antibody against the C terminus of MeCP2 was similar between mEGFP-positive and -negative cells (Figure S4D;  $p = 0.19$ , Student's *t* test). This indicated that the endogenous expression of MeCP2 was not significantly disturbed by the NHEJ-mediated repair.

Next, we applied SLENDR to insert the HA tag to endogenous  $\beta$ -Actin, a major cytoskeletal protein in dendritic spines. As expected, immunostaining using the anti-HA antibody showed that HA- $\beta$ -Actin was highly accumulated in dendritic spines in layer 2/3 pyramidal neurons and microfilaments in astrocytes in the cortex at P28 (Yuste and Bonhoeffer, 2004) (Figures 2C and 2D). In addition, the high contrast images demonstrate that SLENDR enables the precise evaluation of the number and morphology of dendritic spines without the potential morphological phenotypes caused by overexpression of fluorophore tagged-actin or actin binding proteins (Riedl et al., 2008).

To evaluate the timescale of HDR after introducing SLENDR constructs, we targeted doublecortin (DCX), a protein that is expressed early in brain development. DCX is a microtubule-associated protein expressed in postmitotic migrating and differentiating neurons in the developing brain (Gleeson et al., 1999). Intriguingly, HA-DCX was detected in the cortical migrating neurons and accumulated in the growth cone as early as 60 hr after IUE at E12 (Figures 2E and 2F), suggesting that HDR occurred rapidly, possibly within 1 or 2 days after IUE. HA-DCX was detected at both E14 and E18 (Figures 2F and S5B), suggesting SLENDR is suitable for studying protein localization in the embryonic brain. In addition, we evaluated the effect of NHEJ on the expression of DCX by immunostaining using an antibody against the C terminus of the protein. The images showed that 98.7% of the mEGFP-positive neurons exhibited DCX expression ( $n = 151$  cells) (Figure S5A). Together with the negligible NHEJ effects on MeCP2 expression (Figure S4D), these data suggest that our strategy of targeting the Cas9-mediated cleavage at 5'-UTRs minimizes the effect of NHEJ on the expression of the target gene for N-terminal tagging.

We further applied SLENDR to a variety of endogenous proteins with distinct subcellular localizations. Rab11a, a small GTPase involved in the endosomal recycling of proteins, was localized to numerous, small dispersed vesicles throughout the soma and dendrites in cortical neurons and astrocytes at P23, consistent with the pattern expected from the localization of recycling endosomes (Hutagalung and Novick, 2011) (Figures 2G, 2H, and S5C).  $Ca_v1.2$ , the  $\alpha1C$  subunit of the L-type voltage-gated calcium channel, was distributed in clusters on cell bodies and proximal dendrites in layer 2/3 pyramidal neurons at P16 (Figures 2I and 2J) (Hell et al., 1993). Immunofluorescence signal of  $Ca_v1.2$ -HA was also detected in the nucleus. This signal may represent the C-terminal fragment of  $Ca_v1.2$ , which functions as a calcium channel-associated transcription regulator (Gomez-Ospina et al., 2006). 14-3-3 $\epsilon$ , a signaling protein involved in neuronal migration and synaptic plasticity, was

(L) Images of the cerebral cortex at P0 and P28 showing immunoreactivities for NeuN (blue, a neuron marker) and the HA tag (green) fused to the N terminus of 14-3-3 $\epsilon$ .

Scale bars, 5  $\mu$ m (B, D, F, right, H, J, N, P, and R); 50  $\mu$ m (F, left; L).

See also Figures S2, S3, S4, and S5 and Tables S1 and S3.

diffusely distributed in the cytoplasm both at P0 and P28 in neurons and astrocytes (Toyo-oka et al., 2014) (Figures 2K and 2L). FMRP, a polyribosome-associated RNA-binding protein that regulates translation of a large number of mRNAs (Contractor et al., 2015), was found in puncta, which likely reflect FMRP-associated mRNA granules, in the soma as well as dendrites in pyramidal neurons at P7 and P28 (Figures 2M, 2N, and S5D). Arc, an immediate early gene product involved in synaptic plasticity, was localized both in the nucleus and cytoplasm at P29 (Figures 2O and 2P) (Korb et al., 2013; Shepherd and Bear, 2011). Finally, we found that  $\alpha$  isoform of protein kinase C (PKC), a member of a family of serine/threonine kinases implicated in a wide range of cellular functions (Steinberg, 2008), was distributed mostly on the plasma membrane of the soma and dendrites at P9 (Figures 2Q and 2R). Interestingly, PKC $\alpha$  was less accumulated on the plasma membrane at P27, suggesting that endogenous PKC $\alpha$  may be more active at the developing stage (Figures S5E–S5G). Thus, SLENDR enables us to determine the distribution pattern of proteins that has been undefined or controversial by conventional methods.

Notably, we never observed HA-positive cells when we used incorrect ssODNs-sgRNA pairs (Figures S2 and S4). In addition, precise genome editing was confirmed by PCR amplification of the targeted loci followed by DNA sequencing of the PCR products (Figure S3). Furthermore, for each target protein, all HA-positive cells in the same region showed similar HA-staining pattern. These results collectively demonstrate that SLENDR enables us to label endogenous proteins specifically. Taken together, SLENDR is a highly generalizable technique that can be applied to various species of proteins in the brain from embryonic to adult stages for high-quality mapping of subcellular localization.

### Nanometer-Resolution Analysis by SLENDR

Immunoelectron microscopy allows nanoscale visualization of endogenous proteins with defined ultrastructures in cells. However, the lack of reliable antibodies compatible with electron microscopy imaging limits its application to a variety of proteins. Thus, we tested whether SLENDR could be applied to ultrastructural imaging of endogenous proteins using immunoelectron microscopy. To do so, we used cortical tissue in which endogenous CaMKII $\beta$  is fused with HA using SLENDR (Figures 1C and 1D). We applied pre-embedding staining technique: we incubated the tissue with HA-antibody and secondary antibody conjugated with gold, followed by silver enhancement and tissue embedding. We then prepared serial thin-section (50 nm) of the tissue using the automatic tape-collecting ultramicrotome (ATUMtome). ATUMtome permits rapid and automated cutting and collection of serial thin sections onto a continuous reel of tape (Kasthuri et al., 2015). We imaged a number of serial sections on the tape by scanning electron microscopy and reconstructed three-dimensional images of SLENDR-generated knockin cells (Figures 3A–3D). We found that CaMKII $\beta$  was localized near the postsynaptic density (PSD) (mode  $\sim$ 80 nm) in dendritic spines (Figure 3F). Finally, the specificity of the immunogold labeling was examined by comparing the labeling density in HA-positive and surrounding neurons in the same specimen (Figure 3E; HA-positive,  $241.3 \pm 34.5$  particles/ $\mu\text{m}^3$ ; surrounding

control,  $2.6 \pm 2.6$  particles/ $\mu\text{m}^3$ ;  $p < 0.001$ , Student's *t* test). These experiments demonstrate that SLENDR is useful for nanometer scale localization of endogenous proteins.

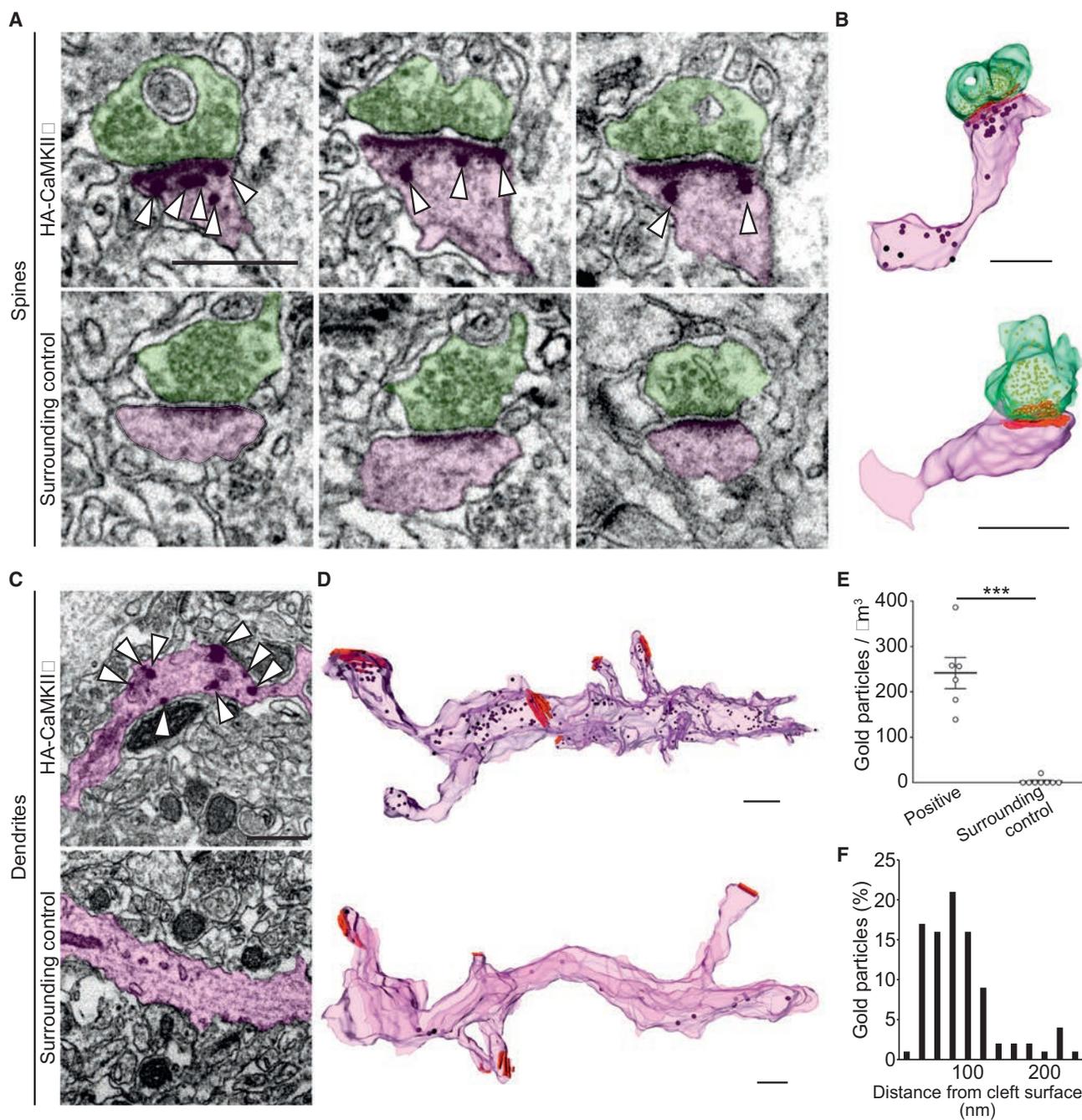
### SLENDR Is Scalable to Various Cell Types in Various Brain Regions

We then tested if SLENDR can be applied to different cell types in different brain regions by targeting distinct progenitor cells present at different locations and timings in the developing brain (Borrell et al., 2005; Chen and LoTurco, 2012; Kitazawa et al., 2014; Nishiyama et al., 2012; Soma et al., 2009). By adjusting the timing and electroporation angle (Figure 4A), various HA-tagged endogenous proteins ( $\beta$ -Actin, CaMKII $\alpha$ , CaMKII $\beta$ , and MeCP2) were observed in various cell types in widespread brain regions (Table S2), including CA1 pyramidal neurons and dentate granule cells in the hippocampus (Figures 4B and 4H), spiny stellate cells in the subiculum (Figures 4C and 4H), granule cells in the olfactory bulb (Figure 4D), medium spiny neurons in the striatum (Figure 4E), basolateral amygdala neurons (Figure 4F), and granule and Purkinje cells in the cerebellum (Figures 4G and 4H). HA-CaMKII $\beta$  in parallel fibers originating from cerebellar granule cells was detected along the dendrites of Purkinje cells in the molecular layer (Figure 4G).

As we showed that staining of endogenous HA- $\beta$ -Actin enables clear visualization of dendritic spine morphology in neurons (Figure 2D), we imaged HA- $\beta$ -Actin in various cell types in the brain to compare the structure of dendritic spines. Morphological diversity of dendritic spines existed among different cell types including dentate granule cells (Figure 4H) and CA1 pyramidal neurons (Figure 4B) in the hippocampus, layer 2/3 pyramidal neurons in the cortex (Figure 2D), spiny stellate cells in the subiculum (Figure 4H), and cerebellar Purkinje cells (Figure 4H). Thus, SLENDR enables comparison of subcellular localization of endogenous proteins in various neuron subtypes and brain regions across developmental stages, providing regional and developmental specific information about protein localization.

### Labeling Multiple Endogenous Proteins by SLENDR

Multiplex labeling of different endogenous proteins with different tags would enable high-resolution co-localization assays. Since CRISPR-Cas9 is able to target multiple genes simultaneously (Cong et al., 2013; Heidenreich and Zhang, 2016; Swiech et al., 2015), we tested the ability of SLENDR to target two different proteins in the same cell (Figure 5A). We performed simultaneous labeling of CaMKII $\alpha$  and CaMKII $\beta$  with the HA and FLAG tag, respectively, by co-introducing SLENDR constructs for HA-CaMKII $\alpha$  and FLAG-CaMKII $\beta$  into the developing cortex at E13 by IUE (Figure 5A). We performed immunostaining of brain slices at P14 using anti-HA and anti-FLAG antibodies. HA-CaMKII $\alpha$  and FLAG-CaMKII $\beta$  were detected in a number of layer 2/3 neurons (HA-positive/mEGFP-positive,  $4.5\% \pm 0.4\%$ ; FLAG-positive/mEGFP-positive,  $4.9\% \pm 0.6\%$ ;  $n = 577$  cells), and a significant fraction of neurons exhibited both HA and FLAG signals (HA and FLAG double-positive/mEGFP-positive,  $0.7\% \pm 0.1\%$ ;  $n = 577$  cells) (Figure 5B). This double labeling efficiency was higher than the simple multiplication of each labeling efficiency ( $4.5\% \times 4.9\% = 0.2\%$ ), demonstrating the practicality of the method for co-localization assay. Thus, SLENDR allows



### Figure 3. Nanometer-Scale Analysis of Endogenous Proteins by SLENDR

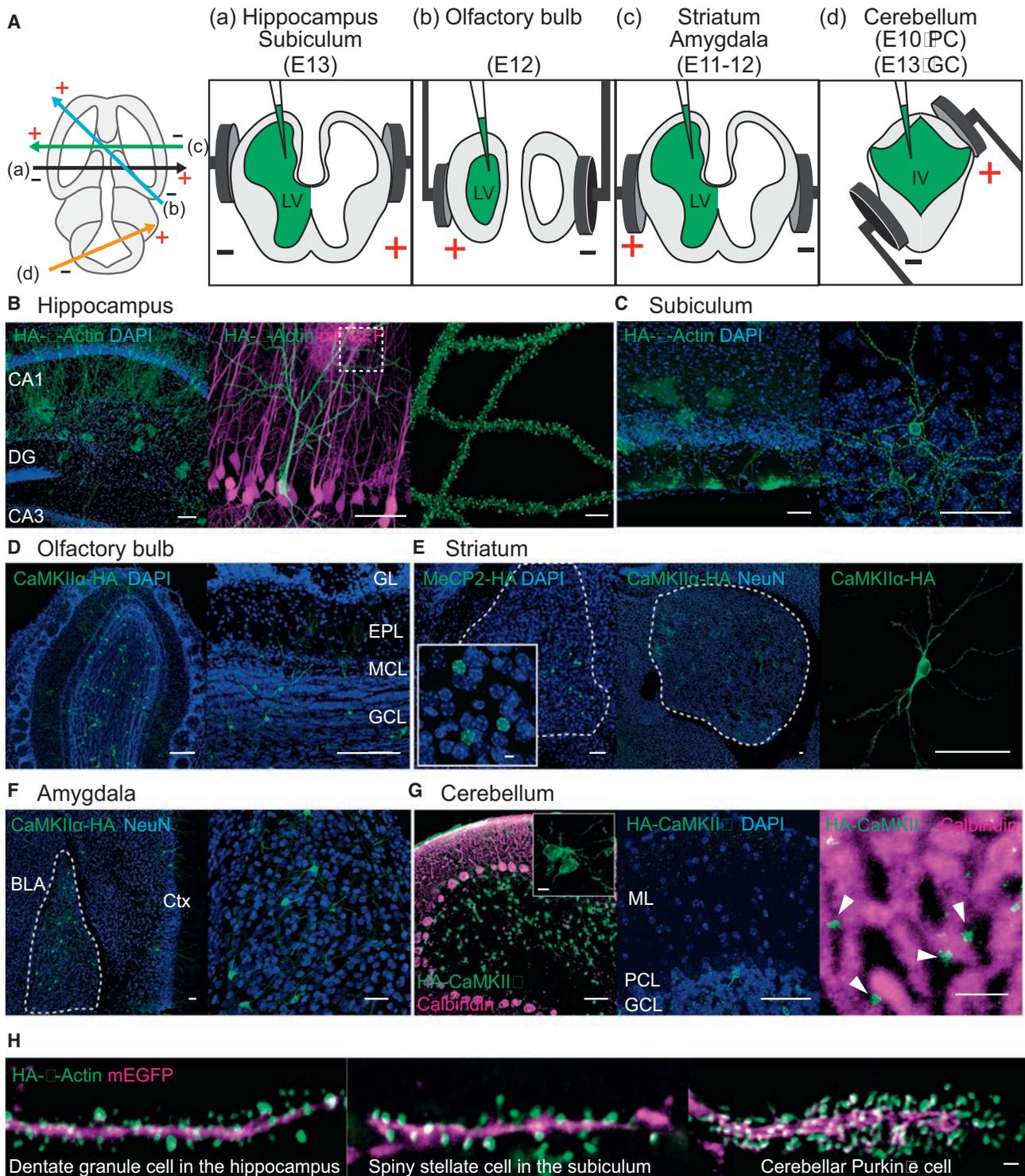
(A–D) Electron microscopic images of dendritic spines (A) and shafts (C) in the cerebral cortex showing immunogold labeling for the HA tag (arrowheads) fused to the N terminus of CaMKII $\beta$  in HA-positive (top) and surrounding control (bottom) cells. Three-dimensional reconstructions of corresponding spines (B) and dendrites (D). Analyzed spines and dendrites are marked in purple and presynaptic terminals are marked in green. Presynaptic vesicles and postsynaptic densities are marked in yellow and red, respectively (B and D).

(E) Density of immunogold particles in HA-positive spines ( $n = 6$ ) and surrounding control spines ( $n = 8$ ). \*\*\* $p < 0.001$ , Student's  $t$  test.

(F) The frequency distribution of distances between individual gold particles and cleft surfaces ( $n = 92$  particles/6 spines).

Data are represented as mean  $\pm$  SEM. Scale bars, 500 nm.

See also Table S3.

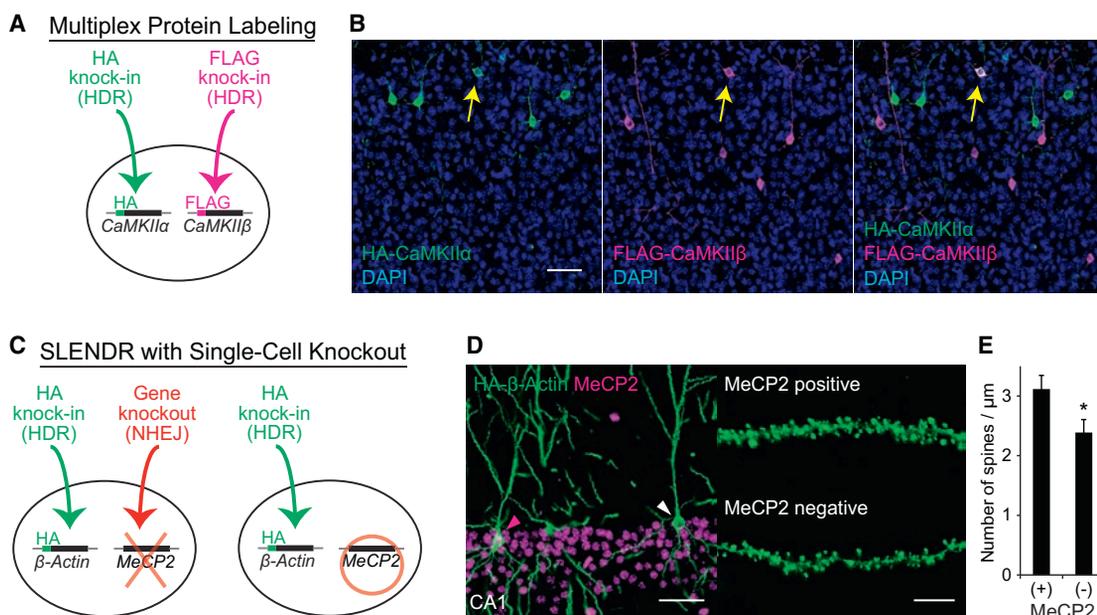


**Figure 4. SLENDR Is Scalable to Various Cell Types in Various Brain Regions**

(A) Schematic illustration of IUE for targeting distinct brain regions. The relative position of the electrodes (+, positive pole; −, negative pole) are shown in the transverse section of the brain to target different brain areas (left). The position of electrode paddles and the injected DNA (green) are shown in the coronal section of the brain to target hippocampus and subiculum (a), olfactory bulb (b), striatum and amygdala (c), and cerebellum (d). LV, lateral ventricle; IV, fourth ventricle; PC, Purkinje cell; GC, granule cell.

(B) Confocal microscopic images of the hippocampus showing the DAPI signal (blue) and immunoreactivities for mEGFP (magenta) and the HA tag (green) fused to β-Actin. DG, dentate gyrus.

(legend continued on next page)



**Figure 5. Application of SLENDR by Multiplex Genome Editing**

(A) Schematics of multiplex labeling of different endogenous proteins with different tags. The HA and FLAG sequences are inserted to *CaMKIIα* and *CaMKIIβ*, respectively, in the same cell through HDR-mediated genome editing.

(B) Multiplex labeling of endogenous *CaMKIIα* and *CaMKIIβ*. Confocal microscopic images of the cerebral cortex at P14 showing the DAPI signal (blue) and immunoreactivities for the HA tag (green) and the FLAG tag (magenta) fused to the N terminus of endogenous *CaMKIIα* and *CaMKIIβ*, respectively. The yellow arrows indicate HA and FLAG double-positive layer 2/3 neurons.

(C) Schematics of combining SLENDR with NHEJ-mediated gene knockout. The HA sequence is inserted to *β-Actin* through HDR-mediated genome editing and a frame-shift mutation is induced in *MeCP2* through NHEJ-mediated genome editing in the same cell (left). In our strategy, some cells in the same tissue undergo only HDR-mediated genome editing (right), allowing comparison of the expression and localization of endogenous proteins within the same brain slice.

(D) Immunofluorescence images of MeCP2 (magenta) and HA-*β-Actin* (green) in the hippocampus CA1 region. Magenta arrowhead, MeCP2-positive; white arrowhead, MeCP2-negative. Representative images of secondary dendrites of MeCP2-positive and -negative neurons (right).

(E) The averaged density of spines on secondary or tertiary apical dendrites in MeCP2-positive (n = 527 spines/6 neurons) and -negative (n = 437/5) neurons. \*p < 0.05, Student's t test.

Data are represented as mean ± SEM. Scale bars, 50 μm (B and D, left); 5 μm (D, right).

See also Table S3.

labeling of two different species of proteins with different tags in single cells, providing a valuable tool for co-localization assays of a pair of endogenous proteins.

### SLENDR in Combination with Single-Cell Knockout

The ability to examine endogenous subcellular protein localization in the context of a knockout of a different protein would provide functional insight into the interaction between the visualized and deleted proteins. In this regard, combining SLENDR with NHEJ-based single-cell knockout would be of particular interest, since it would allow for the study of cell-autonomous

gene function and to compare normal and knockout cells in the same tissue (Zong et al., 2005). Taking advantage of the multiplexity of CRISPR-Cas9, we simultaneously introduced SLENDR constructs to insert the HA tag to *β-Actin* and CRISPR constructs to induce NHEJ-mediated gene knockout of *MeCP2* in progenitors of hippocampal neurons at E13 (Figure 5C) (Incontro et al., 2014; Straub et al., 2014; Swiech et al., 2015). Visualization of HA-tagged *β-Actin* by SLENDR enables visualization of dendritic spine morphology, providing a useful tool to study effects of *MeCP2* gene deletion on dendritic spines in single cells.

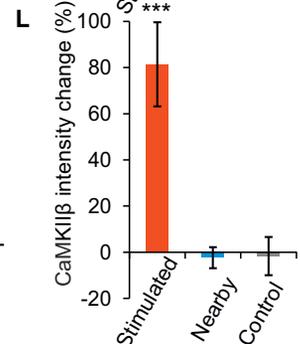
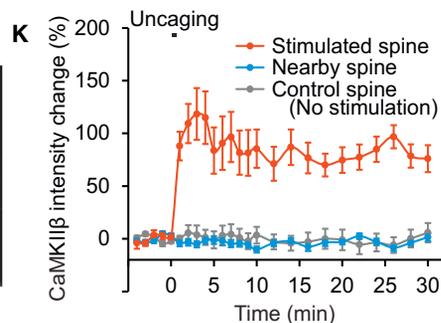
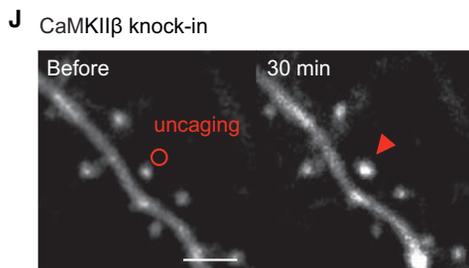
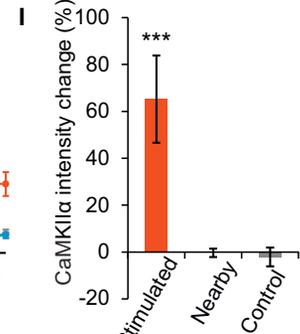
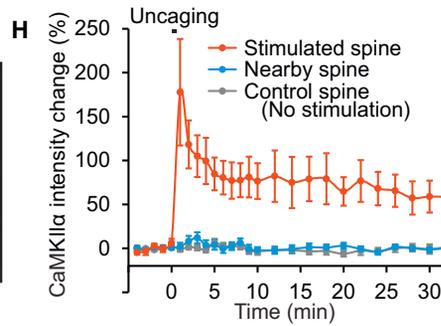
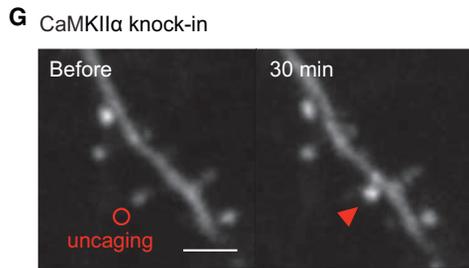
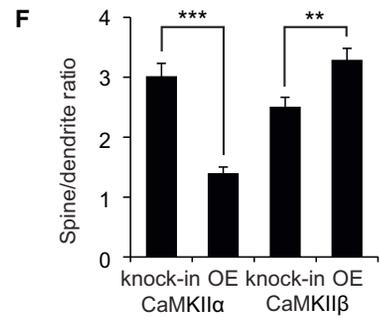
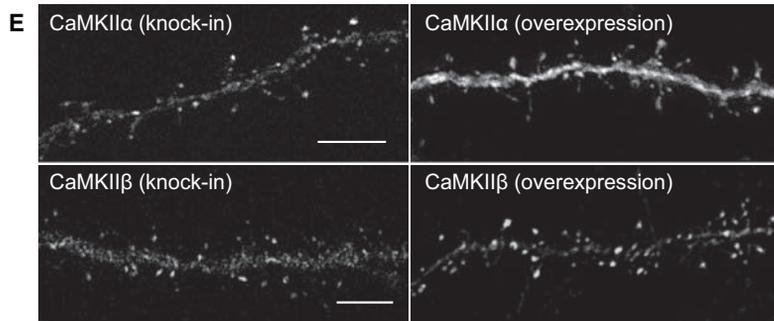
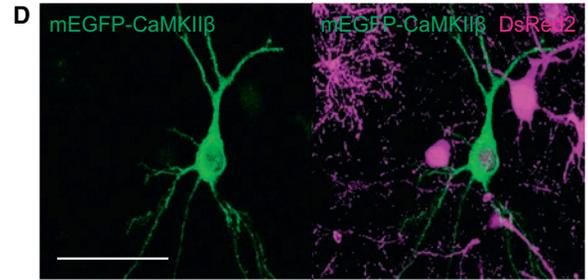
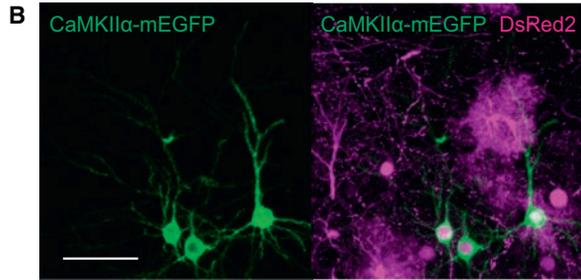
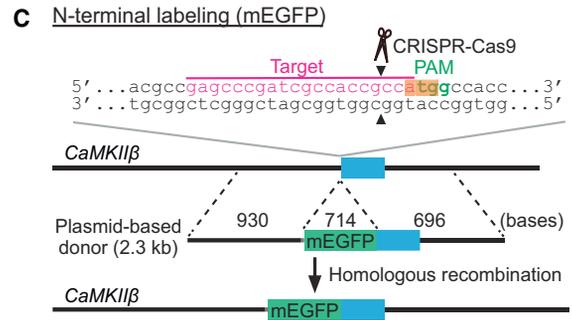
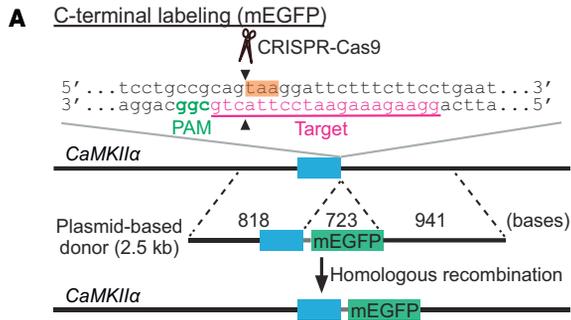
(C–F) Images of the subiculum (C), olfactory bulb (D), striatum (E), and amygdala (F) showing the DAPI signal (C–E, blue) and immunoreactivities for NeuN (E and F, blue) and the HA tag (green) fused to *β-Actin* (C), *CaMKIIα* (D to F), and *MeCP2* (E). GL, glomerular layer; EPL, external plexiform layer; MCL, mitral cell layer; GCL, granule cell layer; BLA, basolateral amygdala; Ctx, cerebral cortex.

(G) Images of the cerebellum showing the DAPI signal (blue) and immunoreactivities for calbindin D-28k (magenta, a Purkinje cell marker) and the HA tag (green) fused to *CaMKIIβ*. ML, molecular layer; PCL, Purkinje cell layer; GCL, granule cell layer.

(H) Images of the dendrites of a dentate granule cell in the hippocampus, a spiny stellate cell in the subiculum and a Purkinje cell in the cerebellum showing immunoreactivities for mEGFP (magenta) and the HA tag (green) fused to *β-Actin*.

Scale bars, 50 μm (B–G); 5 μm (B, right; E, inset; G, inset and right); 1 μm (H).

See also Tables S2 and S3.



(legend on next page)

While knockout efficiency of NHEJ-mediated genome editing has been reported to be very high in neurons (~70%–100%) (Heidenreich and Zhang, 2016; Incontro et al., 2014; Straub et al., 2014; Swiech et al., 2015), immunohistochemistry with anti-MeCP2 antibody showed that MeCP2 was absent in 48.8% of mEGFP-positive CA1 pyramidal neurons at P28. This reduced efficiency was probably due to the dilution of the knockout constructs during cell divisions (Chen and LoTurco, 2012; Yusa et al., 2011). As expected, both MeCP2-negative and positive neurons were observed in HA- $\beta$ -Actin-positive CA1 neurons in the same slices (Figure 5D). Notably, the knockout efficiency in HA- $\beta$ -Actin-positive neurons (24.1%) was substantially lower than that in whole population (48.8%). This is presumably because the knockout constructs are diluted in a population of cells that divides extensively, and this population of cells likely provides more efficient HDR. A loss of function mutation in *MeCP2*, which leads to Rett syndrome in humans, has been reported to impair maturation of dendritic spines (Chen et al., 2001; Swiech et al., 2015). Therefore, we compared the density of dendritic spines between MeCP2-negative and -positive neurons (Figures 5D and 5E). Spine density in the apical dendrite of CA1 neurons was comparable in MeCP2-positive neurons to that reported in these neurons in a previous electron microscopic study (Harris et al., 1992). Importantly, spine density was significantly lower in MeCP2-negative neurons, consistent with previous studies (Figure 5E;  $p = 0.046$ , Student's t test) (Swiech et al., 2015).

Although high-quality antibodies against endogenous proteins are required to distinguish knockout and control cells, our strategy combining SLENDR with NHEJ-mediated gene knockout provides a useful means to investigate cell-autonomous effects of a gene of interest on endogenous proteins.

### Live Imaging of Endogenous Proteins by SLENDR

Labeling endogenous proteins with fluorescent proteins would allow us to image protein dynamics in live cells without overexpression artifacts, further expanding the applicability of SLENDR. Thus, we tested if SLENDR can be used to insert a long sequence encoding a fluorescent protein into a gene of interest. For this, we targeted *mEGFP* to the sequence just upstream of the stop codon of *CaMKII $\alpha$*  or downstream of the start codon of *CaMKII $\beta$*  in the genome (Figures 6A and 6C). We used a plasmid-based template (~2.3–2.5 kb) containing the ~0.7 kb sequence of *mEGFP* (see Supplemental Information). The SLENDR con-

structs together with the transposon vectors expressing DsRed2 were introduced to progenitors of layer 2/3 neurons through IUE at E12. At P14–P28, we found that a sparse subset of layer 2/3 neurons exhibited mEGFP fluorescence fused to endogenous *CaMKII $\alpha$*  and *CaMKII $\beta$*  (Figures 6B and 6D). The efficiency of mEGFP knockin was lower than that of HA knockin (mEGFP-positive/DsRed2-positive, <1%), consistent with previous studies showing that a long sequence insertion through HDR is less efficient (Cox et al., 2015). The specificity of mEGFP knockin was confirmed by control experiments using correct and incorrect template-sgRNA pairs and DNA sequencing following PCR amplification of the targeted locus (Figure S6).

To further verify the benefit of imaging endogenous proteins, we compared the distribution pattern of endogenous and overexpressed mEGFP-tagged *CaMKII* in layer 2/3 cortical neurons at P14. Since *CaMKII $\beta$*  has an actin-binding domain (Lisman et al., 2012), overexpressed *CaMKII $\beta$*  tends to accumulate in spines to a greater extent than overexpressed *CaMKII $\alpha$* . Notably, the degree of enrichment of endogenous *CaMKII $\alpha$*  and *CaMKII $\beta$*  in spines, as measured by the spine/dendrite ratio of the peak intensity, was significantly higher and lower than that of overexpressed *CaMKII $\alpha$*  and *CaMKII $\beta$* , respectively (Figures 6E and 6F; *CaMKII $\alpha$* ,  $p < 0.001$ ; *CaMKII $\beta$* ,  $p = 0.004$ , Student's t test). This is perhaps because overexpressed subunits tend to form homomeric enzymes, while endogenous *CaMKII $\alpha$*  and *CaMKII $\beta$*  tend to form heteromers (Lisman et al., 2012). Thus, the localization of endogenous (heteromeric) subunits would be enriched in spines to an extent that is between homomeric *CaMKII $\alpha$*  and *CaMKII $\beta$* .

We next performed live imaging to monitor the synapse-specific translocation of *CaMKII* after the stimulation of a single dendritic spine (Bosch et al., 2014; Lee et al., 2009; Lisman et al., 2012; Nishiyama and Yasuda, 2015). We prepared organotypic cortical slice cultures from mice in which endogenous *CaMKII $\alpha$*  or *CaMKII $\beta$*  was labeled with mEGFP by SLENDR. When we applied a train of two-photon glutamate uncaging pulses (1 Hz, 30 s) at single spines, mEGFP-tagged *CaMKII $\alpha$*  and *CaMKII $\beta$*  were rapidly accumulated in the stimulated spines and remained for more than 30 min. The translocation was highly restricted to the stimulated spine, with no significant increase in fluorescence intensity in the surrounding spines (Figures 6G–6L; *CaMKII $\alpha$* ,  $p < 0.001$ ; *CaMKII $\beta$* ,  $p < 0.001$ ; Dunnett's test). These results demonstrate that SLENDR enables monitoring of the dynamics of endogenous proteins in living tissue.

### Figure 6. Localization and Dynamics of Endogenous Proteins Labeled with mEGFP by SLENDR

(A and C) Graphical representation of the mouse genomic loci of *CaMKII $\alpha$*  (A) and *CaMKII $\beta$*  (C) showing the targeting sites for Cas9, sgRNA, and HDR donor plasmid. The sgRNA targeting sequences are labeled in magenta. The PAM sequences are labeled in green. The stop and start codons of *CaMKII $\alpha$*  (A) and *CaMKII $\beta$*  (C) are marked in orange.

(B and D) Confocal microscopic images of the somatosensory cortex showing the fluorescence of DsRed2 (magenta) and mEGFP (green) fused to *CaMKII $\alpha$* / $\beta$ .

(E) Images of apical secondary dendrites in layer 2/3 fixed at P14 showing mEGFP-tagged endogenous (knockin) or overexpressed (OE) *CaMKII $\alpha$* / $\beta$ .

(F) The spine/dendrite ratio of the peak intensities of mEGFP-tagged *CaMKII $\alpha$* / $\beta$ . *CaMKII $\alpha$* , knockin,  $n = 51/6$  (spines/neurons); OE,  $n = 40/5$ . *CaMKII $\beta$* , knockin,  $n = 58/4$ ; OE,  $n = 83/5$ .

(G–L) Two-photon microscopic images before and 30 min after glutamate uncaging showing mEGFP-tagged endogenous *CaMKII $\alpha$* / $\beta$  in layer 2/3 neurons (G and J). Red circles, uncaging spots; red arrowheads, stimulated spines. Averaged time courses (H and K) and sustained values (I and L; averaged over 20–30 min) of *CaMKII $\alpha$* / $\beta$  intensity change in the stimulated (red; H and I,  $n = 11/8$ ; K and L,  $n = 9/7$ ), nearby (2–5  $\mu$ m from the stimulated spines; blue; H and I,  $n = 31/8$ ; K and L,  $n = 31/7$ ) and control spines with no stimulation (gray; H and I,  $n = 23/5$ ; K and L,  $n = 13/2$ ). \*\*\* $p < 0.001$ , Student's t test (F) and Dunnett's test (I and L).

Data are represented as mean  $\pm$  SEM. Scale bars, 50  $\mu$ m (B and D); 5  $\mu$ m (E); 2  $\mu$ m (G and J).

See also Figure S6 and Table S3.

## DISCUSSION

In this study, we have developed SLENDR, which allows in vivo genome editing in the mammalian brain for single-cell labeling of endogenous proteins. We demonstrated that SLENDR is a simple and efficient technique to rapidly determine the subcellular localization of endogenous proteins with the resolution of micro- to nanometers in brain tissue. Importantly, the technique is generalizable to a broad spectrum of proteins and various cell types in widespread brain regions. SLENDR also can be used for multiplex labeling of different proteins or for mosaic analysis by combining labeling with single-cell knockout. Furthermore, SLENDR is capable of inserting a long sequence such as that encoding mEGFP in vivo and thus enables live imaging of endogenous proteins during biological processes in the brain.

HDR-mediated genome editing has been a challenge in the brain due to the lack of HDR activity in postmitotic neurons and the inefficient delivery of HDR machinery into the brain. This has limited its application in the field of neuroscience (Heidenreich and Zhang, 2016). We circumvented these problems by targeting mitotic progenitors in the embryonic brain using IUE. In our experiments, the knockin efficiency for the HA tag insertion was sufficient to image the subcellular localization of a protein of interest in single cells in brain tissue. We also showed that SLENDR can be used to knockin a long *mEGFP* sequence, albeit at lower efficiency. To extend the SLENDR technique to broader applications for multiplexed labeling and fluorescent protein fusion, it may be necessary to increase the efficiency of HDR.

It is known that the efficiency of HDR highly depends on the cell cycle and is limited by the competition with NHEJ (Chu et al., 2015; Lin et al., 2014; Maruyama et al., 2015). Consistent with this, we found that SLENDR was efficient in a limited time window when neuronal progenitors were still dividing. CRISPR-Cas9-mediated HDR occurred within a few days after IUE (Figure 2F) and the HDR efficiency was significantly reduced when IUE was performed near the stage of the final cell divisions of target cells (Figures 1B and 1D). In this regard, the direct delivery of pre-assembled Cas9 protein-guide RNA ribonucleoprotein complexes (RNPs), rather than expressing these components from plasmids, might increase the time window for SLENDR, as RNPs have recently been reported to provide rapid nuclease action with high efficiency and low off-target effects for genome editing (Kim et al., 2014; Lin et al., 2014). Alternatively, genetic or pharmacological inhibition of the NHEJ pathway may increase the HDR efficiency (Chu et al., 2015; Maruyama et al., 2015), although potential side effects must be addressed.

We validated the specificity of the SLENDR-mediated sequence insertion in the genome in several ways. First, we designed sgRNAs by unbiased genome-wide analysis to minimize the potential off-target cleavages by Cas9 (Ran et al., 2013). Second, we performed control experiments by using incorrect sgRNAs and found no cells expressing the HA tag. Third, we used distinct sgRNAs targeting the identical gene and observed similar HA staining patterns. Fourth, the localization of HA staining was consistent with that previously reported based on immunohistochemical, biochemical or electron microscopic studies. Fifth, the localization of HA staining was similar from

cell to cell. In addition, although off-target cleavage by Cas9 may be induced in the genome, HDR is unlikely to occur at the off-target sites, because homologous recombination is known to be a highly sequence-specific event. The specificity of SLENDR could be further enhanced by using recently reported Cas9 variants with minimum or no off-target effects while retaining comparable on-target cleavage activity (Kleinstiver et al., 2016; Slaymaker et al., 2016).

One of the merits of SLENDR is that high-quality antibodies can be used for the detection of tags. In addition, once staining conditions have been optimized for a tag, SLENDR can be applied to image various tagged proteins without extensive optimization. Furthermore, since SLENDR allows protein labeling in a sparse subset of cells in the tissue, the specificity of immunostaining can be easily validated by examining surrounding negative control cells in the same specimen. These features are particularly advantageous for immunoelectron microscopy imaging, which often requires extensive optimization of staining conditions and good control samples (e.g., knockout mice).

We have demonstrated that introducing a single epitope tag by SLENDR is sufficient to detect relatively low abundant proteins such as  $Ca_v1.2$  and Arc (Figures 2J and 2P). The sensitivity could be further increased by inserting multiple copies of epitopes or more antigenic probes such as Spaghetti-monster (Viswanathan et al., 2015). Overall, because many studies are constrained by the lack of high-quality antibodies against proteins of interest, SLENDR will provide a generalizable and reliable platform for exploring localization of uncharacterized endogenous proteins using light and electron microscopy.

Multiplexing CRISPR-Cas9-mediated processes provided some insight into how HDR and NHEJ occur following double strand breaks. We found that the cells labeled with a tag with SLENDR, or in other words the cells that underwent HDR, showed less NHEJ-mediated knockout (Figures 5C–5E). This result suggests that HDR occurs more frequently in a population of highly dividing cells. Consistent with this, the efficiency of double-labeling via multiple HDRs was higher than that obtained by the simple multiplication of each labeling efficiency (Figures 5A and 5B). The high efficiency of double-labeling by SLENDR opens the possibility of co-localization assay for a pair of endogenous proteins with high resolution and contrast. Given relatively consistent knockin efficiency (Table S1), the SLENDR-based double labeling technique would be applicable to many pairs of proteins.

There are a few limitations in SLENDR. First, the technique might lead to unintended sequence changes through insertion/deletion mutations mediated by NHEJ in the transfected cells. This problem may cause a large impact particularly in the case of N-terminal tagging, since altered sequence near the N-terminal region could change the transcription and translation of the target gene. Therefore, throughout this study, we attempted to minimize the effect of on-target NHEJ by choosing target sequences at 5'-UTR or near the stop codon for N- or C-terminal tagging, respectively. As expected from this design, the expression of DCX was detected in 98.7% of the mEGFP-positive neurons following SLENDR for the N-terminal tagging of DCX (Figure S5A). In addition, the expression level of MeCP2 was not significantly affected for the N-terminal tagging of MeCP2 (Figure S4D). However, potential effects of NHEJ have to be

carefully evaluated for each gene depending on the purpose of experiments. For example, when the target is a secreted or a type I membrane protein with a signal sequence, one should select a target sequence for CRISPR-Cas9-mediated cleavage to minimize the possibility of deletion or mislocalization of the gene products. Second, immunodetection of a fused epitope tag may be difficult for some proteins, because the accessibility of antibodies may be sterically limited. For example, we failed to detect the HA signal with immunostaining in tissue in which the HA tag was inserted to the C terminus of PSD-95 with SLENDR, whereas PCR detected the HA knockin allele at the DNA level (data not shown). This is likely due to high protein density of PSD, which is known to often prevent antibodies to access to targeted proteins in the structure (Fukaya and Watanabe, 2000). This limitation of immunostaining can be overcome by fusing a fluorescent protein tag with SLENDR and directly observing the fluorescence (Fortin et al., 2014). These limitations would not detract the impact of SLENDR, as the technique should be applicable to most proteins with little optimization as shown in this study.

Since the 5'-NGG protospacer-adjacent motif (PAM) of SpCas9 abundantly exists in the genome and other CRISPR endonucleases with different PAMs are also available to target different sites in the genome (Cong et al., 2013; Hsu et al., 2014; Zetsche et al., 2015), most proteins should be suitable for SLENDR. Furthermore, its throughput and cost-effectiveness are much higher than previous single-cell protein labeling methods (Fortin et al., 2014; Gross et al., 2013). Thus, SLENDR should allow large-scale, potentially genome-wide, determination of precise subcellular localization of endogenous proteins in various cell types and ages, providing a new level of understanding of protein and cellular function in the brain.

## EXPERIMENTAL PROCEDURES

Full experimental details can be found in the Supplemental Experimental Procedures.

### Animals

All experimental procedures were approved by the Max Planck Florida Institute for Neuroscience Institutional Animal Care and Use Committee and were performed in accordance with guidelines from the US NIH. Swiss Webster mice were obtained from Charles River. The day on which the vaginal plug was detected was designated as E0. The first 24 hr after birth was referred to as P0. Both of male and female mice were used.

### DNA Constructs

The human codon-optimized SpCas9 and sgRNA expression plasmid was a gift from F. Zhang (pX330, Addgene plasmid #42230) (Cong et al., 2013). To generate sgRNA-expressing plasmids, a pair of annealed oligos (~20 bp) was ligated into the sgRNA scaffold of pX330 (Ran et al., 2013). We purchased single-stranded oligodeoxynucleotides (ssODNs) for HDR from Integrated DNA Technologies. To fuse mEGFP (monomeric EGFP, A206K) to the C or N terminus of CaMKII $\alpha$  or CaMKII $\beta$ , respectively, plasmid-based donor templates for HDR were prepared. Other details are described in the Supplemental Experimental Procedures.

### In Utero Electroporation

In utero electroporation was performed as previously described (Borrell et al., 2005; Chen and LoTurco, 2012; Imamura and Greer, 2015; Kitazawa et al., 2014; Nishiyama et al., 2012; Soma et al., 2009; Tabata and Nakajima,

2001). The final concentration of each plasmid (pX330-derivatives, pPB-CAG-mEGFP, pPB-CAG-DsRed2, pPB-CAG-tdTomato, and pCAG-hyPBBase), the ssODNs for HDR and the double-stranded DNA template for mEGFP insertion were 1  $\mu$ g/ $\mu$ l, 20  $\mu$ M, and 1  $\mu$ g/ $\mu$ l, respectively (Table S3). The position and angle of the electrode was set as previously described with some modifications (see Figure 4A). Other details are described in the Supplemental Experimental Procedures.

### Histology

Mice were fixed by transcardial perfusion with 4% paraformaldehyde in 0.1 M phosphate buffer and coronal sections (50 or 100  $\mu$ m in thickness) were prepared using a vibratome (VT1200, Leica). Immunohistochemistry was performed as previously described (Fukaya and Watanabe, 2000; Nishiyama et al., 2012). Other details are described in the Supplemental Experimental Procedures.

### Pre-embedding Immunoelectron Microscopy

An HA-CaMKII $\beta$  knockin mouse was transcardially perfused with 4% paraformaldehyde and 0.1% glutaraldehyde in 0.1 M phosphate buffer. The fixed brain was sectioned into 50- $\mu$ m thick slices with a vibratome (VT 1200, Leica), and sections with positive DsRed2 labeling were chosen. Tissue processing for pre-embedding immunoelectron microscopy was performed as previously described (Parajuli et al., 2012). After processing, a region containing HA-positive neurons was trimmed out, and 50-nm thick sections were cut and collected onto a Kapton tape by ATUMtome (RMC/Boeckeler) (Kasthuri et al., 2015; Schalek et al., 2011). The Kapton tape was placed on a silicon wafer, and a layer of 5-nm-thick carbon was coated on the wafer surface with a high vacuum sputter coater (ACE600, Leica). The sections on the wafer were imaged under a scanning electron microscope (SEM, Merlin VP Compact, Zeiss) assisted with Atlas 5 AT software (Zeiss). Other details are described in the Supplemental Experimental Procedures.

### Two-Photon Glutamate Uncaging

Organotypic cortical slice cultures were prepared from mEGFP-CaMKII $\alpha$ / $\beta$  knockin mice as described previously (Stoppini et al., 1991; Yamamoto et al., 1989). Secondary apical dendritic branches of layer 2/3 neurons were imaged at 15–17 days in vitro under a custom-built two-photon microscope with two Ti:Sapphire lasers (Chameleon, Coherent) (Lee et al., 2009). MNI-caged L-glutamate (4-methoxy-7-nitroindolyl-caged L-glutamate, Tocris) was uncaged with a train of 6-ms laser pulses (3.5–4 mW under the objective, 30 times at 1 Hz) near a spine of interest. Experiments were performed at 32°C. Images were analyzed with MATLAB (MathWorks) and ImageJ. Other details are described in the Supplemental Experimental Procedures.

### Statistical Analysis

All statistical values were presented as mean  $\pm$  SEM. The Student's *t* test was used when two independent samples were compared. Dunnett's test was used for multiple comparisons. Statistical analysis was performed with GraphPad Prism 6. Differences between datasets were judged to be significant at *p* < 0.05.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.04.044>.

## AUTHOR CONTRIBUTIONS

T.M. and J.N. conceived the SLENDR idea. T.M., J.N., and R.Y. designed the experiments. T.M. and J.N. performed most experiments and data analysis. Y.S. and N.K. performed electron microscopy imaging and data analysis. T.M., J.N., and R.Y. wrote the paper. All the authors discussed the results and commented on the manuscript.

## ACKNOWLEDGMENTS

The authors thank Sanger Institute for the gift of the piggyBac transposon vector plasmids (pPB-CAG.EBNXN and pCMV-hyPBBase) and Y. Hayashi for CaMKII $\alpha$  and CaMKII $\beta$  cDNA; L. Colgan, M. Yuzaki, S. Soderling, and M. Kano for critical reading of the manuscript; the R.Y. lab members for discussion; J. Richards, K. Liu, and J. Chen for technical assistance; and D. Kloetzer for laboratory management. This work was supported by the NIH (R01MH080047 and DP1NS096787 to R.Y.) and the Human Frontier Science Program (long-term fellowship to T.M.).

Received: February 3, 2016

Revised: March 14, 2016

Accepted: April 13, 2016

Published: May 12, 2016

## REFERENCES

- Borrell, V., Yoshimura, Y., and Callaway, E.M. (2005). Targeted gene delivery to telencephalic inhibitory neurons by directional *in utero* electroporation. *J. Neurosci. Methods* *143*, 151–158.
- Bosch, M., Castro, J., Saneyoshi, T., Matsuno, H., Sur, M., and Hayashi, Y. (2014). Structural and molecular remodeling of dendritic spine substructures during long-term potentiation. *Neuron* *82*, 444–459.
- Chen, F., and LoTurco, J. (2012). A method for stable transgenesis of radial glia lineage in rat neocortex by *piggyBac* mediated transposition. *J. Neurosci. Methods* *207*, 172–180.
- Chen, R.Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat. Genet.* *27*, 327–331.
- Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* *33*, 543–548.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.
- Contractor, A., Klyachko, V.A., and Portera-Cailliau, C. (2015). Altered neuronal and circuit excitability in Fragile X syndrome. *Neuron* *87*, 699–715.
- Cox, D.B., Platt, R.J., and Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. *Nat. Med.* *21*, 121–131.
- Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* *346*, 1258096.
- Fortin, D.A., Tillo, S.E., Yang, G., Rah, J.C., Melander, J.B., Bai, S., Soler-Cedeño, O., Qin, M., Zemelman, B.V., Guo, C., et al. (2014). Live imaging of endogenous PSD-95 using ENABLED: a conditional strategy to fluorescently label endogenous proteins. *J. Neurosci.* *34*, 16698–16712.
- Fukaya, M., and Watanabe, M. (2000). Improved immunohistochemical detection of postsynaptically located PSD-95/SAP90 protein family by protease section pretreatment: a study in the adult mouse brain. *J. Comp. Neurol.* *426*, 572–586.
- Gleeson, J.G., Lin, P.T., Flanagan, L.A., and Walsh, C.A. (1999). Doublecortin is a microtubule-associated protein and is expressed widely by migrating neurons. *Neuron* *23*, 257–271.
- Gomez-Ospina, N., Tsuruta, F., Barreto-Chang, O., Hu, L., and Dolmetsch, R. (2006). The C terminus of the L-type voltage-gated calcium channel Ca<sub>v</sub>1.2 encodes a transcription factor. *Cell* *127*, 591–606.
- Gross, G.G., Junge, J.A., Mora, R.J., Kwon, H.B., Olson, C.A., Takahashi, T.T., Liman, E.R., Ellis-Davies, G.C., McGee, A.W., Sabatini, B.L., et al. (2013). Recombinant probes for visualizing endogenous synaptic proteins in living neurons. *Neuron* *78*, 971–985.
- Harris, K.M., Jensen, F.E., and Tsao, B. (1992). Three-dimensional structure of dendritic spines and synapses in rat hippocampus (CA1) at postnatal day 15 and adult ages: implications for the maturation of synaptic physiology and long-term potentiation. *J. Neurosci.* *12*, 2685–2705.
- Heidenreich, M., and Zhang, F. (2016). Applications of CRISPR-Cas systems in neuroscience. *Nat. Rev. Neurosci.* *17*, 36–44.
- Hell, J.W., Westenbroek, R.E., Warner, C., Ahljianian, M.K., Prystay, W., Gilbert, M.M., Snutch, T.P., and Catterall, W.A. (1993). Identification and differential subcellular localization of the neuronal class C and class D L-type calcium channel alpha 1 subunits. *J. Cell Biol.* *123*, 949–962.
- Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* *157*, 1262–1278.
- Hutagalung, A.H., and Novick, P.J. (2011). Role of Rab GTPases in membrane traffic and cell physiology. *Physiol. Rev.* *91*, 119–149.
- Imamura, F., and Greer, C.A. (2015). Segregated labeling of olfactory bulb projection neurons based on their birthdates. *Eur. J. Neurosci.* *41*, 147–156.
- Incontro, S., Asensio, C.S., Edwards, R.H., and Nicoll, R.A. (2014). Efficient, complete deletion of synaptic proteins using CRISPR. *Neuron* *83*, 1051–1057.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.
- Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., et al. (2015). Saturated reconstruction of a volume of neocortex. *Cell* *162*, 648–661.
- Kim, S., Kim, D., Cho, S.W., Kim, J., and Kim, J.S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* *24*, 1012–1019.
- Kim, K., Lakhanpal, G., Lu, H.E., Khan, M., Suzuki, A., Hayashi, M.K., Narayanan, R., Luyben, T.T., Matsuda, T., Nagai, T., et al. (2015). A temporary gating of actin remodeling during synaptic plasticity consists of the interplay between the kinase and structural functions of CaMKII. *Neuron* *87*, 813–826.
- Kitazawa, A., Kubo, K., Hayashi, K., Matsunaga, Y., Ishii, K., and Nakajima, K. (2014). Hippocampal pyramidal neurons switch from a multipolar migration mode to a novel “climbing” migration mode during development. *J. Neurosci.* *34*, 1115–1126.
- Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* *529*, 490–495.
- Korb, E., Wilkinson, C.L., Delgado, R.N., Lovero, K.L., and Finkbeiner, S. (2013). Arc in the nucleus regulates PML-dependent GluA1 transcription and homeostatic plasticity. *Nat. Neurosci.* *16*, 874–883.
- Lee, S.J., Escobedo-Lozoya, Y., Szatmari, E.M., and Yasuda, R. (2009). Activation of CaMKII in single dendritic spines during long-term potentiation. *Nature* *458*, 299–304.
- Lin, S., Staahl, B.T., Alla, R.K., and Doudna, J.A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife* *3*, e04766.
- Lisman, J., Yasuda, R., and Raghavachari, S. (2012). Mechanisms of CaMKII action in long-term potentiation. *Nat. Rev. Neurosci.* *13*, 169–182.
- Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., and Ploegh, H.L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* *33*, 538–542.
- Nishiyama, J., and Yasuda, R. (2015). Biochemical computation for spine structural plasticity. *Neuron* *87*, 63–75.
- Nishiyama, J., Hayashi, Y., Nomura, T., Miura, E., Kakegawa, W., and Yuzaki, M. (2012). Selective and regulated gene expression in murine Purkinje cells by *in utero* electroporation. *Eur. J. Neurosci.* *36*, 2867–2876.
- Parajuli, L.K., Nakajima, C., Kulik, A., Matsui, K., Schneider, T., Shigemoto, R., and Fukazawa, Y. (2012). Quantitative regional and ultrastructural localization of the Ca<sub>v</sub>2.3 subunit of R-type calcium channel in mouse brain. *J. Neurosci.* *32*, 13555–13567.

- Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., et al. (2014). CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308.
- Riedl, J., Crevenna, A.H., Kessenbrock, K., Yu, J.H., Neukirchen, D., Bista, M., Bradke, F., Jenne, D., Holak, T.A., Werb, Z., et al. (2008). Lifeact: a versatile marker to visualize F-actin. *Nat. Methods* **5**, 605–607.
- Saleh-Gohari, N., and Helleday, T. (2004). Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res.* **32**, 3683–3688.
- Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355.
- Schalek, R., Kasthuri, N., Hayworth, K., Berger, D., Tapia, J., Morgan, J., Turaga, S., Fagerholm, E., Seung, H., and Lichtman, J. (2011). Development of high-throughput, high-resolution 3D reconstruction of large-volume biological tissue using automated tape collection ultramicrotomy and scanning electron microscopy. *Microsc. Microanal.* **17**, 966–967.
- Shepherd, J.D., and Bear, M.F. (2011). New views of Arc, a master regulator of synaptic plasticity. *Nat. Neurosci.* **14**, 279–284.
- Slymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88.
- Soma, M., Aizawa, H., Ito, Y., Maekawa, M., Osumi, N., Nakahira, E., Okamoto, H., Tanaka, K., and Yuasa, S. (2009). Development of the mouse amygdala as revealed by enhanced green fluorescent protein gene transfer by means of *in utero* electroporation. *J. Comp. Neurol.* **513**, 113–128.
- Steinberg, S.F. (2008). Structural basis of protein kinase C isoform function. *Physiol. Rev.* **88**, 1341–1378.
- Stoppini, L., Buchs, P.A., and Muller, D. (1991). A simple method for organotypic cultures of nervous tissue. *J. Neurosci. Methods* **37**, 173–182.
- Straub, C., Granger, A.J., Saulnier, J.L., and Sabatini, B.L. (2014). CRISPR/Cas9-mediated gene knock-down in post-mitotic neurons. *PLoS ONE* **9**, e105584.
- Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M., and Zhang, F. (2015). *In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* **33**, 102–106.
- Tabata, H., and Nakajima, K. (2001). Efficient *in utero* gene transfer system to the developing mouse brain using electroporation: visualization of neuronal migration in the developing cortex. *Neuroscience* **103**, 865–872.
- Toyo-oka, K., Wachi, T., Hunt, R.F., Baraban, S.C., Taya, S., Ramshaw, H., Kaibuchi, K., Schwarz, Q.P., Lopez, A.F., and Wynshaw-Boris, A. (2014). 14-3-3 $\epsilon$  and  $\zeta$  regulate neurogenesis and differentiation of neuronal progenitor cells in the developing brain. *J. Neurosci.* **34**, 12168–12181.
- Viswanathan, S., Williams, M.E., Bloss, E.B., Stasevich, T.J., Speer, C.M., Nern, A., Pfeiffer, B.D., Hooks, B.M., Li, W.P., English, B.P., et al. (2015). High-performance probes for light and electron microscopy. *Nat. Methods* **12**, 568–576.
- Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918.
- Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G., et al. (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature* **514**, 380–384.
- Yamamoto, N., Kurotani, T., and Toyama, K. (1989). Neural connections between the lateral geniculate nucleus and visual cortex *in vitro*. *Science* **245**, 192–194.
- Yang, J., Siao, C.J., Nagappan, G., Marinic, T., Jing, D., McGrath, K., Chen, Z.Y., Mark, W., Tessarollo, L., Lee, F.S., et al. (2009). Neuronal release of proBDNF. *Nat. Neurosci.* **12**, 113–115.
- Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370–1379.
- Yin, H., Xue, W., Chen, S., Bogorad, R.L., Benedetti, E., Grompe, M., Kotliansky, V., Sharp, P.A., Jacks, T., and Anderson, D.G. (2014). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* **32**, 551–553.
- Yusa, K., Zhou, L., Li, M.A., Bradley, A., and Craig, N.L. (2011). A hyperactive *piggyBac* transposase for mammalian applications. *Proc. Natl. Acad. Sci. USA* **108**, 1531–1536.
- Yuste, R., and Bonhoeffer, T. (2004). Genesis of dendritic spines: insights from ultrastructural and imaging studies. *Nat. Rev. Neurosci.* **5**, 24–34.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771.
- Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D., and Luo, L. (2005). Mosaic analysis with double markers in mice. *Cell* **121**, 479–492.

# Ultra-High Resolution 3D Imaging of Whole Cells

Fang Huang,<sup>1,2,14</sup> George Sirinakis,<sup>1,3,14</sup> Edward S. Allgeyer,<sup>1,3</sup> Lena K. Schroeder,<sup>1</sup> Whitney C. Duim,<sup>1,4</sup> Emil B. Kromann,<sup>1,5</sup> Thomy Phan,<sup>1</sup> Felix E. Rivera-Molina,<sup>1</sup> Jordan R. Myers,<sup>1</sup> Irnov Irnov,<sup>6,7</sup> Mark Lessard,<sup>8</sup> Yongdeng Zhang,<sup>1</sup> Mary Ann Handel,<sup>8</sup> Christine Jacobs-Wagner,<sup>6,7,9,10</sup> C. Patrick Lusk,<sup>1</sup> James E. Rothman,<sup>1,11</sup> Derek Toomre,<sup>1,11</sup> Martin J. Booth,<sup>12,13</sup> and Joerg Bewersdorf<sup>1,5,11,\*</sup>

<sup>1</sup>Department of Cell Biology, School of Medicine, Yale University, New Haven, CT 06520, USA

<sup>2</sup>Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907, USA

<sup>3</sup>The Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK

<sup>4</sup>Department of Chemistry, Harvey Mudd College, Claremont, CA 91711, USA

<sup>5</sup>Department of Biomedical Engineering, Yale University, CT 06520, USA

<sup>6</sup>Microbial Sciences Institute, Yale University, West Haven, CT 06516, USA

<sup>7</sup>Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520, USA

<sup>8</sup>The Jackson Laboratory, Bar Harbor, ME 04609, USA

<sup>9</sup>Howard Hughes Medical Institute, Yale University, New Haven, CT 06520, USA

<sup>10</sup>Department of Microbial Pathogenesis, Yale School of Medicine, New Haven, CT 06520, USA

<sup>11</sup>Nanobiology Institute, Yale University, West Haven, CT 06516, USA

<sup>12</sup>Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK

<sup>13</sup>Centre for Neural Circuits and Behaviour, University of Oxford, Oxford OX1 3SR, UK

<sup>14</sup>Co-first author

\*Correspondence: joerg.bewersdorf@yale.edu

<http://dx.doi.org/10.1016/j.cell.2016.06.016>

## SUMMARY

**Fluorescence nanoscopy, or super-resolution microscopy, has become an important tool in cell biological research. However, because of its usually inferior resolution in the depth direction (50–80 nm) and rapidly deteriorating resolution in thick samples, its practical biological application has been effectively limited to two dimensions and thin samples. Here, we present the development of whole-cell 4Pi single-molecule switching nanoscopy (W-4PiSMSN), an optical nanoscope that allows imaging of three-dimensional (3D) structures at 10- to 20-nm resolution throughout entire mammalian cells. We demonstrate the wide applicability of W-4PiSMSN across diverse research fields by imaging complex molecular architectures ranging from bacteriophages to nuclear pores, cilia, and synaptonemal complexes in large 3D cellular volumes.**

## INTRODUCTION

Major advances in cell biology are tightly linked to innovations in microscopy. The development of fluorescence microscopy, for example, enabled sub-cellular localization of specifically labeled proteins of interest (Lichtman and Conchello, 2005). However, the wave nature of light restricts the resolution of conventional light microscopy to ~200 nm, making details of subcellular structures and protein assemblies unresolvable (Hell, 2007). The advent of super-resolution fluorescence microscopy, or nanoscopy, techniques such as stimulated emission depletion (STED) (Hell and Wichmann, 1994) and single-molecule switch-

ing nanoscopy (SMSN) (Betzig et al., 2006; Hess et al., 2006; Rust et al., 2006) has extended the application range of fluorescence microscopy beyond the diffraction limit, achieving up to 10-fold improvement in resolution (Gould et al., 2012a). These methods are now maturing and offering the opportunity to observe biological phenomena never before seen (Chojnacki et al., 2012; Kanchanawong et al., 2010; Liu et al., 2011; Xu et al., 2013). Nanoscopy techniques share a common principle: they spatially separate unresolvable fluorescent molecules by independently switching their emission “on” and “off” (Hell, 2007). In particular, SMSN methods such as photoactivated localization microscopy (PALM), fluorescence photoactivation localization microscopy (FPALM), and stochastic optical reconstruction microscopy (STORM) use a stochastic approach where only a small subset of fluorescent molecules is switched on at any particular moment in time while the majority remains in a non-fluorescent “dark” or “off” state (Gould et al., 2012a). Super-resolved images are reconstructed from the positions of thousands to millions of single molecules that have been recorded in thousands of camera frames.

This imaging strategy was initially applied to single-objective microscopes in two dimensions (2D) (Betzig et al., 2006; Hess et al., 2006; Rust et al., 2006) and later extended to three dimensions (3D) (Huang et al., 2008; Juetten et al., 2008; Pavani et al., 2009). While these instruments achieve 20- to 40-nm resolution in the focal plane (lateral, x-y), the resolution in the depth direction (axial, z) is typically limited to only 50–80 nm. The resolution can, however, be further improved by using a dual-objective “4Pi” detection geometry (Bewersdorf et al., 2006).

Using two objectives doubles the detection efficiency (Xu et al., 2012) and thus improves the localization precision ~1.4-fold in all three dimensions. Additionally, employing two objectives in a 4Pi geometry allows the creation of a single-molecule emission interference pattern at the detector leading to an

~7-fold improvement in axial localization precision over single-objective approaches as demonstrated using interferometric PALM (iPALM) (Shtengel et al., 2009) and 4Pi single marker switching nanoscopy (4Pi-SMSN) (Aquino et al., 2011). This improved resolution enabled, for example, the generation of anatomical maps of focal adhesions at ~10-nm axial resolution (Case et al., 2015; Kanchanawong et al., 2010). However, this method was initially restricted to samples of ~250 nm in thickness (Shtengel et al., 2009) and more recently to 700–1,000 nm (Aquino et al., 2011; Brown et al., 2011). As the typical thickness of a mammalian cell is 5–10  $\mu\text{m}$ , this has limited optical microscopy at the ~10-nm isotropic resolution scale to thin sub-volumes of cells, thus precluding the ability to image organelles that can extend over several microns throughout the whole cell.

Here, we present a new implementation of iPALM/4Pi-SMSN, termed whole-cell 4Pi single-molecule switching nanoscopy (W-4PiSMSN), which extends the imaging capabilities of this technology to whole cells without compromising resolution. W-4PiSMSN allows volumetric reconstruction with 10- to 20-nm isotropic resolution of ~10- $\mu\text{m}$ -thick samples, a 10- to 40-fold improvement in sample thickness over previous iPALM/4Pi-SMSN implementations (Aquino et al., 2011; Brown et al., 2011; Van Engelenburg et al., 2014; Shtengel et al., 2009). Our approach permits ultra-high resolution 3D imaging of virtually any subcellular structure. To demonstrate this, we image the endoplasmic reticulum (ER), bacteriophages, mitochondria, nuclear pore complexes, primary cilia, Golgi-apparatus-associated COPI vesicles, and mouse spermatocyte synaptonemal complexes. By these examples, we show that W-4PiSMSN opens the door to address cell biological questions that were previously unanswerable.

## RESULTS

### Development of W-4PiSMSN

To realize a system that achieves 10- to 20-nm 3D resolution across the thickness of entire mammalian cells, we expanded on previous iPALM and 4Pi-SMSN developments (Aquino et al., 2011; Shtengel et al., 2009). In these systems, fluorescence emission is collected by two opposing objective lenses and combined to interfere (Aquino et al., 2011; von Middendorff et al., 2008; Shtengel et al., 2009). Depending on the axial position of a molecule, the light will interfere constructively or destructively, as indicated by the brightness of the molecule's image on the detector. However, molecules at axial positions that differ by multiples of half the wavelength of light lead to the same interference pattern and cause ambiguity in determining their axial positions. This localization ambiguity leads to scrambled images that contain axially shifted image artifacts, known as ghost images, in samples thicker than ~250 nm. This can be avoided by using not only the brightness but also the z-position-dependent shape of the single-molecule images to determine a molecule's axial position. To address this, a higher-moment based analysis (Aquino et al., 2011) and analysis of the point-spread function (PSF) eccentricity in a hyperbolic mirrors-modified system (Brown et al., 2011) were developed, which extended the image volume thickness to 700–1,000 nm. However, these methods pose significant drawbacks such as

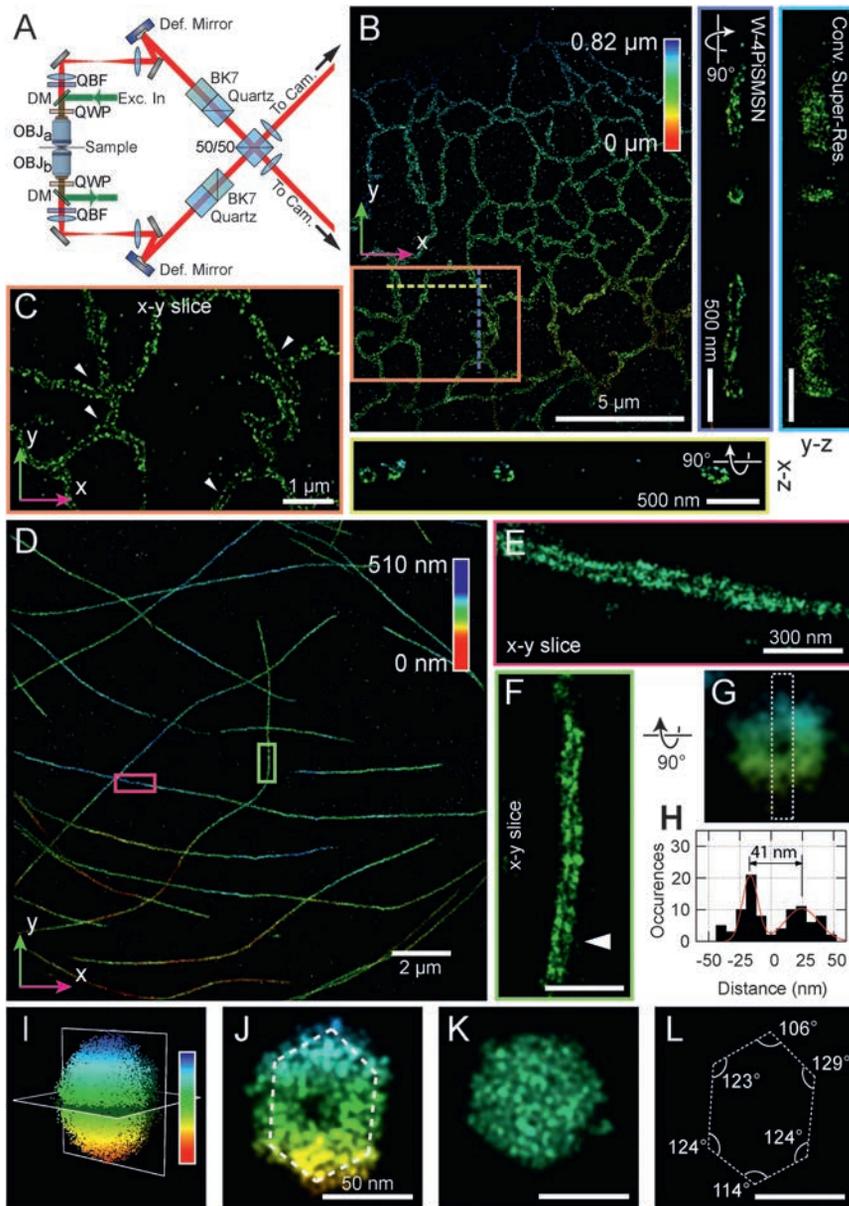
poor localization density because of the highly selective computational processes focusing on subtle features of PSFs. The methods are also susceptible to sample-induced optical aberrations, which change the PSF shape when imaging biological structures deeper in the sample (Burke et al., 2015; von Diezmann et al., 2015). As a result, applications have been restricted to thin structures close to the coverslip (Case et al., 2015; Van Engelenburg et al., 2014; Kanchanawong et al., 2010).

To enable 4Pi-SMSN to probe deeper into the cell and extend the application of this technology to larger cellular features, we have developed W-4PiSMSN. First, expanding on the optical design by Aquino et al. (2011), we included deformable mirrors in both arms of the 4Pi-interferometric cavity (Figure 1A; Figure S1). We use these mirrors to correct for imperfections in the instrument beam path and optimize the PSF quality for samples with various thicknesses (Supplemental Information; Figure S2). Deformable mirrors also allow us to compensate for sample-induced aberration modes, such as spherical aberrations (Burke et al., 2015; Gould et al., 2012b), which vary from sample to sample and with depth. Additionally, we can use these mirrors to introduce astigmatism in both interference arms without adding further complexity to the system (Supplemental Information). Thus, the deformable mirrors enable compromise-free, reproducible PSFs in a depth- and sample-independent manner.

Second, building on an earlier approach by Brown et al. (2011), we developed an analysis method that combines information from (1) the 4Pi-PSF's interference phase, which allows for precise axial localization but does not distinguish between different interference peaks, and (2) the eccentricity of the astigmatic 4Pi-PSF, which narrows axial localizations down to individual interference peaks but in itself does not offer the precision of 4Pi interference. Our new analysis algorithm interprets the large number of molecules imaged in each time and z-depth window as an *ensemble measurement* of the concurrent state of the W-4PiSMSN system (Supplemental Information; Figure S3A) and determines the relationship between the eccentricity of the astigmatic PSF and the interference phase of the 4Pi-PSF. Then the axial positions of all corresponding molecules can be assigned with high precision and unambiguously using a monotonic metric designed to describe the overall shape of the PSF and maintain its monotonicity in the presence of moderate amounts of aberrations (Supplemental Information; Figure S3B). Since this analysis is performed for well-defined temporal and axial data subsets, we have generalized it to identify and correct for drift (from both the system and the sample) over the course of imaging. Our method is robust against aberrations and improves the reliability and efficiency of axial position assignment as it automatically adapts to changes in the shape and interference pattern of the 4Pi-PSF.

### Ultra-High Resolution Imaging with W-4PiSMSN

To demonstrate the resolution capabilities of our new system, we first imaged the ER. ER membranes were stained using anti-GFP antibodies to the overexpressed transmembrane protein, mEmerald-Sec61 $\beta$ , in COS-7 cells. We visualized the ER as a connected network of hollow tubes with 60–100 nm diameters (Figures 1B and 1C; Movie S1). Both horizontal and vertical



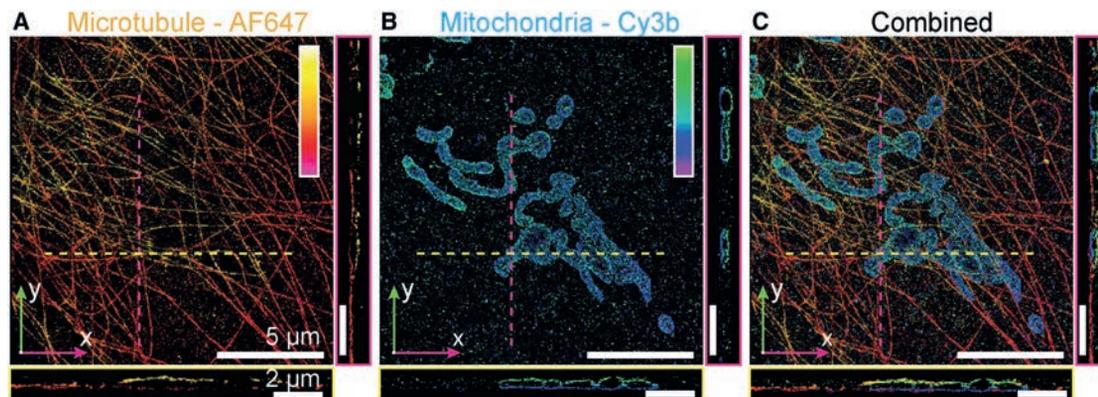
**Figure 1. W-4PiSMSN Design and Resolution Demonstrations with ER, Microtubules, and Bacteriophages**

(A) Simplified optical diagram of W-4PiSMSN. (B) Overview and cross-sections of the ER network in an immunolabeled COS-7 cell. Cross-sections of the W-4PiSMSN reconstruction show clearly separated membranes of the tubular structures, which cannot be resolved with conventional astigmatism-based nanoscopy (light blue frame). (C) x-y slice through the mid-section of the ER network shown in (B) highlights the distinct membrane contour of ER tubules (arrowheads). (D) Overview of immunolabeled microtubules in a COS-7 cell. (E and F) 20-nm-thin x-y slices of the red (E) and green (F) segments shown in (D) demonstrate that microtubules can be easily resolved as hollow cylinders in W-4PiSMSN. (G) A look through a 120-nm-long segment of the microtubule of (F). (H) A histogram showing the number of localizations in a cross-section of the microtubule, white dotted box in (G). (I) A bacteriophage reconstructed from 115 averaged viral particles rendered in 3D. (J and K) 5-nm-thin vertical (J) and horizontal (K) slices through the averaged dataset corresponding to the planes shown in (I). (L) The internal angle measurements of the hexagon shape identified from the viral capsid shown in (J). OBJ, objective; QWP, quarter-wave plate; DM, dichroic mirror; QBF, quad-band band-pass filter; Def. Mirror, deformable mirror; Cam, camera; 50/50, beam splitter cube.

cross-sections reveal the 3D membrane contour that was previously resolvable only with electron tomography (Frey and Manella, 2000). This high 3D resolution is quantitatively supported by a Fourier Shell Correlation value of 22 nm (Figure S4) (Nieuwenhuizen et al., 2013). To test our approach on even smaller structures, we imaged antibody-labeled microtubules in COS-7 cells, a gold standard in SMSN (Figures 1D–1H). Without any detectable imaging artifacts, W-4PiSMSN resolves this 25-nm microtubule filament, which appears in all orientations as a hollow core coated with antibodies (Figures 1E–1H). In addition, the dataset features a high localization density of  $\sim 5.5$  localization events per  $10 \times 10 \text{ nm}^2$  of surface area. Displaying the localization events by their radial distance from the tubule axis shows a Gaussian peak with a full width at half maximum (FWHM) of 16–24 nm (Figure S4). Considering that the use of pri-

mary and secondary antibodies adds uncertainty to the actual position of the imaged dye molecules, we conclude that the 3D resolution of our instrument is well below 20 nm (FWHM).

To demonstrate our approach on another challenging target, we imaged T7 bacteriophages. They feature an icosahedral-shaped capsid of  $\sim 60$ -nm diameter, which has only been visible by cryo-electron microscopy (cryo-EM) techniques before (Hu et al., 2013). We non-specifically labeled proteins on the surface of purified T7 phages using an Alexa Fluor 647 NHS Ester, which reacts with primary amines, and mounted the phages on a coverslip (Figure S5). Image slices of a single phage in the x-y, y-z, and x-z directions show a hollow center in all orientations. To further refine the details of the detected phage structures, we adapted the tomogram-averaging approach originally developed for cryo-EM (Briggs, 2013; Broeken et al., 2015). By combining 115 T7 images, our averaged reconstruction reveals the icosahedral shape of the T7 phages (Figures 1I–1L; Figure S4). As presented in Figures 1J–1L, a slice perpendicular to the major axis shows the expected pentagonal shape while a slice parallel to the major axis reveals a hexagonal shape. Our approach, however, has yet to clearly resolve the



**Figure 2. Two-Color Reconstruction of Mitochondria and Microtubules**

(A and B) W-4PiSMSN reconstruction of microtubules (A) and mitochondria (TOM20) (B) in a COS-7 cell immunolabeled with Alexa Fluor 647 and Cy3B, respectively. An x-y overview and x-z and y-z slices (yellow and magenta lines, respectively) are shown. (C) The combined two-color image reveals microtubules running adjacent to the mitochondria surface.

~23-nm tail and fiber structures of the T7 phage (Hu et al., 2013). This is likely due to either incomplete labeling of the surface proteins or the flexibility of these structures. Nonetheless, our W-4PiSMSN system has enabled the visualization of the ultrastructure of bacteriophages using light microscopy.

We tested the two-color imaging capability of W-4PiSMSN by imaging microtubules and mitochondria in a COS-7 cell immunolabeled with Alexa Fluor 647 and Cy3B, respectively. Our reconstructions show microtubules running in close proximity ~10–20 nm from the mitochondria top and bottom surfaces (Figures 2A–2C; Movie S2; Figure S5). Our system decouples axial localization from the PSF shape, the latter being susceptible to depth-dependent distortions caused by sample-induced optical aberrations (Liu et al., 2013; McGorty et al., 2014). While single-objective systems rely on the PSF shape, the W-4PiSMSN approach uses the relative interference amplitudes to determine the axial location of individual molecules. However, multicolor imaging is challenging because the spatial interference modulation frequency is wavelength dependent and differs between color channels. To address this, we derived the modulation frequency using a pupil-function based approach (Supplemental Information). Our simulation results were verified experimentally by registering two color channels from an affine transformation matrix, which was calibrated using two-color-labeled mitochondria in fixed cells (Supplemental Information).

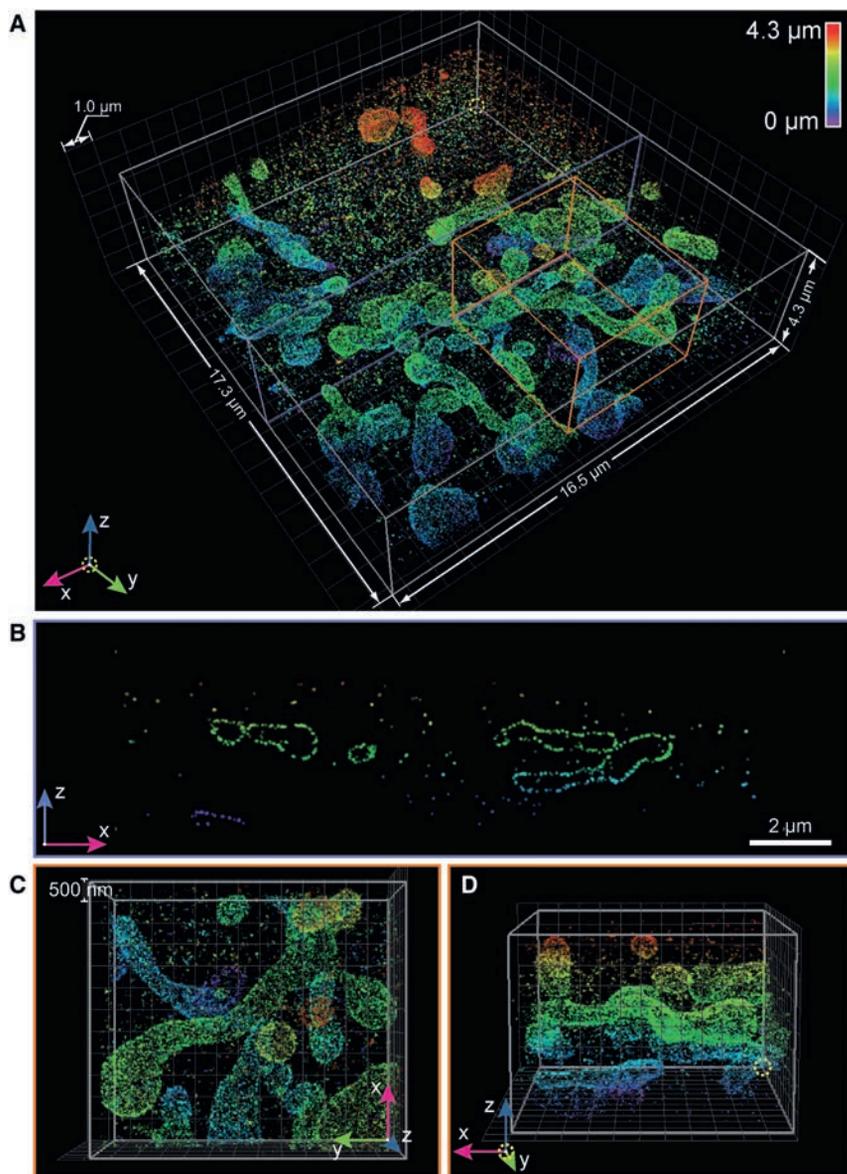
### Whole-Cell 3D Imaging with W-4PiSMSN

Imaging volumes thicker than ~1.2 μm requires axial sample scanning, because molecules more than ~600 nm out of focus cannot be identified and localized efficiently (Huang et al., 2008; Juetten et al., 2008). Thus, optical sections must be recorded at different axial sample positions and subsequently merged to obtain the complete cellular volume. Compared to conventional 3D nanoscopes, the superior localization precision of W-4PiSMSN puts high demands on the localization accuracy in each volume section (i.e., avoiding volume distortions) and the merging process. In the section-merging process, small misalignments of neighboring optical sections caused by sam-

ple-induced aberrations or drift can lead to significant deterioration of the resolution and distortions of the super-resolved volume (Huang et al., 2008; Mlodzianowski et al., 2011).

We designed our system to minimize drift: our instrument design takes advantage of a horizontal symmetry plane coinciding with the common focal planes of the objectives and the beam splitter cube of the interference cavity. This symmetric design desensitizes the interferometric cavity of the microscope to temperature changes leading to approximately equal thermal expansion in both arms of the interference cavity. To compensate for any remaining instrument and sample drift caused by mechanical and thermal fluctuations, we developed a set of hardware and software tools (Figure S6). The objectives are stabilized in 3D relative to each other by focusing a near-infrared laser beam by one objective and detecting the focus with the other objective in a “biplane” configuration (Figure S6) (Juetten et al., 2008; Ram et al., 2008). This allows the detection of relative objective movement in 3D, which can then be compensated for via a feedback loop. Furthermore, we cross-correlate 3D volume data segments of 1- to 2-min windows using a redundancy-based drift correction method (Li et al., 2013; Wang et al., 2014) in an extended correlation volume. Within each of these short data segments, an independent relationship between astigmatism and interference phase is established. Any discrepancies between these relationships for different segments are treated as drift (Supplemental Information). The above-described methods enable us to fully compensate for sample and instrument drift and changes in the optical path between the two arms of the interferometric 4Pi cavity due to the axial scanning nature of the measurements.

To demonstrate the whole-cell imaging capabilities of the W-4PiSMSN system, we imaged mitochondria using antibodies against the outer membrane protein TOM20 over the whole thickness of a COS-7 cell. Figure 3 reveals the outer membrane contour and the remarkably interconnected mitochondrial network over a depth of 4.3 μm (Figures 3B–3D; Movie S3). We were not able to detect any significant ghost images within the volume (Figures 3A–3D).



**Figure 3. W-4PiSMSN Reconstruction of TOM20 on Mitochondria in COS-7 Cell**

(A) Overview of the mitochondria network visualized by immunolabeling TOM20 with Alexa Fluor 647. The dataset is assembled from 11 optical sections with 500-nm step sizes.

(B) x-z cross-section of the purple plane in (A) showing the distribution of TOM20 on the outer mitochondrial membrane. Ghost images are completely negligible.

(C and D) Top (C) and side (D) views of the orange box in (A) show the 3D arrangement of the organelle.

Moreover, as the Golgi complex is located close to the middle of the cell, recording high-quality data in a central z-plane is a challenging test of the instrument's 3D resolution capabilities. Figure 5 shows the  $\beta'$  COP, a protein in the outer COPI complex, immunostained using Alexa Fluor 647 in BSC-1 cells. Strikingly, we visualized distinct hollow COPI-coated spheres within cells (Figures 5B, 5C, 5E, 5F, and 5G–5N). Our 3D images resolve individual COPI vesicles with  $\sim 100$ -nm diameter, consistent with previous measurements (Pellett et al., 2013). Additionally, a 300-nm-thick section shows that COPI-coated structures are packed around a 500- to 1,000-nm (x and y) by 500-nm (z) area devoid of COPI labeling, presumably containing a Golgi stack (Figures 5D–5F).

### Revealing Ciliary Membrane GPCR Organization

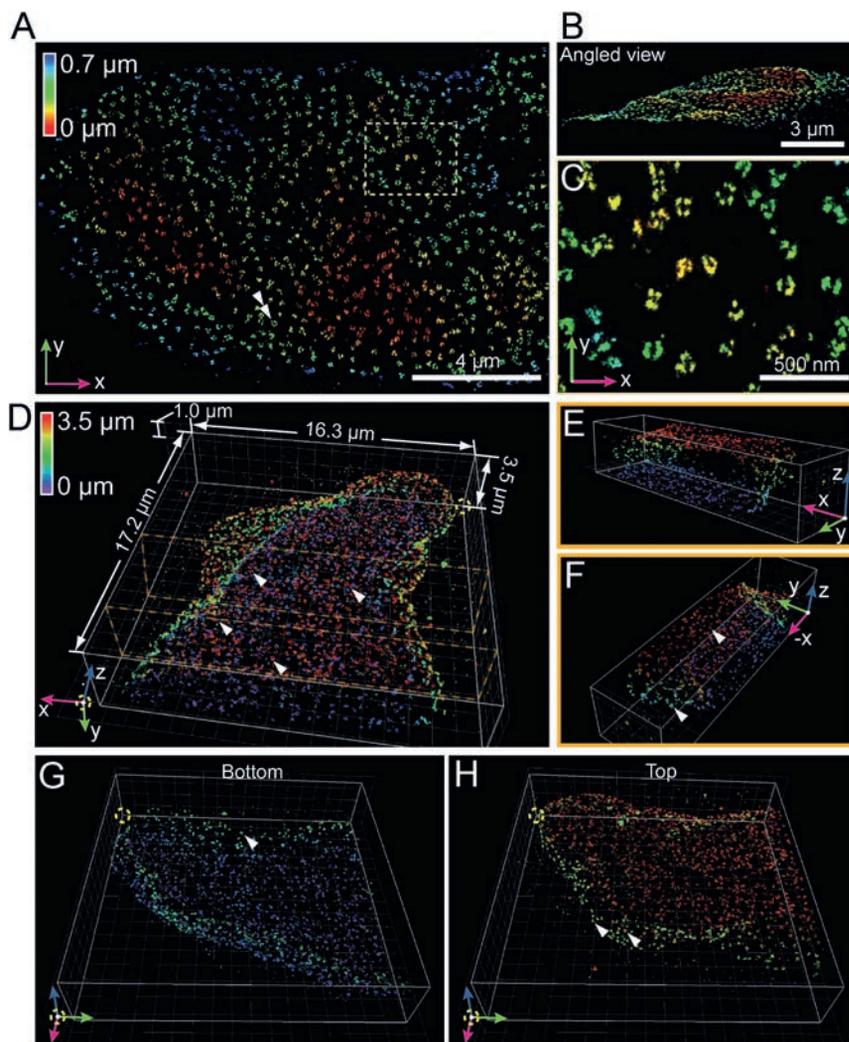
Most high-resolution studies of the primary cilium, a solitary microtubule-based organelle that protrudes from the cell surface and acts as a cellular antenna, have relied on EM (Wood and Rosenbaum,

2015). A transmission EM image typically shows only a small subsection of a cilium as the sample is a random oblique  $\sim 70$ - to 100-nm-thick section through the structure, which can be up to 10  $\mu\text{m}$  long and  $\sim 250$  nm wide. Scanning EM images can easily show an entire cilium with high resolution; however, these images completely lack information about specific protein localization. Previous nanoscopy studies on cilia relied on inferring the 3D organization from 2D datasets (Yang et al., 2013, 2015). Here, we used W-4PiSMSN to image the G-protein-coupled receptor Smoothed (SMO) on whole primary cilia in hTERT-RPE1 cells with high 3D resolution (Figure 6; Movie S5). SMO was tagged with a pH-sensitive GFP (pH-SMO), which was used as an epitope for antibody labeling with Alexa Fluor 647 (Figure 6; Supplemental Information). We observe that over-expressed pH-SMO localizes to the membranes of cilia, which form hollow cylinders 3–10  $\mu\text{m}$  long (Figure 6) and vary in

To further demonstrate that image quality is maintained throughout the thickness of whole cells, we imaged nuclear pore complexes (NPCs) on the nuclear envelope. By immunolabeling with an antibody that recognizes a component of the cytoplasmic filaments (Nup358) of NPCs (von Appen et al., 2015), we can reconstruct NPCs on the top, side, and bottom of the nucleus (Figure 4; Movie S4). As with mitochondria, our approach reveals the contours of almost the entire nuclear surface, where both prominent invaginations and subtle undulations (typically visualized only by electron microscopy [EM]) are apparent (Figures 4A and 4B).

### Revealing Golgi-Apparatus-Associated COPI Vesicles

We next imaged COPI vesicles, which have traditionally been resolved only by EM as they have  $\sim 100$ -nm diameters and are densely packed around the Golgi cisternae (Orci et al., 1997).



**Figure 4. W-4PiSMSN Imaging of Nuclear Pore Complexes over the Thickness of a Cell Nucleus**

Nup358 was immunolabeled with Alexa Fluor 647 in hTERT-RPE1 cells.

(A) Overview of a region of the nucleus. The axial location of the nuclear pores is color coded.

(B) Side view of (A).

(C) A subregion indicated by the dashed box in (A) shows a zoomed in view of multiple nuclear pores.

(D) Overview of a 3D reconstruction of the nucleus obtained by combining nine optical sections.

(E) A section of the reconstruction in (D) confirms that the labeling is largely limited to the nuclear envelope.

(F) Different view of the same section.

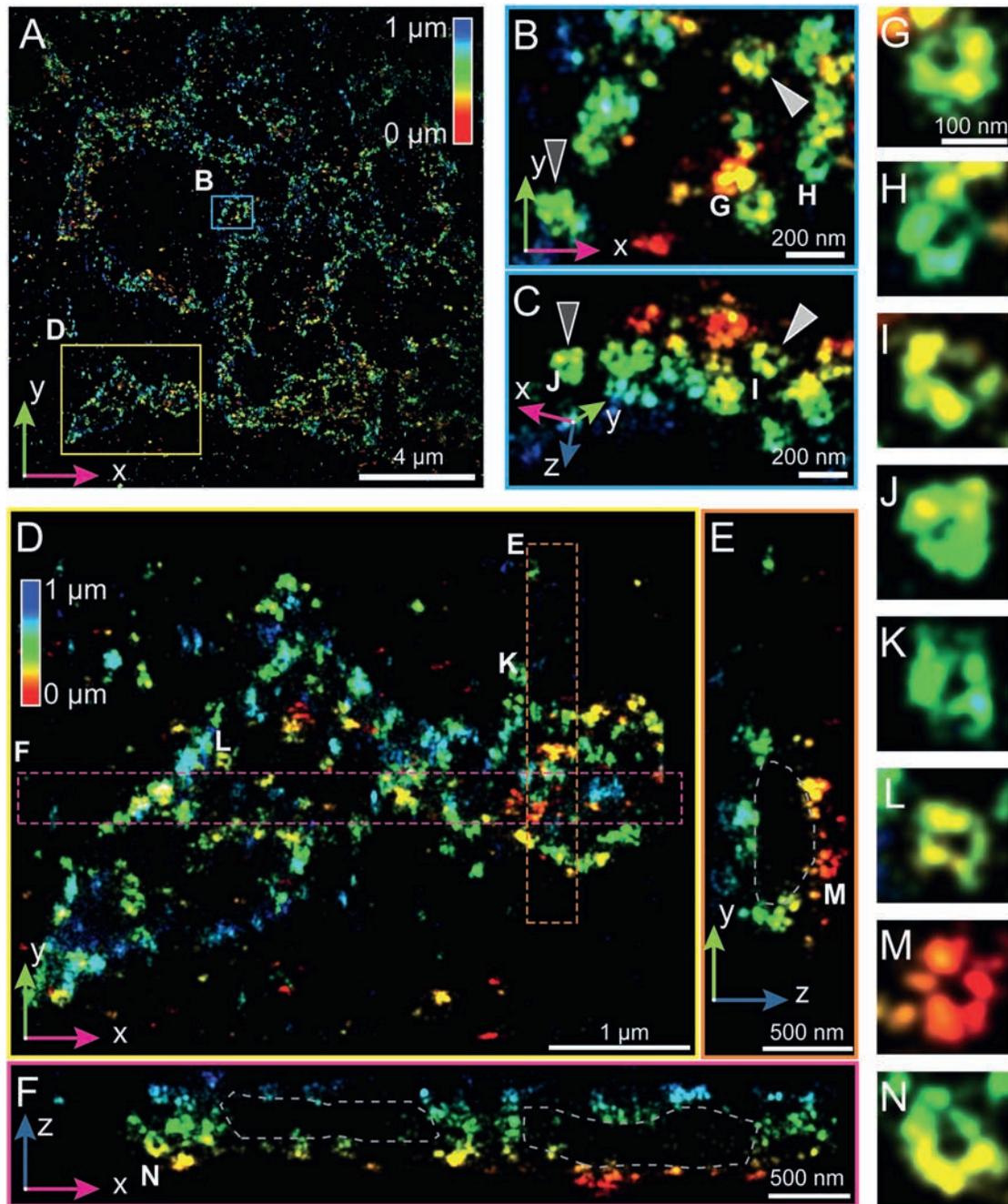
(G and H) Bottom (G) and top (H) half of the nucleus shown in (D). The images reveal ring-like nuclear pores on the top and the bottom nuclear envelope as well as at the sides of the nucleus (arrowheads).

sions may be vesicles (diameter  $\sim$ 150–200 nm) budding from the cilium (Figures 6F–6H), as ectosomes have been reported to bud from some cilia (Wood and Rosenbaum, 2015).

### Resolving Synaptonemal Complexes in Whole-Mouse Spermatocytes

As a final demonstration of the capacity of our instrument to image deep into cells as thick as 10  $\mu$ m, we stained synaptonemal complexes in mouse spermatocyte nuclei in the pachytene phase of meiotic prophase (Figure 7; Movies S6). While synaptonemal complexes have been imaged using structured illumination (Carlton, 2008; Qiao et al., 2012) and 4Pi microscopy at 100- to 200-nm resolution (Fritsche et al., 2012), higher resolution optical images have been limited to chromosome spreads of  $<$ 1- $\mu$ m thickness (Schücker et al., 2015). Here, we examined spermatocytes harvested from testes of 17- to 18-day-old mice with W-4PiSMSN and imaged the twisting band of the paired lateral elements of autosomal synaptonemal complexes, highlighted by immunolabeling synaptonemal complex protein 3 (SYCP3), a constituent component of the lateral elements (Page and Hawley, 2004). Reconstructed from a total of 126 optical sections (21 depth positions imaged in six repetition cycles), the entire 3D image spanned nearly 9  $\mu$ m in depth and resolved SYCP3 substructure of the individual autosomal synaptonemal complexes with unprecedented clarity independent of their orientation or depth (Figures 7A–7E; Movie S6). Furthermore, 19 synaptonemal complexes representing pairs of individual autosomal homologs could be isolated using a Euclidian metric-based clustering algorithm on the individual single-molecule localizations (Supplemental Information). Thus, our approach promises the capacity to visualize the nanoscale spatial organization of chromosomal scaffolds in the context of

diameter from  $\sim$ 160 to 280 nm (Figures 6A–6E). Our W-4PiSMSN images of the ciliary membrane allow us to precisely measure the cilium's diameter along its entire length. Interestingly, we find that cilia diameters are not always constant. Rather, one example cilium shows an abrupt contraction of  $\sim$ 50 nm midway along its length (Figures 6C–6E; Movie S5). We speculate this change in diameter may correspond to the thinning of the 9+0 microtubule axoneme, which is known to transition from triplet microtubules, to doublets, and eventually singlets. The ciliary tip is not narrow but has a bulbous shape, consistent with structures observed in EM (He et al., 2014; Wang et al., 2013). The high-resolution reconstruction of the ciliary membrane also allowed us to “unwrap” the membrane tube into a flat surface (Figure 6H; Supplemental Information) permitting data quantification such as cluster analysis and co-localization measurement in a complex geometry. Next, we examined the local density of molecules along the ciliary membrane to identify regions with higher concentrations of pH-SMO. Higher local density is present around the base, on small bulbous protrusions, and on stripes along the cilium length (Figures 6F–6H; Figure S7). These protrusions



#### Figure 5. W-4PiSMSN Resolves Individual COPI-Coated Vesicles

COPI complexes were immunolabeled with an antibody against  $\beta'$  COP and imaged with Alexa Fluor 647 in BSC-1 cells.

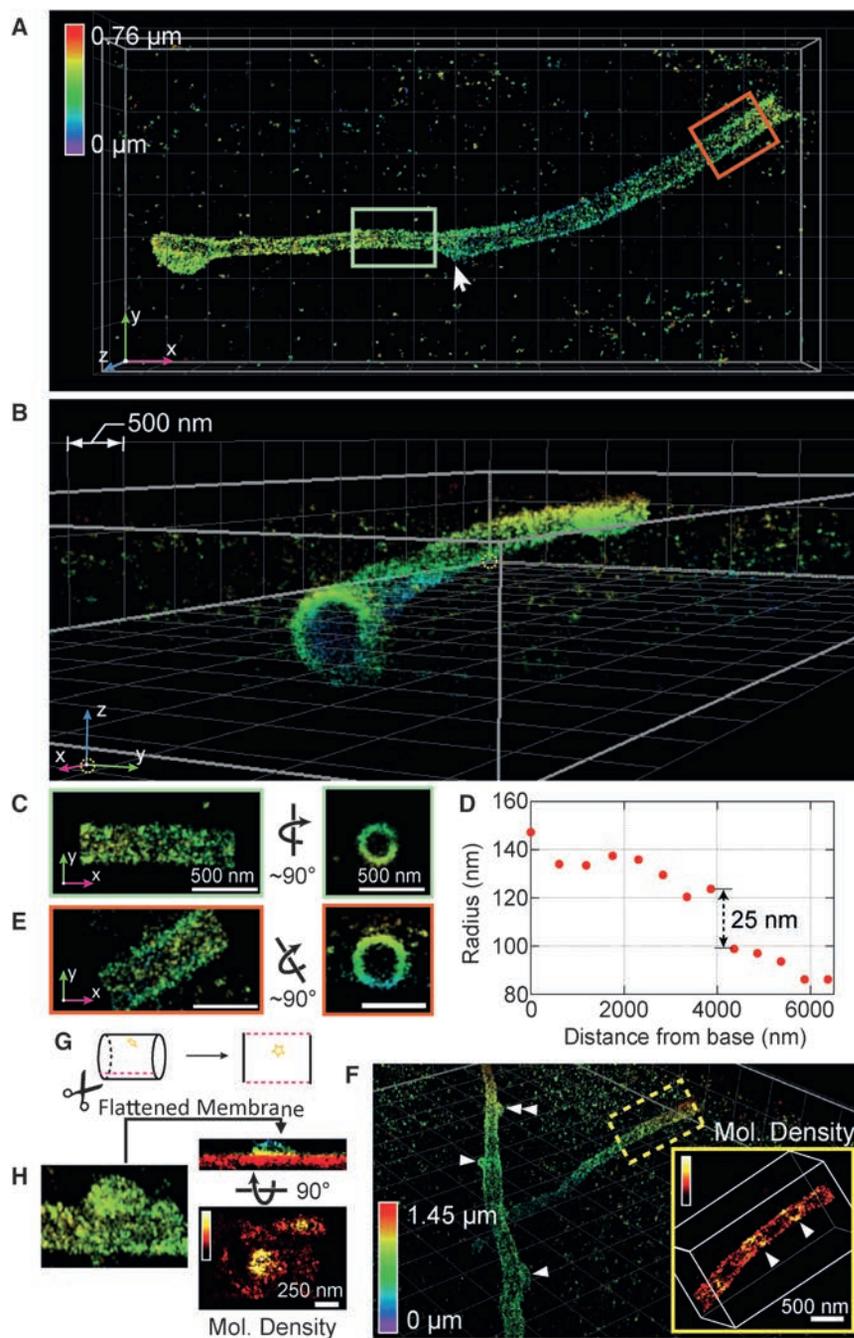
(A) Overview of a region of the field of view, with axial location of molecules color coded.

(B and C) Top (B) and side (C) view of the blue-boxed subregion indicated in (A) showing that COPI often forms round and hollow sphere-like structures. Dark gray and light gray arrowheads indicate the same COPI structures.

(D) x-y view of the area devoid of COPI as indicated by the yellow box in (A).

(E and F) x-z and y-z view of the orange (E) and magenta (F) boxed regions shown in (D) show that COPI surrounds an area presumably containing the Golgi cisternae.

(G–N) COPI vesicle structures as indicated by the respective labels in (B)–(F) shown at the same enlarged scale reveal circular structures.



**Figure 6. GPCR Smoothed on a Primary Cilium**

(A and B) Top (A) and side (B) views of a primary cilium on an hTERT-RPE1 cell expressing pH-SMO, which was immunolabeled with Alexa Fluor 647.

(C and E) Views of sections close to the tip (C) and the base (E) as indicated by the light green and orange boxes in (A) show the localization of pH-SMO to the cilium membrane.

(D) Radius of different sections of the cilium as a function of their distance from the base.

(F) Overview of a cilium in another region of the sample, showing vesicle-like buds on the ciliary membrane surface (arrowheads). The inset shows the local density of the boxed region, which suggests a helical stripe organization of pH-SMO (arrowheads in inset).

(G and H) A bud-like profile shown in (F) can be unwrapped as depicted in (G), showing the height of the vesicle above the cilium membrane and the high molecular density of pH-SMO at the bud (H).

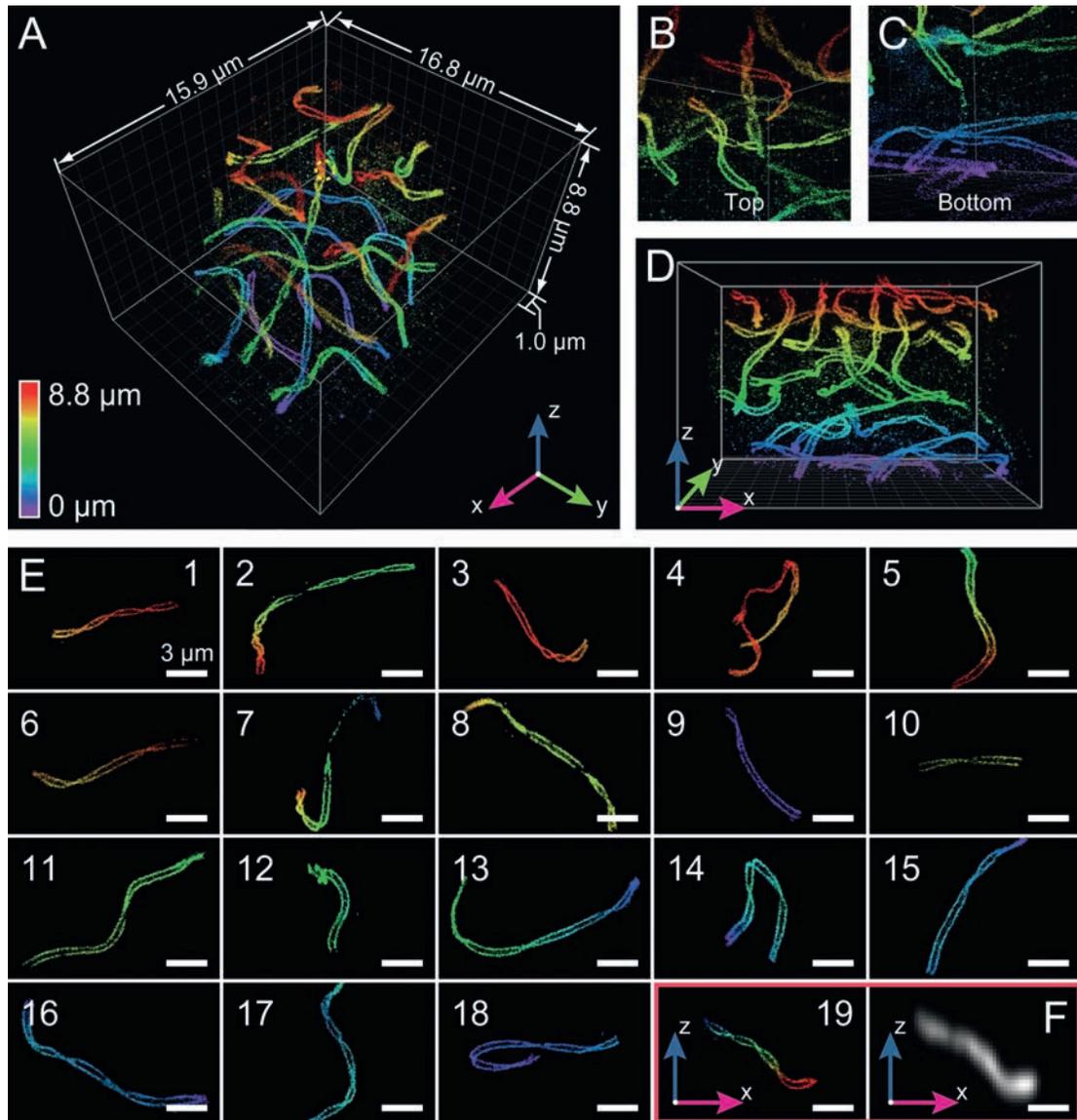
depth improved  $\sim 10$ -fold over state-of-the-art iPALM and 4Pi-SMSN. This development extends the application range of 4Pi-based SMSN dramatically: imaging is no longer limited to features within small sub-volumes of cells. Instead, we are capable of imaging organelles that span large volumes, exemplified by the mitochondrial network, the nuclear envelope, and synaptonemal complexes, which we capture in virtual entirety. Thus, W-4PiSMSN is a versatile and powerful tool that promises a new perspective on how proteins distribute across entire organelles throughout whole cells, a key unmet challenge in cell biology.

Is there room to further increase the spatial resolution of SMS nanoscopy? First, SMS resolution depends on the precision with which one can localize blinking molecules. The precision is approximately proportional to the sharpness of the PSF and, for negligible background noise, is inversely proportional to the square root of the number of detected photons. Our approach has

focused mainly on creating the sharpest PSF and detecting as much of the emitted fluorescence light as possible. Recently, there have been promising developments that increase the number of emitted photons per molecule (Klehs et al., 2014; Ong et al., 2015; Vaughan et al., 2012), which we have not exploited here. Unfortunately, these advances have so far come at the expense of an increase in recording time. We anticipate, however, that with new generations of fluorophores or refined imaging buffers, these approaches can further improve image quality.

## DISCUSSION

Through a confluence of several technological innovations, we have demonstrated that W-4PiSMSN provides unprecedented access to the ultrastructure of cells with  $\sim 10$ - to 20-nm isotropic resolution throughout their entire volume. This resolution is 20–50 times higher than conventional microscopy with imaging



**Figure 7. W-4PiSMSN of the Synaptonemal Complexes in a Whole-Mouse Spermatocyte**

(A) Overview reconstructed from 21 optical sections. Lateral elements of the synaptonemal complex, spaced  $\sim 200$  nm apart, are resolved throughout the  $\sim 9$ - $\mu$ m depth of the spermatocyte at uniform resolution.

(B and C) Different views from locations inside the spermatocyte centered on top (B) and bottom (C) regions of the dataset.

(D) x-z view of (A).

(E) The 19 synaptonemal complexes from an entire mouse spermatocyte haploid genome were computationally isolated using a Euclidian distance-based clustering algorithm.

(F) A conventional image of the 19<sup>th</sup> synaptonemal complex in x-z view.

Scale bars in (E-19) and (F), 2  $\mu$ m.

Second, image quality, or spatial resolution, of SMSN images depends on the density of localized molecules (Patterson et al., 2010; Shroff et al., 2008). The application examples we presented demonstrate that small features such as cylinder-shaped immunolabeled microtubules ( $\sim 40$ -nm diameter) or ER tubules ( $\sim 60$ -nm diameter) can now be easily resolved in 3D using light microscopy. This image interpretation is aided by the fact that the observer fills the gaps in the distribution of molecules along

a tubule by mentally extrapolating from the expected tubular structure. We utilized several computational image processing techniques, including particle averaging of the bacteriophage data (Figure 1), clustering of the synaptonemal complexes (Figure 7; Movie S6), and “unwrapping” of cilia (Figure 6), to reconstruct complex structures. These approaches, which add constraints to data interpretation (e.g., the fact that cilia are tubular) and can be tailored to the cell biological question at

hand, allow us to extract structural details, which are more subtle than the labeling density suggests. Ultimately, labeling density is limited by the density of probe targets, usually proteins. The development of new labeling approaches that allow membrane targeting (Erdmann et al., 2014) overcome this restriction and additionally offer the possibility of revealing the membrane boundaries of individual organelles. Complementary approaches utilizing stochastic transient chemical binding can further address the limited pool of fluorescent labels and theoretically allow unlimited numbers of localized molecules (Giannone et al., 2010; Jungmann et al., 2014; Lew et al., 2011; Sharonov and Hochstrasser, 2006). However, even with the use of conventional labeling techniques, W-4PiSMSN is capable of visualizing otherwise inaccessible structures in a multitude of settings as demonstrated by the large range of presented applications.

In conclusion, we believe that the development of W-4PiSMSN represents the culmination of more than a decade's research on high-resolution fluorescence imaging techniques and establishes 3D biological imaging with molecular specificity and resolution in the 10-nm range as a general imaging technique.

## EXPERIMENTAL PROCEDURES

### Microscope Setup

A detailed description is provided as Supplemental Information. In brief, the 4Pi cavity of the system was set up vertically around two opposing high-NA objective lenses (Movie S7). Excitation light from three laser lines at wavelengths of 642, 561, and 405 nm was coupled into the upper objective for wide-field illumination. Following the concept by Aquino et al. (2011), fluorescence was collected by both objectives and passed through quarter wave plates, which enforced equal fractions of *s* and *p*-polarized light independent of the dipole emitter orientation. A custom Babinet-Soleil compensator corrected for dispersion and allowed adjusting the phase delay between the upper and lower interferometer arm independently for the two polarization components before the light was combined at a 50/50 beam splitter cube. We added deformable mirrors (Boston Micromachines, Multi-5.5) in planes conjugate to the back pupils of the objectives, which allowed for aberration correction, optimization of the PSF and introduction of astigmatism for artifact-free 3D localization. *s* and *p*-polarized fluorescence exiting the beam splitter cube at two sides was split with a polarizing beam splitter cube into four images featuring different interference phases. The four images were recorded simultaneously by a scientific complementary metal-oxide semiconductor (sCMOS) camera (Hamamatsu, ORCA-Flash 4.0v2). The entire system was controlled by a custom-written program in LabVIEW.

### Aberration Correction Using Deformable Mirrors

Two deformable mirrors were independently adjusted as follows. For each interfering arm, starting from the flat voltage map (provided by the manufacturer), 28 theoretically generated membrane modes were applied sequentially with ten different voltage amplitudes per mode. The detected peak signals (0<sup>th</sup> moment Gaussian weighted sum) of a single fluorescent bead in focus were extracted for all amplitudes of the applied modes. The optimal amplitude for each mode was determined as the value that gave the maximum signal level by fitting a quadratic function to the measurements. These newly obtained amplitudes were added to the flat voltage map and serve as a starting point of another iteration. We used five iterations to achieve optimal system aberration correction. Details are provided the Supplemental Information.

### Biological Sample Preparation

A complete description of cell culture, fluorescence labeling, coverslip, and buffer preparation is included in sections 15 to 26 of the Supplemental Information.

In short, 25-mm-diameter coverslips were cleaned by sonication in 1 M KOH for 15 min before use. Fluorescent 100-nm-diameter crimson beads were attached to the coverslip surface using poly-L-lysine. Cultured mammalian cells were grown on coverslips for 24–48 hr before fixation. COS-7 cells were used for microtubule, ER, and mitochondria samples. RPE-hTERT cells were used for nuclear pore complex and cilia samples. BSC1 cells were used for COPI samples. For synaptonemal complex samples, spermatocytes were isolated from the testes of mice and settled on coverslips before being fixed. T7 bacteriophages were isolated from *E. coli* cultures and labeled with Alexa Fluor 647 NHS Ester before being placed on coverslips.

Microtubule, mitochondria, and ER samples were fixed with 3% paraformaldehyde (PFA) + 0.1% glutaraldehyde before antibody labeling. A saponin pre-extraction step preceded fixation when only microtubules were labeled. Nuclear Pore Complex samples were fixed in  $-20^{\circ}\text{C}$  methanol. COPI and synaptonemal samples were fixed in 4% PFA.

Antibodies against endogenous proteins were used to label microtubules (anti- $\alpha$ -tubulin), mitochondria (anti-TOM20), nuclear pore complexes (anti-Nup358), COPI (anti- $\beta'$  COP), and synaptonemal complexes (anti-SYCP3). Overexpressed proteins were labeled with antibodies in ER samples (mEmerald-Sec61 $\beta$  using anti-GFP) and Cilia samples (pHlourin-mSmoothened using anti-GFP).

All one-color samples were labeled with Alexa Fluor 647, either using a commercial secondary antibody or an NHS ester. Two-color microtubule and mitochondria samples were labeled with Alexa Fluor 647 and Cy3B, respectively. Cy3B-labeled secondary antibodies were made by conjugating reactive Cy3B with unlabeled antibodies. After secondary antibody labeling, COS-7 and BSC1 cell samples were post-fixed with 3% PFA + 0.1% GA.

Labeled biological samples were placed in an aluminum sample frame and covered with a second cleaned coverslip. A thin spacer and imaging buffer was placed between the two coverslips. The coverslips were held in the sample frame using an addition-curing silicone. The imaging buffer was either conventional or COT containing thiol buffer (Supplemental Information). Samples were imaged immediately after the silicone solidified.

### W-4PiSMSN Data Acquisition

Four phase images are arranged along the splitting line of the sCMOS camera's upper and lower readout region. 50,000 to 320,000 camera frames were recorded at 50 or 100 fps, resulting in 10 min to 1.5 hr total acquisition time per dataset. For sample volumes thinner than 1.2  $\mu\text{m}$ , the sample stage was not translated during data acquisition. For thicker samples, the stage was moved axially in 500-nm steps every 1,000–3,000 frames until the whole targeted imaging volume was covered, resulting in up to 21 z-steps and imaging volume depths of up to 9  $\mu\text{m}$ . This axial scan was automatically repeated six to 19 times, and the data from the scans were combined. The laser intensity was manually adjusted during each experiment to optimize the emitter density per frame and to maximize detectable emissions starting at intensities as high as 35 kW/cm<sup>2</sup> to transfer emitters efficiently into dark states and decreasing to typically 5 kW/cm<sup>2</sup> near the end of the acquisition when the pool of blinking molecules had declined (Lin et al., 2015). Additionally, the laser intensity of the 405-nm laser was manually controlled over the course of imaging to optimize the active emitter density.

### Single-Molecule Interference Phase Estimation

Raw camera frames were first isolated into four different phase images. The four phase images were then merged into a single image using a transformation matrix obtained from a combination of algorithms using log-polar and affine transformations (Supplemental Information). Single-molecule candidates from the merged frames were isolated and fitted with an elliptical Gaussian model using a maximum likelihood estimator accounting for the camera-specific noise associated with sCMOS cameras (Huang et al., 2013) (Supplemental Information). Estimates of single-molecule positions, width, total number of detected photons, background, and log-likelihood ratio were obtained. Intensities of the 0<sup>th</sup> moment Gaussian (Supplemental Information) were calculated by a weighted least-square fitting of a Gaussian with the amplitude being the only fitting parameter and weighted to take the sCMOS noise into account (Huang et al., 2013). Subsequently, the reduced moments of each polarization were extracted using a previously described approach

(Aquino et al., 2011). For two-color imaging, phase shifts between *s* and *p* polarization for the two color channels differ by 0.3 radians. These phase shift differences were measured independently from images of fluorescent beads recorded in two detection channels. With the obtained phase differences between the *s* and *p* polarization, the phase shifts between the four different phase images are known. Subsequently, the interference phases of the detected molecules were obtained by solving a set of equations describing the 0<sup>th</sup> moment intensities in *s* and *p* polarization channels as functions of an unknown offset and their relative phase shifts as characterized above (Supplemental Information).

#### Axial Position Analysis with Ridge-Finding Algorithm

We developed a metric  $m: (\sigma_x^3/\sigma_y) - (\sigma_y^3/\sigma_x)$ , where  $\sigma_x$  and  $\sigma_y$  are the estimated SDs of the 2D Gaussian for a single-molecule emission pattern. The metric preserves its monotonicity in the presence of aberrations, and we used it to estimate an unambiguous (but still rough) position of each single emitter, before the phase estimate was used to pinpoint the exact axial position of a molecule.

For every 3,000–5,000 camera frames, we generated a 2D histogram image from phase estimations and the metric  $m$  from all detected single molecules. As  $m$  is monotonic and the single-molecule phase is periodic with a period of  $2\pi$ , the resulting 2D histogram looks like a pattern of tilted repeating stripes (Figure S3). We developed a ridge-finding algorithm to determine a series of connected vectors defining the correspondence between the determined phase values and the values of  $m$ . This allowed us to unwrap the periodic phase for unambiguous axial localization. A detailed description of the algorithm is provided in the Supplemental Information.

#### Drift Correction

3D drift correction was performed by first calculating the distance pairs between image segments (3,000–5,000 frames), and subsequently forming three sets of equations for *x*, *y*, and *z*, respectively (Li et al., 2013; Wang et al., 2014). A least-square solution that minimizes the overall error of the set of equations was obtained and back-substituted into all equations. Errors can be calculated from each of these equations and a specific equation within the set is removed from the stack if its error is larger than 7 nm. This process was repeated until no error was larger than 7 nm or the matrix was no longer at its full rank. This allowed us to correct system and sample induced drift in 3D with short segments of data (Supplemental Information).

#### Optical Segment Alignment

To image thick samples, optical sections were recorded at different axial positions of the sample by axially translating the *z*-piezo holding the sample stage. The localization data contain *x*, *y*, and *z* position estimates of different optical sections and must be aligned/stitched seamlessly to support the high precision obtained in W-4PiSMSN. Previous methods (Huang et al., 2008) that shifted each optical section by a constant in the axial direction have been prone to introduce misalignment of the optical sections and subsequently deteriorate the resolution achievable in thick samples. Here, we developed an optical alignment method based on 3D cross-correlation. In the W-4PiSMSN system, optical sections are  $\sim 1.2 \mu\text{m}$  thick. Whole-cell samples were scanned in the axial direction with 500-nm step sizes. This allowed for abundant overlapping regions between adjacent optical sections and provided critical information for precise optical section alignment using the developed 3D cross-correlation methods as described in Supplemental Information.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and seven movies and can be found with this article online at <http://dx.doi.org/10.1016/j.cell.2016.06.016>.

#### AUTHOR CONTRIBUTIONS

F.H., G.S., E.S.A., M.J.B., and J.B. designed the microscope. F.H., G.S., E.S.A., T.P., M.J.B., and J.B. built the instrument and developed the computer

code. F.H., G.S., E.S.A., L.K.S., W.C.D., E.B.K., Y.Z., and J.B. optimized and tested the microscope. L.K.S., W.C.D., F.E.R.-M., J.R.M., I.I., M.L., M.A.H., C.J.-W., C.P.L., J.E.R., D.T., and J.B. designed biological experiments. F.H., G.S., L.K.S., W.C.D., F.E.R.-M., J.R.M., I.I., and M.L. optimized sample preparation protocols and prepared samples. F.H., G.S., E.S.A., L.K.S., E.B.K., T.P., and J.B. visualized the data. All authors contributed to writing the manuscript.

#### ACKNOWLEDGMENTS

We would like to thank Travis Gould (Bates College) for initial discussion of the instrument design, Hideo Takakura, Hoong Chuin Lim, Stephanie Baguley (Yale University), and Yasuhiro Fujiwara (The Jackson Laboratory) for help in sample preparation, and Wen Jiang (Purdue University) for providing purified T7 phage samples. We thank Elias Coutavas and Günter Blobel (Rockefeller University) for the Nup358 antibody. We thank Mike Mosbacher and Christopher Jost for help on early CAD design and thermal simulations of the instrument, and Markus Schwarzhuber, Florian Schroeder, Stefan Politzka, Tobias Wagner, and Robert Muth for help on mechanical part design, custom electronics, sample chamber optimization, and data visualization. We thank George Takahashi from the Purdue Envision Center for help on data visualization. We thank Michael Mlodzianoski for discussions on the manuscript. This work was primarily supported by a grant from the Wellcome Trust (095927/A/11/Z). F.H. received funding from a James Hudson Brown–Alexander Brown Coxé Postdoctoral Fellowship. E.B.K. was supported by The Denmark–America Foundation (Coloplast), Civilingeniør Frants Allings Legat, Knud Højgaard's Fond, Reinholdt W. Jorcks Fond, Berg Nielsens Legat, and Ingeniør Alexandre Haynman og Hustru Nina Haynmans Fond. G.S. and E.S.A. were supported by the Wellcome Trust (092096) and Cancer Research UK (C6946/A14492). F.E.R.-M. and D.T. were supported by a grant from the NIH (R21 HD078851-01). I.I. was supported by a grant from the NIH (R01 GM065835) to C.J.-W. C.J.-W. is an investigator of the Howard Hughes Medical Institute. M.A.H. was supported by a grant from NIH (P01 GM99640). J.R.M. and C.P.L. were supported by grants from the NIH (R01 GM105672 and R01 HL124402). J.B. declares significant financial interest in Bruker, which manufactures SMSN instruments. F.H. and J.B. have filed a patent application for sCMOS-related data analysis (WO 2014144443 A2). The invention and software is licensed to Hamamatsu Photonics K.K.

Received: February 29, 2016

Revised: May 2, 2016

Accepted: June 3, 2016

Published: July 7, 2016

#### REFERENCES

- Aquino, D., Schönlé, A., Geisler, C., Middendorff, C.V., Wurm, C.A., Okamura, Y., Lang, T., Hell, S.W., and Egner, A. (2011). Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores. *Nat. Methods* 8, 353–359.
- Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacio, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642–1645.
- Bewersdorf, J., Schmidt, R., and Hell, S.W. (2006). Comparison of I5M and 4Pi-microscopy. *J. Microsc.* 222, 105–117.
- Briggs, J.A.G. (2013). Structural biology in situ—the potential of subtomogram averaging. *Curr. Opin. Struct. Biol.* 23, 261–267.
- Broeken, J., Johnson, H., Lidke, D.S., Liu, S., Nieuwenhuizen, R.P.J., Stallinga, S., Lidke, K.A., and Rieger, B. (2015). Resolution improvement by 3D particle averaging in localization microscopy. *Methods Appl. Fluoresc.* 3, 014003.
- Brown, T.A., Tkachuk, A.N., Shtengel, G., Kopek, B.G., Bogenhagen, D.F., Hess, H.F., and Clayton, D.A. (2011). Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction. *Mol. Cell Biol.* 31, 4994–5010.

- Burke, D., Patton, B., Huang, F., Bewersdorf, J., and Booth, M.J. (2015). Adaptive optics correction of specimen-induced aberrations in single-molecule switching microscopy. *Optica* 2, 177.
- Carlton, P.M. (2008). Three-dimensional structured illumination microscopy and its application to chromosome structure. *Chromosome Res.* 16, 351–365.
- Case, L.B., Baird, M.A., Shtengel, G., Campbell, S.L., Hess, H.F., Davidson, M.W., and Waterman, C.M. (2015). Molecular mechanism of vinculin activation and nanoscale spatial organization in focal adhesions. *Nat. Cell Biol.* 17, 880–892.
- Chojnacki, J., Staudt, T., Glass, B., Bingen, P., Engelhardt, J., Anders, M., Schneider, J., Müller, B., Hell, S.W., and Krüsslich, H.-G. (2012). Maturation-dependent HIV-1 surface protein redistribution revealed by fluorescence nanoscopy. *Science* 338, 524–528.
- Erdmann, R.S., Takakura, H., Thompson, A.D., Rivera-Molina, F., Allgeyer, E.S., Bewersdorf, J., Toomre, D., and Schepartz, A. (2014). Super-resolution imaging of the Golgi in live cells with a bioorthogonal ceramide probe. *Angew. Chem. Int. Ed. Engl.* 53, 10242–10246.
- Frey, T.G., and Mannella, C.A. (2000). The internal structure of mitochondria. *Trends Biochem. Sci.* 25, 319–324.
- Fritsche, M., Reinholdt, L.G., Lessard, M., Handel, M.A., Bewersdorf, J., and Heermann, D.W. (2012). The impact of entropy on the spatial organization of synaptonemal complexes within the cell nucleus. *PLoS ONE* 7, e36282.
- Giannone, G., Hosy, E., Levet, F., Constals, A., Schulze, K., Sobolevsky, A.I., Rosconi, M.P., Gouaux, E., Tampé, R., Choquet, D., and Cognet, L. (2010). Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Biophys. J.* 99, 1303–1310.
- Gould, T.J., Hess, S.T., and Bewersdorf, J. (2012a). Optical nanoscopy: from acquisition to analysis. *Annu. Rev. Biomed. Eng.* 14, 231–254.
- Gould, T.J., Burke, D., Bewersdorf, J., and Booth, M.J. (2012b). Adaptive optics enables 3D STED microscopy in aberrating specimens. *Opt. Express* 20, 20998–21009.
- He, M., Subramanian, R., Bangs, F., Omelchenko, T., Liem, K.F., Jr., Kapoor, T.M., and Anderson, K.V. (2014). The kinesin-4 protein Kif7 regulates mammalian Hedgehog signalling by organizing the cilium tip compartment. *Nat. Cell Biol.* 16, 663–672.
- Hell, S.W. (2007). Far-field optical nanoscopy. *Science* 316, 1153–1158.
- Hell, S.W., and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* 19, 780–782.
- Hess, S.T., Girirajan, T.P.K., and Mason, M.D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* 91, 4258–4272.
- Hu, B., Margolin, W., Molineux, I.J., and Liu, J. (2013). The bacteriophage  $\tau 7$  virion undergoes extensive structural remodeling during infection. *Science* 339, 576–579.
- Huang, B., Jones, S.A., Brandenburg, B., and Zhuang, X. (2008). Whole-cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution. *Nat. Methods* 5, 1047–1052.
- Huang, F., Hartwich, T.M.P., Rivera-Molina, F.E., Lin, Y., Duim, W.C., Long, J.J., Uchil, P.D., Myers, J.R., Baird, M.A., Mothes, W., et al. (2013). Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nat. Methods* 10, 653–658.
- Juette, M.F., Gould, T.J., Lessard, M.D., Mlodzianowski, M.J., Nagpure, B.S., Bennett, B.T., Hess, S.T., and Bewersdorf, J. (2008). Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples. *Nat. Methods* 5, 527–529.
- Jungmann, R., Avedaño, M.S., Woehrstein, J.B., Dai, M., Shih, W.M., and Yin, P. (2014). Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat. Methods* 11, 313–318.
- Kanchanawong, P., Shtengel, G., Pasapera, A.M., Ramko, E.B., Davidson, M.W., Hess, H.F., and Waterman, C.M. (2010). Nanoscale architecture of integrin-based cell adhesions. *Nature* 468, 580–584.
- Klehs, K., Spahn, C., Endesfelder, U., Lee, S.F., Fürstenberg, A., and Heilemann, M. (2014). Increasing the brightness of cyanine fluorophores for single-molecule and superresolution imaging. *ChemPhysChem* 15, 637–641.
- Lew, M.D., Lee, S.F., Ptacin, J.L., Lee, M.K., Twieg, R.J., Shapiro, L., and Moerner, W.E. (2011). Three-dimensional superresolution colocalization of intracellular protein superstructures and the cell surface in live *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. USA* 108, E1102–E1110.
- Li, X., Mooney, P., Zheng, S., Booth, C.R., Braunfeld, M.B., Gubbens, S., Agard, D.A., and Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* 10, 584–590.
- Lichtman, J.W., and Conchello, J.-A. (2005). Fluorescence microscopy. *Nat. Methods* 2, 910–919.
- Lin, Y., Long, J.J., Huang, F., Duim, W.C., Kirschbaum, S., Zhang, Y., Schroeder, L.K., Rebane, A.A., Velasco, M.G.M., Virrueta, A., et al. (2015). Quantifying and optimizing single-molecule switching nanoscopy at high speeds. *PLoS ONE* 10, e0128135.
- Liu, K.S.Y., Siebert, M., Mertel, S., Knoche, E., Wegener, S., Wichmann, C., Matkovic, T., Muhammad, K., Depner, H., Mettke, C., et al. (2011). RIM-binding protein, a central part of the active zone, is essential for neurotransmitter release. *Science* 334, 1565–1569.
- Liu, S., Kromann, E.B., Krueger, W.D., Bewersdorf, J., and Lidke, K.A. (2013). Three dimensional single molecule localization using a phase retrieved pupil function. *Opt. Express* 21, 29462–29487.
- McGorty, R., Schnitzbauer, J., Zhang, W., and Huang, B. (2014). Correction of depth-dependent aberrations in 3D single-molecule localization and super-resolution microscopy. *Opt. Lett.* 39, 275–278.
- Mlodzianowski, M.J., Schreiner, J.M., Callahan, S.P., Smolková, K., Dlasková, A., Santorová, J., Ježek, P., and Bewersdorf, J. (2011). Sample drift correction in 3D fluorescence photoactivation localization microscopy. *Opt. Express* 19, 15009–15019.
- Nieuwenhuizen, R.P.J., Lidke, K.A., Bates, M., Puig, D.L., Grünwald, D., Stallinga, S., and Rieger, B. (2013). Measuring image resolution in optical nanoscopy. *Nat. Methods* 10, 557–562.
- Ong, W.Q., Citron, Y.R., Schnitzbauer, J., Kamiyama, D., and Huang, B. (2015). Heavy water: a simple solution to increasing the brightness of fluorescent proteins in super-resolution imaging. *Chem. Commun. (Camb.)* 51, 13451–13453.
- Orci, L., Starnes, M., Ravazzola, M., Amherdt, M., Perrelet, A., Söllner, T.H., and Rothman, J.E. (1997). Bidirectional transport by distinct populations of COPI-coated vesicles. *Cell* 90, 335–349.
- Page, S.L., and Hawley, R.S. (2004). The genetics and molecular biology of the synaptonemal complex. *Annu. Rev. Cell Dev. Biol.* 20, 525–558.
- Patterson, G., Davidson, M., Manley, S., and Lippincott-Schwartz, J. (2010). Superresolution imaging using single-molecule localization. *Annu. Rev. Phys. Chem.* 61, 345–367.
- Pavani, S.R.P., Thompson, M.A., Biteen, J.S., Lord, S.J., Liu, N., Twieg, R.J., Piestun, R., and Moerner, W.E. (2009). Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl. Acad. Sci. USA* 106, 2995–2999.
- Pellett, P.A., Dietrich, F., Bewersdorf, J., Rothman, J.E., and Lavieu, G. (2013). Inter-Golgi transport mediated by COPI-containing vesicles carrying small cargoes. *eLife* 2, e01296.
- Qiao, H., Chen, J.K., Reynolds, A., Höög, C., Paddy, M., and Hunter, N. (2012). Interplay between synaptonemal complex, homologous recombination, and centromeres during mammalian meiosis. *PLoS Genet.* 8, e1002790.
- Ram, S., Prabhat, P., Chao, J., Ward, E.S., and Ober, R.J. (2008). High accuracy 3D quantum dot tracking with multifocal plane microscopy for the study of fast intracellular dynamics in live cells. *Biophys. J.* 95, 6025–6043.
- Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* 3, 793–795.

- Schücker, K., Holm, T., Franke, C., Sauer, M., and Benavente, R. (2015). Elucidation of synaptonemal complex organization by super-resolution imaging with isotropic resolution. *Proc. Natl. Acad. Sci. USA* *112*, 2029–2033.
- Sharonov, A., and Hochstrasser, R.M. (2006). Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proc. Natl. Acad. Sci. USA* *103*, 18911–18916.
- Shroff, H., Galbraith, C.G., Galbraith, J.A., and Betzig, E. (2008). Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nat. Methods* *5*, 417–423.
- Shtengel, G., Galbraith, J.A., Galbraith, C.G., Lippincott-Schwartz, J., Gillette, J.M., Manley, S., Sougrat, R., Waterman, C.M., Kanchanawong, P., Davidson, M.W., et al. (2009). Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proc. Natl. Acad. Sci. USA* *106*, 3125–3130.
- Van Engelenburg, S.B., Shtengel, G., Sengupta, P., Waki, K., Jarnik, M., Ablan, S.D., Freed, E.O., Hess, H.F., and Lippincott-Schwartz, J. (2014). Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits. *Science* *343*, 653–656.
- Vaughan, J.C., Jia, S., and Zhuang, X. (2012). Ultrabright photoactivatable fluorophores created by reductive caging. *Nat. Methods* *9*, 1181–1184.
- von Appen, A., Kosinski, J., Sparks, L., Ori, A., DiGiulio, A.L., Vollmer, B., Mackmull, M.-T., Banterle, N., Parca, L., Kastiris, P., et al. (2015). In situ structural analysis of the human nuclear pore complex. *Nature* *526*, 140–143.
- von Diezmann, A., Lee, M.Y., Lew, M.D., and Moerner, W.E. (2015). Correcting field-dependent aberrations with nanoscale accuracy in three-dimensional single-molecule localization microscopy. *Optica* *2*, 985–993.
- von Middendorff, C., Egner, A., Geisler, C., Hell, S.W., and Schönle, A. (2008). Isotropic 3D Nanoscopy based on single emitter switching. *Opt. Express* *16*, 20774–20788.
- Wang, W.-J., Tay, H.G., Soni, R., Perumal, G.S., Goll, M.G., Macaluso, F.P., Asara, J.M., Amack, J.D., and Tsou, M.-F.B. (2013). CEP162 is an axoneme-recognition protein promoting ciliary transition zone assembly at the cilia base. *Nat. Cell Biol.* *15*, 591–601.
- Wang, Y., Schnitzbauer, J., Hu, Z., Li, X., Cheng, Y., Huang, Z.-L., and Huang, B. (2014). Localization events-based sample drift correction for localization microscopy with redundant cross-correlation algorithm. *Opt. Express* *22*, 15982–15991.
- Wood, C.R., and Rosenbaum, J.L. (2015). Ciliary ectosomes: transmissions from the cell's antenna. *Trends Cell Biol.* *25*, 276–285.
- Xu, K., Babcock, H.P., and Zhuang, X. (2012). Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nat. Methods* *9*, 185–188.
- Xu, K., Zhong, G., and Zhuang, X. (2013). Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons. *Science* *339*, 452–456.
- Yang, T.T., Hampilos, P.J., Nathwani, B., Miller, C.H., Sutaria, N.D., and Liao, J.-C. (2013). Superresolution STED microscopy reveals differential localization in primary cilia. *Cytoskeleton* *70*, 54–65.
- Yang, T.T., Su, J., Wang, W.-J., Craige, B., Witman, G.B., Tsou, M.F., and Liao, J.-C. (2015). Superresolution pattern recognition reveals the architectural map of the ciliary transition zone. *Sci. Rep.* *5*, 14096.

## Creating a Selection to Fit Your Needs

Scissors · Retractors · Magnifiers · Probes & Hooks · Bone Instruments · Animal Identification  
Hemostats · Forceps · Surgical & Laboratory Equipment · Feeding Needles · Spatulae & Spoons  
Wound Closure · Surgical Plates · Instrument Care & Sterilization · Rongeurs · Scalpels & Knives  
Clamps · Pins & Holders · Needles & Needle Holders · Student Quality Instruments & Much More



***CAN'T FIND THE MOUSE MODEL YOU NEED?***

BUILD IT WITH  
**CRISPR!**

From simple indel KOs to more complex conditional KOs and reporter insertions, we've created a wide range of models using CRISPR/ Cas9 technology. If we aren't able to generate your CRISPR model, you don't pay.

**TO LEARN MORE,  
VISIT [JAX.ORG/CRISPR](http://JAX.ORG/CRISPR).**

1.800.422.6423

1.207.288.5845



**The Jackson  
Laboratory**

*Leading the search  
for tomorrow's cures*