





# IT'S LIGHTING THE WAY TO A NEW LEVEL OF UNDERSTANDING

Your everyday and innovative cellular assays can reveal deep biological understanding, if you've got the sensitive imaging and intuitive data analysis to uncover it. The Operetta CLS™ high-content analysis system features a unique combination of microscopy technologies – automated water-immersion objectives, high-power 8x LED

illumination, true confocal optics, and an ultrasensitive sCMOS camera – to deliver all the speed, sensitivity, and resolution you need to reveal subcellular details. And with our simple, powerful Harmony™ 4.5 software, Operetta CLS lets you discover subtle phenotypic changes. Operetta CLS: When great technologies combine, what happens is illuminating.



Operetta CLS High-Content Analysis System

Check out the new Operetta CLS HCA System at [www.perkinelmer.com/LightTheWay](http://www.perkinelmer.com/LightTheWay)

# Cell Reports

Best of 2017



Year 6 in Review

# Accuracy by Design; Not by Chance.

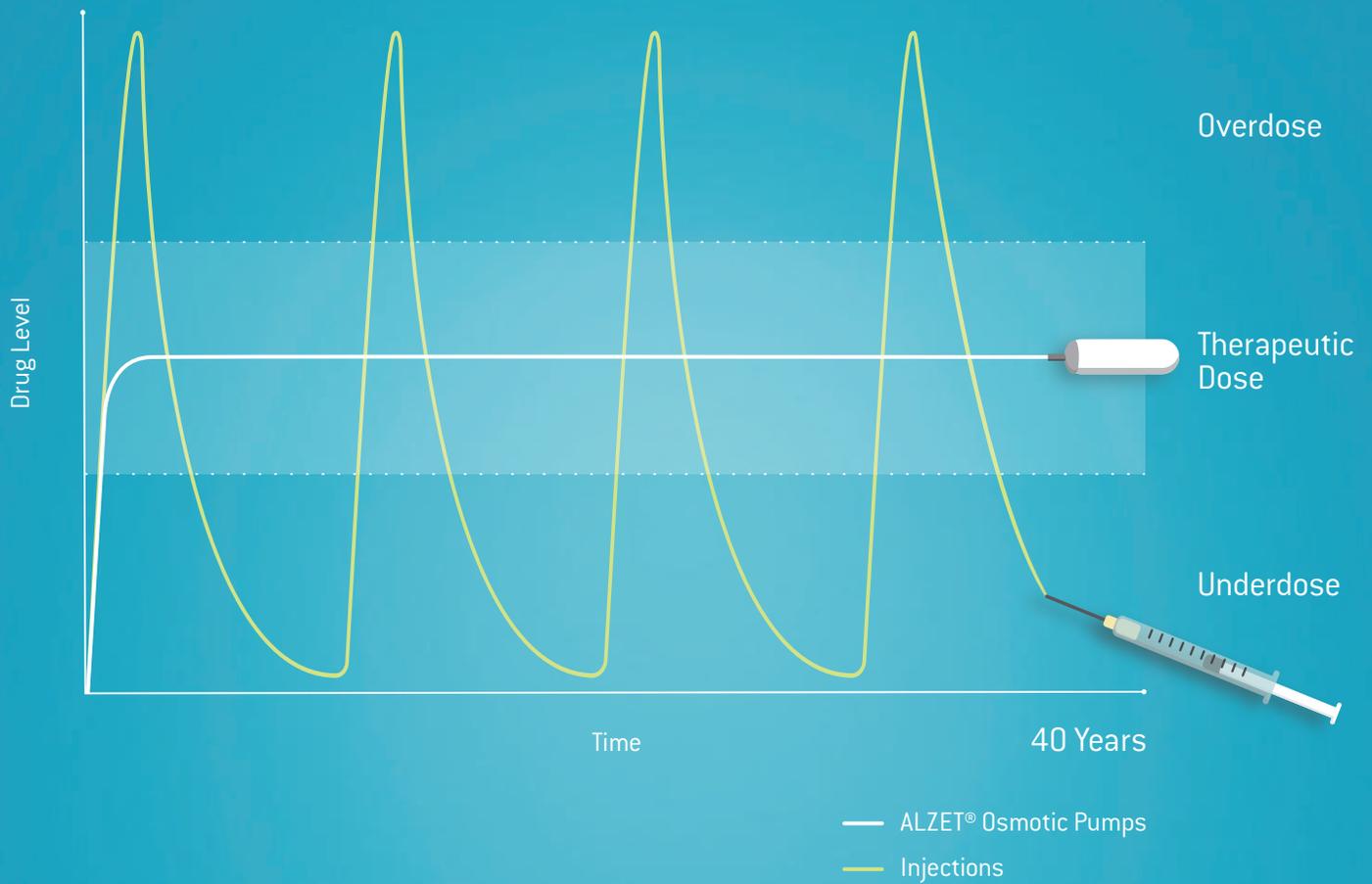
## Human XL Cytokine Discovery Luminex® High Performance Assay

- Broad selection – 45 analytes
- Superior accuracy – see the data
- Flexible – choose your analytes
- Easy to order – save time

**NEW**



Learn more | [rndsystems.com/human-xl](http://rndsystems.com/human-xl)



# 40 years of consistent dosing



ALZET Osmotic Pump, 100 µl  
[actual size]

**alzet**<sup>®</sup>  
OSMOTIC PUMPS

Find out more at  
[alzet.com/40years](http://alzet.com/40years)

A product of



©2017 DURECT Corporation  
All Rights Reserved





# sartorius

No matter the time of day or night,  
see exactly what happened  
to your cells and when.

Biological processes are dynamic, and a single snapshot in time may not capture rare or transient events, causing you to miss a relevant response. With the IncuCyte® S3 live-cell analysis system and reagents, automatically follow the sequence of biological events continuously, then 'rewind and replay' the experiment to see what really happened to your cells while you were away.

The IncuCyte S3 combines image-based measurements, a physiologically relevant environment, and microplate throughput to enable researchers to visualize and analyze cell behaviors at a scale and in ways that were previously not possible—all without ever removing cells from the incubator.

Visit [www.essenbio.com/IncuCyte](http://www.essenbio.com/IncuCyte) to learn about the next-generation IncuCyte S3 System and the benefits of real-time, automated live-cell analysis.



**IncuCyte**®  
by SARTORIUS

# F · S · T<sup>®</sup>

FINE SCIENCE TOOLS

## ADVANCING THE FIELD OF RESEARCH

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



FINE SURGICAL INSTRUMENTS FOR RESEARCH™

VISIT US AT [FINESCENCE.COM](http://FINESCENCE.COM) OR CALL 800 521 2109

# Focus on Results not Preparing Samples

*New Bioluminescent Assays for Energy Metabolism*

- In-well detection with no sample preprocessing
- Detect small metabolic changes in cells
- Use many sample types including: *media, tissue, plasma, serum and cells*

**Metabolite Detection Assays: *Glucose, Glucose Uptake, Lactate, Glutamate, Glutamine***

**Nucleotide/Co-Factor Detection Assays**

**Oxidative Stress Assays**

Sample an Assay Today:

**[www.promega.com/Energy](http://www.promega.com/Energy)**

# Antibody Validation

## DEFINE the Intended Antibody Application

Determine the type of immunoassay required.

Consider how the target protein may be modified in sample preparation. Consider how the antigen is identified by an antibody in a given application (e.g. cell type differentiation, protein localization, protein expression, or quantification).

## UNDERSTAND the Attributes of an Antigen

Find approved nomenclature of an antigen. Confirm protein sequence and molecular weight. Identify isoforms or variants.

Identify PTMs and other protein processing. Examine the presence of proteins closely related to target protein. Identify cellular localization.

## ESTABLISH the Controls for Antigen Detection

Evaluate peer-reviewed literature for intended target detection. Review databases for expression levels of intended target.

Determine types of cell lines, tissues, stimulation agents, transgenic or knockout models to use for confirmation. Establish the best controls for the intended antigen.

## REVIEW Antibody Vendor Information

Determine the type of immunogen used for antibody production. Establish whether monoclonal or polyclonal. Determine antibody isotype. Identify clone name or product code.

Establish host species. Find species reactivity. Analyze performance data on tested applications.

## PERFORM In-house Validation

Determine whether existing validation data is available. If data is unavailable, perform additional validation using the appropriate controls. Optimize the antibody concentrations and buffer conditions. Start with conditions suggested by the vendor.

Test the antibody against established controls to ensure specificity and reproducibility.

## Foreword

---

The end of 2017 marked 6 years of publishing excellent biology at *Cell Reports*, and so—drum roll—we present the *Best of Cell Reports 2017*. It was a year when we saw excellent work from across the life sciences at a variety of scales, from single-cell analyses of microglia and profiling aging changes in glial regional identity through to whole-body profiling of metastasis in mice. *Cell Reports* papers were also in the news in 2017, from genomic analyses of dog breeds to a topical agent that can induce melanin production. This made it particularly tough to select only ten papers, and as always, we based the collection on Altmetric scores, reader downloads, and citations. However, I would urge you to look at the journal and find out more about the full scope and reach of *Cell Reports*.

As we look forward to year 7, we want to use this chance to reach out and thank the authors, reviewers, editorial board, and advisors involved with the journal. Now kick back, grab a favorite beverage, and enjoy this collection of reports, resources, and articles published in *Cell Reports* in 2017.

Lastly, we are grateful for the generosity of our sponsors, who helped to make this reprint collection possible.

# Cell Reports

**YOU DON'T WORK  
PART TIME,**



**NEITHER SHOULD YOUR  
ANTIBODIES.**

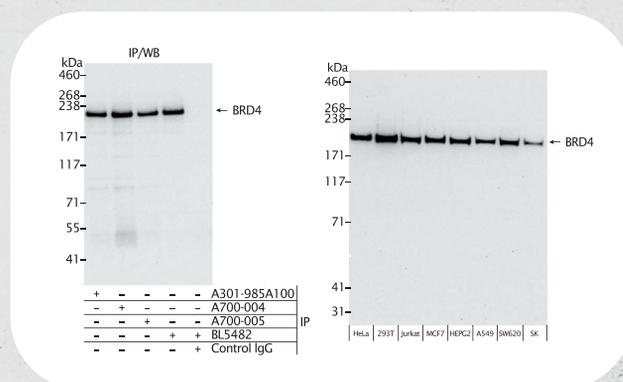
Studies show only 50% of antibodies can be trusted to work the way they're designed to.\* That's where Bethyl is different. With the standards and experience of over 40 years, we produce antibodies that deliver reliable results. We manufacture and validate everything on-site to ensure target specificity and sensitivity. All to guarantee our antibodies will function as designed in your assay 100% of the time. Choose Bethyl. We put a lot in every drop.

Discover free shipping<sup>†</sup> with your next trial size order:

[BETHYL.COM/SHIPTRIALFREE17](http://BETHYL.COM/SHIPTRIALFREE17)

**BETHYL**  
LABORATORIES, INC

<sup>†</sup>Terms & Conditions Apply. Please see website for trial sizes and complete details.  
©2017 Bethyl Laboratories, Inc. All rights reserved.



Detection of human BRD4 by rabbit anti-BRD4 recombinant monoclonal antibody Cat# A700-004 [BL-149-2H5] in WB of IPs (left) and WB of whole cell lysates (right). Rabbit anti-BRD4 recombinant monoclonal antibodies Cat# A700-004 [BL-149-2H5] and Cat# A700-005 [BL151-6F11], and affinity purified polyclonal antibody Cat# A301-985A100 used for IP (left).

\*Berglund, L., et al. A Genecentric Human Protein Atlas for Expression Profiles Based on Antibodies. *Molecular & Cellular Proteomics*, 7, 2019-27 (2009).

# Cell Reports

Best of 2017

## Reports

**A UV-Independent Topical Small-Molecule Approach for Melanin Production in Human Skin**

*Nisma Mujahid, Yanke Liang, Ryo Murakami, Hwan Geun Choi, Allison S. Dobry, Jinhua Wang, Yusuke Suita, Qing Yu Weng, Jennifer Allouche, Lajos V. Kemeny, Andrea L. Hermann, Elisabeth M. Roider, Nathanael S. Gray, and David E. Fisher*

**Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells**

*Yiyuan Liu, Bobbie Pelham-Webb, Dafne Campigli Di Giammartino, Jiexi Li, Daleum Kim, Katsuhiko Kita, Nestor Saiz, Vidur Garg, Ashley Doane, Paraskevi Giannakakou, Anna-Katerina Hadjantonakis, Olivier Elemento, and Effie Apostolou*

## Articles and Resources

**Whole-Body Profiling of Cancer Metastasis with Single-Cell Resolution**

*Shimpei I. Kubota, Kei Takahashi, Jun Nishida, Yasuyuki Morishita, Shogo Ehata, Kazuki Tainaka, Kohei Miyazono, and Hiroki R. Ueda*

**Major Shifts in Glial Regional Identity Are a Transcriptional Hallmark of Human Brain Aging**

*Lilach Soreq, UK Brain Expression Consortium, North American Brain Expression Consortium, Jamie Rose, Eyal Soreq, John Hardy, Daniah Trabzuni, Mark R. Cookson, Colin Smith, Mina Ryten, Rickie Patani, and Jernej Ule*

**Genomic Analyses Reveal the Influence of Geographic Origin, Migration, and Hybridization on Modern Dog Breed Development**

*Heidi G. Parker, Dayna L. Dreger, Maud Rimbault, Brian W. Davis, Alexandra B. Mullen, Gretchen Carpintero-Ramirez, and Elaine A. Ostrander*

**Entosis Is Induced by Glucose Starvation**

*Jens C. Hamann, Alexandra Surcel, Ruoyao Chen, Carolyn Teragawa, John G. Albeck, Douglas N. Robinson, and Michael Overholtzer*

**Host Genotype and Gut Microbiome Modulate Insulin Secretion and Diet-Induced Metabolic Phenotypes**

*Julia H. Kreznar, Mark P. Keller, Lindsay L. Traeger, Mary E. Rabaglia, Kathryn L. Schueler, Donald S. Stapleton, Wen Zhao, Eugenio I. Vivas, Brian S. Yandell, Aimee Teo Broman, Bruno Hagenbuch, Alan D. Attie, and Federico E. Rey*

**Clonal Expansion of Lgr5-Positive Cells from Mammalian Cochlea and High-Purity Generation of Sensory Hair Cells**

*Will J. McLean, Xiaolei Yin, Lin Lu, Danielle R. Lenz, Dalton McLean, Robert Langer, Jeffrey M. Karp, and Albert S.B. Edge*

**Temporal Tracking of Microglia Activation in Neurodegeneration at Single-Cell Resolution**

*Hansruedi Mathys, Chinnakkaruppan Adaikkan, Fan Gao, Jennie Z. Young, Elodie Manet, Martin Hemberg, Philip L. De Jager, Richard M. Ransohoff, Aviv Regev, and Li-Huei Tsai*

**$\beta$ -Hydroxybutyrate Deactivates Neutrophil NLRP3 Inflammasome to Relieve Gout Flares**

*Emily L. Goldberg, Jennifer L. Asher, Ryan D. Molony, Albert C. Shaw, Caroline J. Zeiss, Chao Wang, Ludmilla A. Morozova-Roche, Raimund I. Herzog, Akiko Iwasaki, and Vishwa Deep Dixit*

# Go figure

With author-narrated animations



Introducing Figure360, an author-narrated animation of select figures in Cell Press primary research and review journals.

A short, digestible synopsis puts the figure in context and helps you zoom in on the most important take-home message in a matter of minutes. Why go it alone when the author can help you figure it out in less than half the time?

**Check it out at [www.cell.com/figure360](http://www.cell.com/figure360)**

**Figure360**

**CellPress**



Nikon  
100<sup>th</sup>  
anniversary

# A1ways Evolving

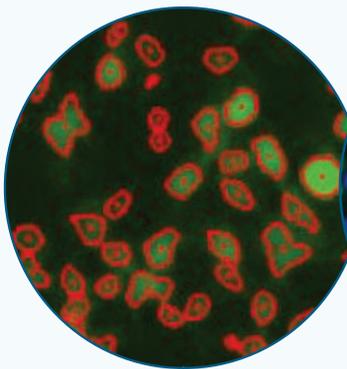
**NEW** High Definition 1K Resonant Scanner

Open up a world of new imaging strategies and possibilities with the latest upgrade to Nikon's always evolving A1R+ confocal microscope system. The all-new High Definition 1K Resonant Scanner delivers high resolution images at ultra-high speed. The new scanner also provides 4x the field of view at the same resolution usually generated by a normal 512x512 scanner. The wide dynamic range and reduced noise level raises the bar for image quality in resonant scanners.

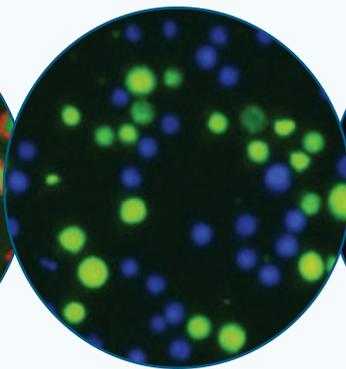
For more information, go to [www.nikoninstruments.com/a1r](http://www.nikoninstruments.com/a1r)  
or call 1-800-52-NIKON.



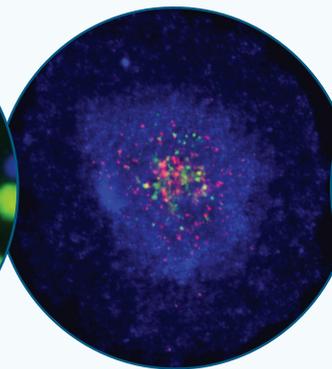
# Pioneers in Automated Cell Counting and Analysis



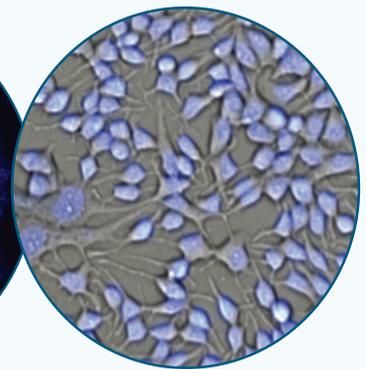
VIROLOGY



IMMUNO-ONCOLOGY



3D MODELS



SINGLE CELL ANALYSIS

Making your analysis count and your research excel in:

- ✓ High throughput foci and plaque counting
- ✓ 3D multicellular tumor-spheroid screening
- ✓ Cell mediated cytotoxicity assays for Immuno-Oncology
- ✓ Imaging, identification and quantification for single cell analyses

**Learn more** at [www.nexcelom.com](http://www.nexcelom.com)

Light re-engineered

An Employee-Owned Company

## Multiline Laser Illuminator - High Power, Low Cost

### LDI<sup>\*</sup>

- Up to 1W of power per channel
- Feedback controlled stability
- Up to 7 wavelengths
- Ultimate price to performance ratio

### LDI Applications

- Spinning Disk Confocal
- Photoactivation
- Optogenetics
- TIRF
- PALM / STORM

Works with X-Light V2 & OptoTIRF

[www.89north.com](http://www.89north.com)  
[sales@89north.com](mailto:sales@89north.com)  
+1.802.881.0302  
1.877.417.8313

## The World's Finest Optical Filters

### TIRF Filter Sets

- Highest possible signal to noise
- Custom filter cubes for precise dichroic alignment

### New Multiband Filter Sets

- Dedicated LED engine sets
- 5-band set for new Brilliant Violet™ dyes

Lifetime  
Warranty

Filter sets optimized for the LDI

[www.chroma.com](http://www.chroma.com)  
[sales@chroma.com](mailto:sales@chroma.com)  
+1.802.428.2500  
1.800.824.7662

# REDEFINING more

FLUORESCENCE | CHEMILUMINESCENCE | PHOSPHOR IMAGING



With three imaging modalities and outstanding performance, the **Sapphire™ Biomolecular Imager** expands what you can accomplish with a single image scanning instrument.

---

Find out what more really means—visit  
[azurebiosystems.com/sapphire](http://azurebiosystems.com/sapphire)

---



# A UV-Independent Topical Small-Molecule Approach for Melanin Production in Human Skin

Nisma Mujahid,<sup>1,2,5</sup> Yanke Liang,<sup>3,4,5</sup> Ryo Murakami,<sup>2</sup> Hwan Geun Choi,<sup>3,4</sup> Allison S. Dobry,<sup>2</sup> Jinhua Wang,<sup>3,4</sup> Yusuke Suita,<sup>2</sup> Qing Yu Weng,<sup>2</sup> Jennifer Allouche,<sup>2</sup> Lajos V. Kemeny,<sup>2</sup> Andrea L. Hermann,<sup>2</sup> Elisabeth M. Roider,<sup>2</sup> Nathanael S. Gray,<sup>3,4</sup> and David E. Fisher<sup>2,6,\*</sup>

<sup>1</sup>Department of Pathology and Laboratory Medicine, Boston University School of Medicine, Boston, MA 02118, USA

<sup>2</sup>Cutaneous Biology Research Center, Department of Dermatology, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129, USA

<sup>3</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

<sup>4</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

<sup>5</sup>These authors contributed equally

<sup>6</sup>Lead Contact

\*Correspondence: dfisher3@mgh.harvard.edu  
<http://dx.doi.org/10.1016/j.celrep.2017.05.042>

## SUMMARY

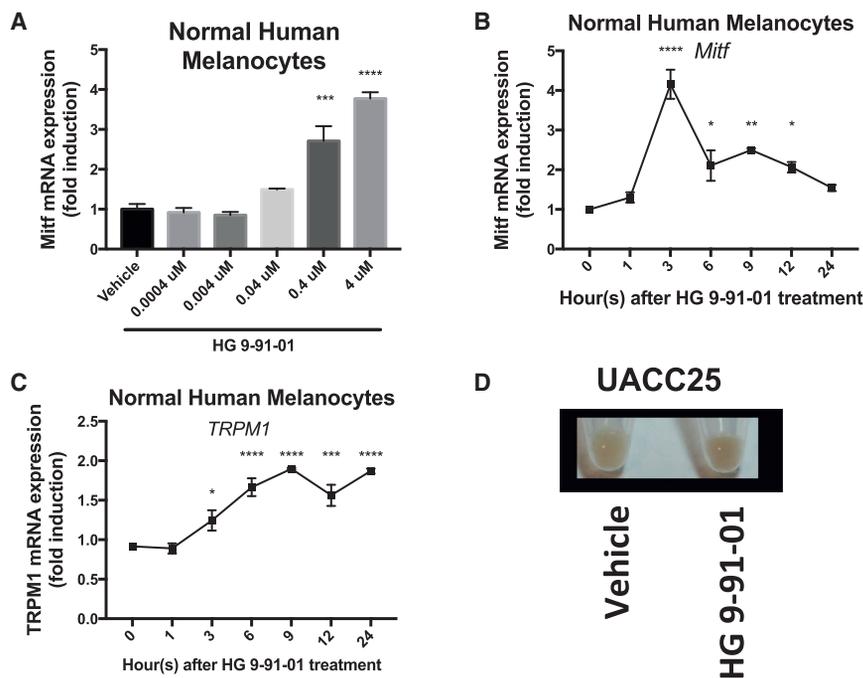
The presence of dark melanin (eumelanin) within human epidermis represents one of the strongest predictors of low skin cancer risk. Topical rescue of eumelanin synthesis, previously achieved in “redhaired” *Mc1r*-deficient mice, demonstrated significant protection against UV damage. However, application of a topical strategy for human skin pigmentation has not been achieved, largely due to the greater barrier function of human epidermis. Salt-inducible kinase (SIK) has been demonstrated to regulate MITF, the master regulator of pigment gene expression, through its effects on CRTC and CREB activity. Here, we describe the development of small-molecule SIK inhibitors that were optimized for human skin penetration, resulting in MITF upregulation and induction of melanogenesis. When topically applied, pigment production was induced in *Mc1r*-deficient mice and normal human skin. These findings demonstrate a realistic pathway toward UV-independent topical modulation of human skin pigmentation, potentially impacting UV protection and skin cancer risk.

## INTRODUCTION

The incidence of nonmelanoma and melanoma skin cancers has been increasing in the United States over recent decades (Rogers et al., 2015; Ryerson et al., 2016; Watson et al., 2016). Epidemiological evidence suggests that there is a causal relationship between sun/UV exposure and the three major histologic forms of skin cancer: squamous cell carcinoma, basal cell carcinoma, and cutaneous melanoma (Gandini et al., 2005; Kennedy et al., 2003; Wu et al., 2014). Individuals with fair skin and/or poor tanning ability are at higher risk for developing these malig-

nancies (Armstrong and Kricger, 2001), which are uncommon in darkly pigmented individuals (Pennello et al., 2000). During UV-induced tanning, DNA damage in keratinocytes triggers p53-mediated transcription of the pro-opiomelanocortin (*POMC*) gene (Cui et al., 2007). Proteolytic cleavage of POMC produces alpha-MSH (melanocyte-stimulating hormone), which binds to the melanocortin-receptor-1 (MC1R) on melanocytes, activating adenylate cyclase. Elevated cyclic AMP (cAMP) activates protein kinase A (PKA), which phosphorylates the cAMP-responsive-element-binding protein (CREB) (Newton et al., 2005; Tsatmalia et al., 1999), which, in turn, stimulates the transcription of the microphthalmia-associated transcription factor (*MITF*) gene (Bertolotto et al., 1998; Price et al., 1998). *MC1R* non-signaling variants are associated with lighter skin tones and red hair and are linked to poor tanning responses (Valverde et al., 1995). Previously, topical application of the cAMP agonist forskolin was shown to rescue the cAMP-MITF-eumelanin pathway in *Mc1r*-deficient mice (D’Orazio et al., 2006). Subsequent studies identified the phosphodiesterase PDE4D3 as a key regulator of melanocytic cAMP homeostasis, and its suppression produced hyperpigmentation similar to forskolin treatment in red-haired mice (Khaled et al., 2010). However, attempts to apply both of these small-molecule approaches to human skin have been unsuccessful, likely related to poor skin penetration of the active species.

Genetic data in mice have suggested the presence of a pathway in which CREB-regulated transcription co-activator (CRTC) positively regulates and salt-inducible kinase 2 (SIK2) negatively regulates MITF and pigment synthesis independently of CREB phosphorylation by PKA (Horike et al., 2010). In macrophages, the small-molecule SIK inhibitor HG 9-91-01 has been shown to regulate CREB-dependent gene transcription by suppressing phosphorylation of CRTC (Clark et al., 2012), thereby inhibiting cytoplasmic sequestration and permitting its nuclear translocation. We hypothesized that small-molecule SIK inhibitors could be generated and optimized as topical agents capable of inducing cutaneous pigmentation independently of UV irradiation in human skin.



**Figure 1. Inhibition of SIK by HG 9-91-01 Promotes *MITF* Transcription and Pigmentation In Vitro**

(A) mRNA expression of *MITF* relative to *RPL11* mRNA and vehicle control in normal human melanocytes 3 hr after HG 9-91-01 or vehicle control (70% ethanol, 30% propylene glycol) treatment, quantified by qRT-PCR (n = 3, mean ± SEM).

(B and C) mRNA expression of *MITF* (B) and *MITF*-dependent gene *TRPM1* (C) relative to *RPL11* mRNA and vehicle control at each time point, in normal human melanocytes over 24 hr after 4 μM HG 9-91-01 or vehicle control treatment, quantified by qRT-PCR (n = 3, mean ± SEM).

(D) Cell pellets of UACC257 melanoma cells after 3 days of treatment with vehicle control or 4 μM SIK inhibitor HG 9-91-01 (image is representative of n = 3 experiments).

For the graph in (A), statistical significance is reported as follows: \*\*\*\*p < 0.0001, one-way ANOVA with Dunnett's multiple comparisons test comparing treatment dose to vehicle control. For the graphs in (B) and (C), statistical significance is reported as follows: \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001; \*\*\*\*p < 0.0001, repeated-measures one-way ANOVA with Dunnett's multiple comparisons test comparing each time point to time point 0.

## RESULTS

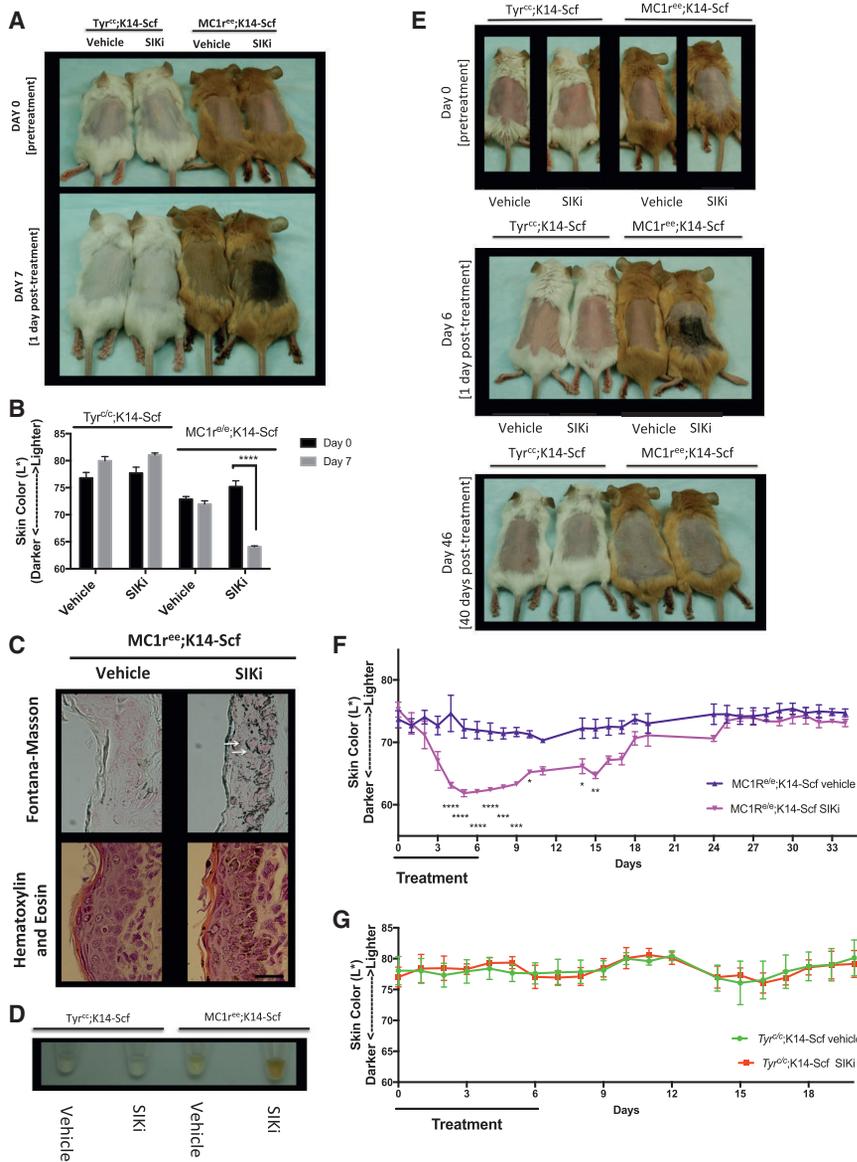
### Small-Molecule Inhibition of SIK Induces *MITF* Expression In Vitro

To test regulation of the pigmentation pathway by the previously published SIK inhibitor HG 9-91-01 (HG) (Clark et al., 2012) in vitro, we treated normal human melanocytes, UACC62 human melanoma cells, and UACC257 human melanoma cells. Dose-dependent increases in expression of *MITF* were observed in these cells in response to SIK inhibitor application (Figures 1A, S1A, and S1D). RNA levels of the *MITF* target gene *TRPM1* (Miller et al., 2004) also increased and followed the anticipated delayed kinetics relative to *MITF* induction in normal human melanocytes (Figures 1B and 1C) and UACC257 human melanoma cells (Figures S1G and S1H). Gross pigmentation was observed in cell pellets of UACC257 human melanoma cells after 3 days of HG 9-91-01 treatment (Figure 1D). Since SIK kinase activity is known to be dependent on LKB1 (Kato et al., 2006) we next evaluated whether SIK-inhibitor treatment of LKB1-null G361 melanoma cells would induce *MITF*. In LKB1-null G361 melanoma cells, there is no *MITF* induction with SIK-inhibitor treatment (Figure S1J). In contrast, when LKB1 is introduced in G361 melanoma cells (Figure S1I), we observed a 6-fold induction of *MITF* expression with SIK-inhibitor treatment (Figure S1K), demonstrating the dependence of SIK-inhibitor effect on active SIK. These data suggest that small-molecule SIK inhibition can stimulate the pigmentation pathway in vitro.

### HG 9-91-01 Rescues Melanogenesis in Mice with Inactive Melanocortin 1 Receptor

Since our in vitro results demonstrated that inhibition of SIK by HG 9-91-01 positively regulated *MITF* transcription, we next

evaluated whether topical application of this compound could induce pigmentation independent of MC1R in vivo. To test this, we utilized a previously described mouse "red hair" model that carries the inactivating *Mc1r<sup>el/e</sup>* mutant allele and a transgene, K14-SCF, in which stem cell factor expression is driven by the keratin-14 promoter, allowing for epidermal homing of melanocytes (D'Orazio et al., 2006; Kunisada et al., 1998). Albino mice harboring a mutation in the *tyrosinase* gene were combined with the K14-SCF transgene (*Tyr<sup>cl/c</sup>*;K14-SCF mice) and served as controls to evaluate whether the pigmentation afforded by topical SIK inhibitor was dependent upon the canonical tyrosinase-melanin pathway. Daily application of the SIK inhibitor HG 9-91-01 for 7 days caused robust darkening in *Mc1r<sup>el/e</sup>*;K14-SCF mice (Figures 2A and S2A). No visible change in skin pigmentation was observed in *Mc1r<sup>el/e</sup>*;K14-SCF mice treated with vehicle or in *Tyr<sup>cl/c</sup>*;K14-SCF mice treated with vehicle or HG 9-91-01 (Figures 2A, S2A, and S2B). Reflective colorimetry analysis (Commission Internationale de l'Eclairage [CIE] L\* white-black color axis (Park et al., 1999)) revealed significant darkening in *Mc1r<sup>el/e</sup>*;K14-SCF mice treated with SIK inhibitor, but not in vehicle-treated *Mc1r<sup>el/e</sup>*;K14-SCF mice or in *Tyr<sup>cl/c</sup>*;K14-SCF mice treated with either SIK inhibitor or vehicle control (70% ethanol, 30% propylene glycol) (Figure 2B). Fontana-Masson staining, a specialized melanin stain, revealed strong induction of melanin production in *Mc1r<sup>el/e</sup>*;K14-SCF mice only in areas treated with HG 9-91-01 (Figures 2C and S2D) but no pigment induction in *Mc1r<sup>el/e</sup>*;K14-SCF mice treated with vehicle (Figure 2C) or in albino (*Tyr<sup>cl/c</sup>*;K14-SCF) mice treated with either vehicle or SIK inhibitor (Figure S2C). Nuclear capping of melanin-laden melanosomes was observed within epidermal keratinocytes in *Mc1r<sup>el/e</sup>*;K14-SCF mice treated with HG 9-91-01 (indicated by white arrows) and represents a known sub-cellular localization typical of physiologic skin pigmentation



**Figure 2. Topical Treatment with HG9-91-01 Causes Robust Darkening that Is Progressive and Reversible in *Mc1r<sup>e/e</sup>;K14-SCF* Mice**

(A–D) Shown here: (A) *Mc1r<sup>e/e</sup>;K14-SCF* mice and *Tyr<sup>c/c</sup>;K14-SCF* mice before treatment (day 0) and after 7 days of treatment (day 7) with 30  $\mu$ L vehicle control (70% ethanol, 30% propylene glycol) or 37.5 mM HG 9-91-01 (image is representative of  $n = 4$  experiments). (B) Reflective colorimetry measurements ( $L^*$  white-black color axis;  $n = 4$ , mean  $\pm$  SEM) and (D) melanin extraction (image is representative of  $n = 4$  experiments) of the *Mc1r<sup>e/e</sup>;K14-SCF* mice and *Tyr<sup>c/c</sup>;K14-SCF* mice described in (A). (C) Skin sections of *Mc1r<sup>e/e</sup>;K14-SCF* mice described in (A) stained with Fontana-Masson (eumelanin) (top two panels) or H&E (bottom two panels); (magnification, 400 $\times$ ). White arrows represent nuclear capping; scale bar represents 25  $\mu$ m.

(E) *Mc1r<sup>e/e</sup>;K14-SCF* mice and *Tyr<sup>c/c</sup>;K14-SCF* mice before treatment (day 0) and after 6 days of treatment with 30  $\mu$ L vehicle control (70% ethanol, 30% propylene glycol) or 37.5 mM HG 9-91-01 (day 6), and 40 days post-treatment (day 46) (vehicle mouse in day-46 photo is different from that in the day-0 and day-6 photos).

(F and G) Reflective colorimetry measurements (CIE  $L^*$  white-black color axis) of (F) *Mc1r<sup>e/e</sup>;K14-SCF* mice and (G) *Tyr<sup>c/c</sup>;K14-SCF* mice treated as described in (E). Vehicle-treated *Mc1r<sup>e/e</sup>;K14-SCF* mice:  $n = 5$  (days 0–19), and  $n = 4$  (days 24–34); HG 9-91-01-treated *Mc1r<sup>e/e</sup>;K14-SCF* mice:  $n = 3$ ; vehicle-treated *Tyr<sup>c/c</sup>;K14-SCF* mice:  $n = 3$  (days 0–10), and  $n = 2$  (days 11–20); HG 9-91-01-treated *Tyr<sup>c/c</sup>;K14-SCF* mice:  $n = 3$  (mean  $\pm$  SEM).

For the graph in (B), statistical significance is reported as follows: \*\*\*\* $p < 0.0001$ , multiple t test analysis with the two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli. For the graphs in (F) and (G), statistical significance is reported as follows: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ , two-way ANOVA with Sidak's multiple comparisons test comparing treatment to vehicle control at each time point.

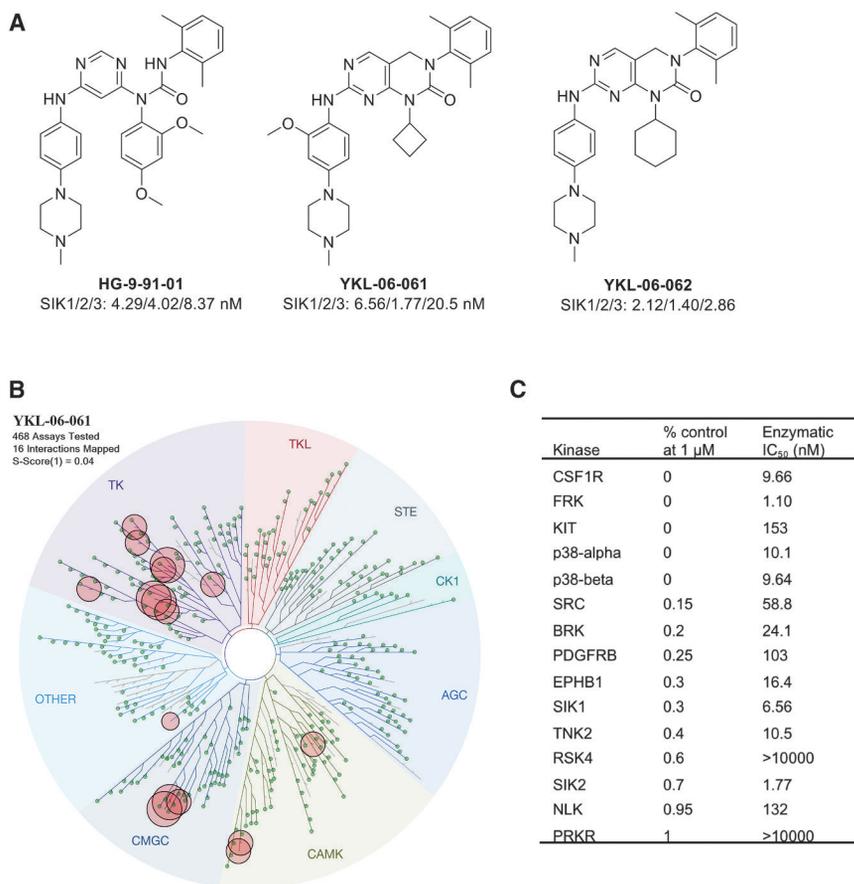
(Kobayashi et al., 1998) (Figure 2C). This feature suggests that SIK-inhibitor treatment stimulates not only melanocytic pigment synthesis but also the export of melanin in a fashion that closely mimics the known pathway of UV melanogenesis. H&E staining revealed normal morphology of HG 9-91-01-treated *Mc1r<sup>e/e</sup>;K14-SCF* (Figure 2C) and *Tyr<sup>c/c</sup>;K14-SCF* epidermis (Figure S2C). NaOH lysis of skin samples (Wakamatsu and Ito, 2002) revealed a visible increase in extractable eumelanin from *Mc1r<sup>e/e</sup>;K14-SCF* mice treated with HG 9-91-01, compared with all other treatment groups (Figure 2D).

Darkening induced by topical application of HG 9-91-01 to *Mc1r<sup>e/e</sup>;K14-SCF* mice was progressive over 6 days of treatment and gradually reversed over the 2 weeks after treatment was stopped (Figure 2F). Skin pigmentation remained in its pretreatment state 26 days later (40 days after treatment ended) (Figure 2E). No change was observed in *Tyr<sup>c/c</sup>;K14-SCF* mice during

treatment or 14 days after treatment was stopped (Figure 2G). Forty days after treatment was stopped, Fontana-Masson staining of skin sections of *Mc1r<sup>e/e</sup>;K14-SCF* mice and *Tyr<sup>c/c</sup>;K14-SCF* mice revealed no differences between vehicle and treatment groups, and H&E staining illustrated normal morphology for all mice (Figure S2E). These findings combined with the small-molecule and lipophilic nature of the SIK inhibitors led us to further investigate the use of SIK inhibitors for topical eumelanization of human skin.

### Second-Generation SIK Inhibitors Are as Efficacious Inducing the Pigmentation Pathway as HG 9-91-01

Since there are limitations to topical delivery of HG 9-91-01 into human skin epidermis (Figures 4A–4D), we derived SIK inhibitors designed to enhance passive epidermal permeation utilizing Lipinski's Rule of Five, which predicts greater absorption of



**Figure 3. Characterization of SIK Inhibitors**

(A) Structures of HG-9-91-01, YKL-06-061, and YKL-06-062 and their biochemical  $IC_{50}$ s against SIKs.

(B) KinomeScan kinase selectivity profile for YKL-06-061. YKL-06-061 was profiled at a concentration of 1  $\mu$ M against a diverse panel of 468 kinases by DiscoverX. Kinases that exhibited a score of 1 or below are marked in red circles. (Score is percent relative to DMSO control. Smaller numbers indicate stronger binding.) See Table S1 for full kinome profile.

(C) Biochemical kinase  $IC_{50}$ s of YKL-06-061 top hits as shown in (B).

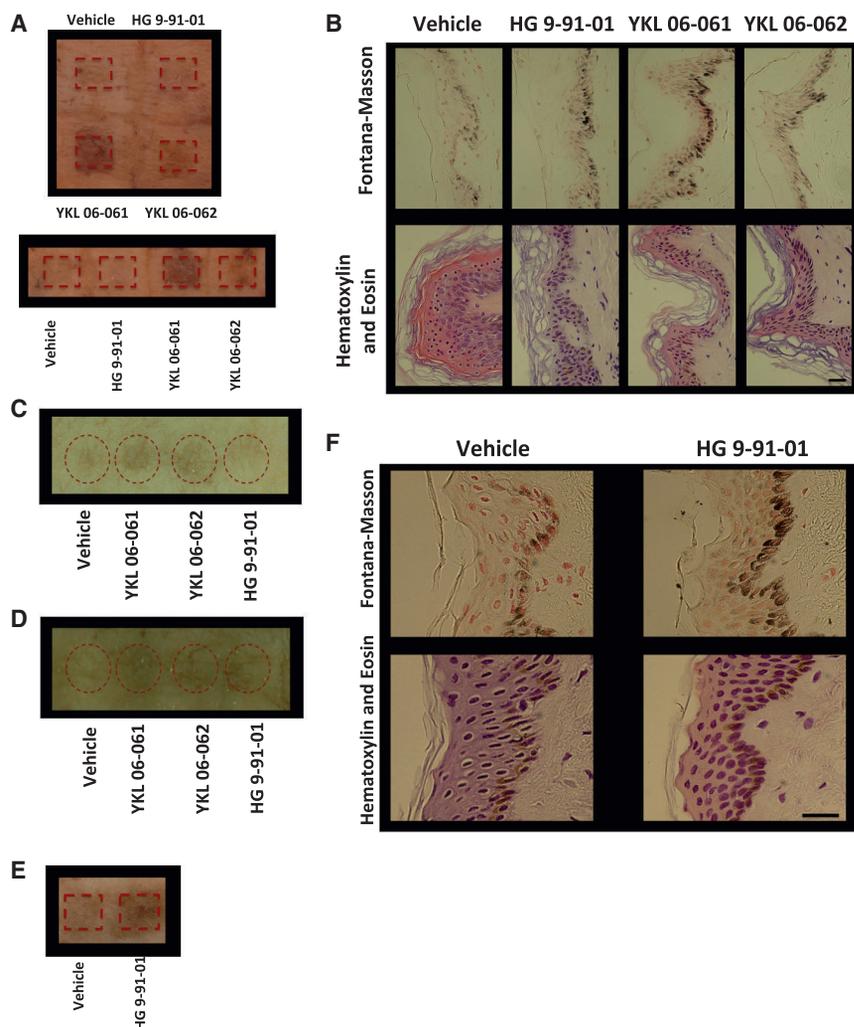
TK, tyrosine kinase; TKL, tyrosine kinase-like; STE, homologs of yeast sterile 7, sterile 11, sterile 20 kinases; CK1, casein kinase 1; AGC, containing PKA, PKG, and PKC families; CAMK, calcium/calmodulin-dependent protein kinase; CMGC, containing CDK, MAPK, GSK3, and CLK families. See also Table S1.

compounds if they have fewer than five H-bond donors, fewer than ten H-bond acceptors, a molecular weight less than 500 g/mol, and calculated log P (CLogP) less than 5 (Bos and Meinardi, 2000; Choy and Prausnitz, 2011; Lipinski et al., 2001) (Figure S3A). In an initial screen, two second-generation SIK inhibitors, YKL 06-061 and YKL 06-062, induced darkening as measured by reflective colorimetry analysis after topical treatment of human breast skin explants (Figure S3B). Furthermore, both YKL 06-061 and YKL 06-062 have a lower molecular weight than HG 9-91-01, and the more efficacious YKL 06-061 has a lipophilicity closer to Lipinski's Rule of Five, possibly explaining the drug's enhanced penetration capabilities (Figure S3A). Second-generation inhibitors had half maximal inhibitory concentration ( $IC_{50}$ ) values for the inhibition of SIK1, SIK2, and SIK3 that were comparable to those of HG 9-91-01 (Figure 3A). To assess the kinome selectivity information of new analogs, YKL-06-061 was screened across a panel of 468 human kinases at a concentration of 1  $\mu$ M using the KinomeScan methodology (DiscoverX). YKL-06-061 exhibited an S(1) score of 0.02, with 16 kinases displaying tight binding to it (Ambic scores of  $\leq 1$ ) (Figure 3B). As the KinomeScan assays measure binding, we also performed enzymatic assays for these targets either in house or using the SelectScreen Kinase Profiling Service at Thermo Fisher Scientific (Figure 3C). YKL-06-061 inhibited only one kinase, fyn-related kinase (FRK), more strongly than SIKs, which demonstrates its

high overall selectivity (Figure 3C). We anticipate that YKL-06-062 has similar kinase selectivity, considering their high structural similarity. Similar to observations with HG 9-91-01, treatment of normal human melanocytes (Figures S3C and S3D), UACC62 human melanoma cells, and UACC257 human melanoma cells (Figures S1B, S1C, S1E, and S1F) with YKL 06-061 or YKL 06-062 for 3 hr yielded a dose-dependent increase in MITF mRNA expression. Levels of

### Topical SIK Inhibitors Induce Human Skin Eumelanization

*TRPM1* mRNA increased after MITF induction upon treatment with YKL 06-061 or YKL 06-062 in normal human melanocytes and UACC257 human melanoma cells (Figures S1G, S1H, S3E, and S3F). Treatment of human skin explants with passive topical application of the second-generation SIK inhibitors, YKL 06-061 and YKL 06-062, induced significant pigmentation after 8 days of treatment (1 $\times$ /day), but no significant gross pigmentation was observed in skin treated with HG 9-91-01 (Figure 4A). Fontana-Masson staining revealed increased melanin content in skin treated with YKL 06-061 or YKL 06-062 and marginally increased melanin in skin treated with HG 9-91-01, as compared with control (Figure 4B). This effect was reproducible with independent preparations of synthesized drugs applied passively (via pipette) to the top of different human skin explants (Figure 4C and 4D). Mechanical application of the first-generation SIK inhibitor HG 9-91-01, by rubbing via an applicator, induced significant gross pigmentation (Figure 4E), and increased melanin content was observed upon Fontana-Masson staining of skin sections (Figure 4F), suggesting that HG 9-91-01's limited human skin penetration can be, at least partially, overcome through mechanical application. YKL 06-061 and YKL 06-062 did not require



**Figure 4. Treatment of Human Skin Explants with 37.5 mM of SIK Inhibitor Induces Pigmentation**

(A) Human breast skin explants treated with passive application of vehicle control (70% ethanol, 30% propylene glycol) or 37.5 mM SIK inhibitor YKL 06-061, YKL 06-062, or HG 9-91-01 for 8 days (10  $\mu$ L; 1 $\times$ /day). Image was taken 2 days after the end of treatment (image is representative of two of  $n = 3$  experiments).

(B) Fontana-Masson (top panel) and H&E (bottom panel) staining (magnification, 400 $\times$ ) of breast skin described in (A). Scale bar represents 25  $\mu$ m.

(C) Human breast skin explants treated with passive application of vehicle control or 37.5 mM SIK inhibitor YKL 06-061, YKL 06-062, or HG 9-91-01 for 5 days (10  $\mu$ L; 2 $\times$ /day). Image was taken 1 day after the end of treatment (image is representative of  $n = 1$  experiment).

(D) Human breast skin explants treated with passive application of vehicle control or 37.5 mM SIK inhibitor YKL 06-061, YKL 06-062, or HG 9-91-01 for 6 days (10  $\mu$ L; 2 $\times$ /day). Image was taken 1 day after the end of treatment (image is representative of  $n = 1$  experiment).

(E) Human breast skin explants treated with mechanical application of vehicle control or 50 mM (50  $\mu$ L for 1 day; 1 $\times$ /day) or 25 mM (50  $\mu$ L for 3 days; 3 $\times$ /day) HG 9-91-01. Image was taken 4 days after the start of treatment (image is representative of  $n = 1$  experiment).

(F) Fontana-Masson (top panels) and H&E (bottom panels) staining (magnification, 400 $\times$ ) of human skin explants described in (E). Scale bar represents 25  $\mu$ m.

mechanical application (rubbing) to induce significant human epidermal darkening.

## DISCUSSION

These results illustrate the development and successful application of small-molecule SIK inhibitors for topical induction of skin pigmentation independently of UV irradiation in human skin. SIK inhibitors were shown to induce enhanced expression of the MITF transcription factor, which is known to regulate expression of numerous pigment enzymes that promote biosynthesis of eumelanin. A new generation of SIK inhibitors was developed, based on strategies for enhancing the likelihood of skin penetration through optimizing molecular size and lipophilicity. Two such SIK-targeted inhibitors, YKL 06-061 and YKL 06-062, were shown to induce similar responses both in vitro and when applied to human skin explants. In addition to upregulating mRNA levels of *MITF* and *TRPM1*, topical SIK inhibitors were seen to trigger the transfer of melanosomes into epidermal keratinocytes in a manner that recapitulates the perinuclear capping (subcellular localization) seen in normal human

epidermal pigmentation. Thus, SIK-inhibitor treatments appear to induce not only synthesis of melanin but also melanosomal maturation, export, and localization features, even after import into keratinocytes. These features closely resemble the previously observed behavior of forskolin treatment in red-haired mice (D’Orazio et al., 2006).

Topical application of small-molecule, UV-independent pigment inducers has not yet been examined in humans and would require careful considerations of safety. For example, the induction of dark pigmentation is associated with the lowest risk of most skin cancers in humans (Armstrong and Kricger, 2001; Pennello et al., 2000), and this pigment synthesis is believed to be dependent upon MITF (Bertolotto et al., 1998). However, fixed genomic mutation or amplification of the *MITF* gene can be oncogenic in certain contexts (Bertolotto et al., 2011; Garraway et al., 2005; Yokoyama et al., 2011). Reversible upregulation of MITF, as reported here, is also likely to occur in routine instances of UV tanning, and constitutive elevation of MITF is likely in the skin of individuals with darker pigmentation levels; neither would be anticipated to trigger genomic mutation of the *MITF* gene. Analogously, transient administration of recombinant hematopoietic growth factors has not been associated with formation of oncogenic transformation or leukemia (Dombret et al., 1995; Ohno et al., 1990). In mice, topical

forskolin's pigmentary rescue in "redheads" resulted in significant protection from UV carcinogenesis, without apparent associated toxicities over many months of treatment (D'Orazio et al., 2006). A recent study has utilized injections of the synthetic alpha-MSH analog, afamelanotide, for treatment of photosensitivity associated with erythropoietic protoporphyria. Pigmented lesions/melanoma were carefully evaluated and reported not to occur at elevated risk (Langendonk et al., 2015).

Our *in vivo* studies demonstrate that topical SIK inhibitor can be applied with localized SIK inhibition and no detected systemic effects in mice (such as failure to thrive); and although it has been previously shown that SIK1 inhibition leads to cell-cycle arrest in epithelial cells, this was dependent on the presence of transforming growth factor  $\beta$  (TGF- $\beta$ ) (Lönn et al., 2012), and we observed normal skin turnover with no morphological changes of the skin (measured grossly or histologically, other than pigment/color). During normal UV-induced tanning, MC1R activation leads to enhanced PKA activity (Newton et al., 2005), and PKA-dependent phosphorylation of SIK1 (Takemori et al., 2002), SIK2 (Horike et al., 2003), and SIK3 (Kato et al., 2006) decreases their kinase activity. Since our compounds' activity is analogous to the on/off switch of UV-induced tanning, we believe that it will be a safe, viable method of topical pigment production, though it may be important to assure localized delivery to skin.

The half-life of melanin in skin is thought to be several weeks and diminishes primarily after superficial keratinocyte sloughing. Most epidermal melanin resides within keratinocytes after transfer of melanosomes from melanocytes. Therefore, it is possible that small-molecule approaches like that described here might be achievable, or maintained, through intermittent pulse-dosing strategies, thereby further limiting systemic drug exposure. In conclusion, these studies describe a small-molecule, topical approach to the rescue of eumelanin synthesis in a UV-independent manner. Future studies will be needed to examine the optimal applications of such agents in a variety of clinical settings.

## EXPERIMENTAL PROCEDURES

See the Supplemental Information for detailed methods.

### Materials

SIK inhibitors were dissolved in 30% propylene glycol plus 70% ethanol. HG-91-01 was purchased from Medchem Express, and all other SIK inhibitors were synthesized by the authors.

### Kinome Profiling

Kinome profiling was performed using KinomeScan ScanMAX at a compound concentration of 1  $\mu$ M. Data are reported in the Supplemental Information. Protocols are available from DiscoverX.

### Kinase Activity *In Vitro* Assay

The biochemical activities against SIK2 were measured with a Caliper-based mobility shift assay (PerkinElmer).

### Real-Time qPCR

The relative expression of each gene was calculated with 7500 Fast Real-Time PCR System software, which utilizes *Ct* normalized to mRNA levels of *RPL11* to calculate relative expression. Results are reported relative to control cells.

### Mice

C57BL/6J *Mc1r*<sup>+/e</sup> mice were crossed with K14-SCF transgenic mice, and C57BL/6J *Tyr*<sup>c2/c2</sup> were crossed with K14-SCF transgenic mice (D'Orazio et al., 2006; Kunisada et al., 1998). Mixed-gender adult mice were used. All animal experiments were performed in accordance with institutional policies and Institutional Animal Care and Use Committee-approved protocols.

### Human Tissue Samples

Skin samples considered surgical waste were obtained de-identified from healthy donors undergoing reconstructive surgery, according to institutional regulation.

### Colorimeter Measurements

Differences in darkening of the skin were measured by reflective colorimetry (Commission Internationale de l'Eclairage [CIE] L\* white-black color axis) utilizing a CR-400 Colorimeter (Minolta) calibrated to a white standard background calibration plate, with calibration date set to Y 93.1, x 0.3133, y 0.3194, before each set of measurements.

### Statistical Analysis

Data are presented as the mean  $\pm$  SEM.

Statistical significance of differences between experimental groups for *in vitro* experiments of cell lines treated with varying doses of SIKi or vehicle control were assessed by one-way ANOVA with Dunnett's multiple comparisons post-test. *In vitro* time course experiments were assessed by repeated-measures one-way ANOVA with Dunnett's multiple comparisons post-test.

Statistical significance for colorimeter readings in Figure 2B was determined by multiple t test analysis between day 0 and day 7 for each treatment group, with the two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli to correct for false discovery rate (FDR)—desired FDR (Q) = 1%—with no assumption of consistent SD. Statistical significance for colorimeter readings in Figure S3B were assessed by one-way ANOVA with Dunnett's multiple comparisons post-test.

For the G361 melanoma cells transduced with LKB1, a one-way ANOVA was used, with Dunnett's multiple comparisons test to assess the statistical significance of LKB1 expression and a two-tailed paired t test to assess the statistical significance of MITF induction with SIK-inhibitor treatment.

Statistical significance of differences between experimental groups for *in vivo* time course experiments was assessed by two-way ANOVAs with Sidak's multiple comparisons test.

Multiplicity-adjusted p values were reported for each comparison, and differences of means were considered significant if  $p < 0.05$ .

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and one data file and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.05.042>.

## AUTHOR CONTRIBUTIONS

N.M., Y.L., R.M., N.S.G., and D.E.F. contributed to conception and design of the study; N.M., Y.L., H.G.C., A.S.D., J.W., Y.S., Q.Y.W., J.A., L.V.K., A.L.H., and E.M.R. contributed to acquisition and analysis of data. N.M., D.E.F., and J.W. wrote the article.

## ACKNOWLEDGMENTS

The authors thank Nicholas Lowe, Xunwei Wu, Jie Wen, Yang Feng, and Vivien Igras for technical advice and assistance. This work was supported by grants from NIH (5P01 CA163222 and 5R01 AR043369-19), the Melanoma Research Alliance, the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation, and the Canadian Institutes of Health Research (DFS-140391).

N.M., R.M., N.S.G., and D.E.F. declare that parts of the work are subject of a U.S. provisional patent application titled “Pyrimidopyrimidinones as SIK Inhibitors.”

Received: March 16, 2017

Revised: May 2, 2017

Accepted: May 12, 2017

Published: June 13, 2017

## REFERENCES

- Armstrong, B.K., and Kricger, A. (2001). The epidemiology of UV induced skin cancer. *J. Photochem. Photobiol. B* 63, 8–18.
- Bertolotto, C., Abbe, P., Hemesath, T.J., Bille, K., Fisher, D.E., Ortonne, J.P., and Ballotti, R. (1998). Microphthalmia gene product as a signal transducer in cAMP-induced differentiation of melanocytes. *J. Cell Biol.* 142, 827–835.
- Bertolotto, C., Lesueur, F., Giuliano, S., Strub, T., de Lichy, M., Bille, K., Desse, P., d’Hayer, B., Mohamdi, H., Remenieras, A., et al.; French Familial Melanoma Study Group (2011). A SUMOylation-defective MITF germline mutation predisposes to melanoma and renal carcinoma. *Nature* 480, 94–98.
- Bos, J.D., and Meinardi, M.M. (2000). The 500 dalton rule for the skin penetration of chemical compounds and drugs. *Exp. Dermatol.* 9, 165–169.
- Choy, Y.B., and Prausnitz, M.R. (2011). The rule of five for non-oral routes of drug delivery: ophthalmic, inhalation and transdermal. *Pharm. Res.* 28, 943–948.
- Clark, K., MacKenzie, K.F., Petkevicius, K., Kristariyanto, Y., Zhang, J., Choi, H.G., Pegg, M., Plater, L., Pedrioli, P.G., Mclver, E., et al. (2012). Phosphorylation of CRTC3 by the salt-inducible kinases controls the interconversion of classically activated and regulatory macrophages. *Proc. Natl. Acad. Sci. USA* 109, 16986–16991.
- Cui, R., Widlund, H.R., Feige, E., Lin, J.Y., Wilensky, D.L., Igras, V.E., D’Orazio, J., Fung, C.Y., Schanbacher, C.F., Granter, S.R., and Fisher, D.E. (2007). Central role of p53 in the suntan response and pathologic hyperpigmentation. *Cell* 128, 853–864.
- D’Orazio, J.A., Nobuhisa, T., Cui, R., Arya, M., Spry, M., Wakamatsu, K., Igras, V., Kunisada, T., Granter, S.R., Nishimura, E.K., et al. (2006). Topical drug rescue strategy and skin protection based on the role of Mc1r in UV-induced tanning. *Nature* 443, 340–344.
- Dombret, H., Chastang, C., Fenaux, P., Reiffers, J., Bordessoule, D., Bouabdallah, R., Mandelli, F., Ferrant, A., Auzanneau, G., Tilly, H., et al.; AML Cooperative Study Group (1995). A controlled study of recombinant human granulocyte colony-stimulating factor in elderly patients after treatment for acute myelogenous leukemia. *N. Engl. J. Med.* 332, 1678–1683.
- Gandini, S., Sera, F., Cattaruzza, M.S., Pasquini, P., Picconi, O., Boyle, P., and Melchi, C.F. (2005). Meta-analysis of risk factors for cutaneous melanoma: II. Sun exposure. *Eur. J. Cancer* 41, 45–60.
- Garraway, L.A., Widlund, H.R., Rubin, M.A., Getz, G., Berger, A.J., Ramaswamy, S., Beroukhi, R., Milner, D.A., Granter, S.R., Du, J., et al. (2005). Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* 436, 117–122.
- Horike, N., Takemori, H., Katoh, Y., Doi, J., Min, L., Asano, T., Sun, X.J., Yamamoto, H., Kasayama, S., Muraoka, M., et al. (2003). Adipose-specific expression, phosphorylation of Ser794 in insulin receptor substrate-1, and activation in diabetic animals of salt-inducible kinase-2. *J. Biol. Chem.* 278, 18440–18447.
- Horike, N., Kumagai, A., Shimono, Y., Onishi, T., Itoh, Y., Sasaki, T., Kitagawa, K., Hatano, O., Takagi, H., Susumu, T., et al. (2010). Downregulation of SIK2 expression promotes the melanogenic program in mice. *Pigment Cell Melanoma Res.* 23, 809–819.
- Katoh, Y., Takemori, H., Lin, X.Z., Tamura, M., Muraoka, M., Satoh, T., Tsuchiya, Y., Min, L., Doi, J., Miyauchi, A., et al. (2006). Silencing the constitutive active transcription factor CREB by the LKB1-SIK signaling cascade. *FEBS J.* 273, 2730–2748.
- Kennedy, C., Bajdik, C.D., Willemze, R., De Gruij, F.R., and Bouwes Bavinck, J.N.; Leiden Skin Cancer Study (2003). The influence of painful sunburns and lifetime sun exposure on the risk of actinic keratoses, seborrheic warts, melanocytic nevi, atypical nevi, and skin cancer. *J. Invest. Dermatol.* 120, 1087–1093.
- Khaled, M., Levy, C., and Fisher, D.E. (2010). Control of melanocyte differentiation by a MITF-PDE4D3 homeostatic circuit. *Genes Dev.* 24, 2276–2281.
- Kobayashi, N., Nakagawa, A., Muramatsu, T., Yamashina, Y., Shirai, T., Hashimoto, M.W., Ishigaki, Y., Ohnishi, T., and Mori, T. (1998). Supranuclear melanin caps reduce ultraviolet induced DNA photoproducts in human epidermis. *J. Invest. Dermatol.* 110, 806–810.
- Kunisada, T., Lu, S.Z., Yoshida, H., Nishikawa, S., Nishikawa, S., Mizoguchi, M., Hayashi, S., Tyrrell, L., Williams, D.A., Wang, X., and Longley, B.J. (1998). Murine cutaneous mastocytosis and epidermal melanocytosis induced by keratinocyte expression of transgenic stem cell factor. *J. Exp. Med.* 187, 1565–1573.
- Langendonk, J.G., Balwani, M., Anderson, K.E., Bonkovsky, H.L., Anstey, A.V., Bissell, D.M., Bloomer, J., Edwards, C., Neumann, N.J., Parker, C., et al. (2015). Afamelanotide for erythropoietic protoporphyria. *N. Engl. J. Med.* 373, 48–59.
- Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 46, 3–26.
- Lönn, P., Vanlandewijck, M., Raja, E., Kowanetz, M., Watanabe, Y., Kowanetz, K., Vasilaki, E., Heldin, C.H., and Moustakas, A. (2012). Transcriptional induction of salt-inducible kinase 1 by transforming growth factor  $\beta$  leads to negative regulation of type I receptor signaling in cooperation with the Smurf2 ubiquitin ligase. *J. Biol. Chem.* 287, 12867–12878.
- Miller, A.J., Du, J., Rowan, S., Hershey, C.L., Widlund, H.R., and Fisher, D.E. (2004). Transcriptional regulation of the melanoma prognostic marker melastatin (TRPM1) by MITF in melanocytes and melanoma. *Cancer Res.* 64, 509–516.
- Newton, R.A., Smit, S.E., Barnes, C.C., Pedley, J., Parsons, P.G., and Sturm, R.A. (2005). Activation of the cAMP pathway by variant human MC1R alleles expressed in HEK and in melanoma cells. *Peptides* 26, 1818–1824.
- Ohno, R., Tomonaga, M., Kobayashi, T., Kanamaru, A., Shirakawa, S., Matsuoka, T., Omine, M., Oh, H., Nomura, T., Sakai, Y., et al. (1990). Effect of granulocyte colony-stimulating factor after intensive induction therapy in relapsed or refractory acute leukemia. *N. Engl. J. Med.* 323, 871–877.
- Park, S.B., Suh, D.H., and Youn, J.I. (1999). A long-term time course of colorimetric evaluation of ultraviolet light-induced skin reactions. *Clin. Exp. Dermatol.* 24, 315–320.
- Pennello, G., Devesa, S., and Gail, M. (2000). Association of surface ultraviolet B radiation levels with melanoma and nonmelanoma skin cancer in United States blacks. *Cancer Epidemiol. Biomarkers Prev.* 9, 291–297.
- Price, E.R., Horstmann, M.A., Wells, A.G., Weibaecher, K.N., Takemoto, C.M., Landis, M.W., and Fisher, D.E. (1998). Alpha-melanocyte-stimulating hormone signaling regulates expression of microphthalmia, a gene deficient in Waardenburg syndrome. *J. Biol. Chem.* 273, 33042–33047.
- Rogers, H.W., Weinstock, M.A., Feldman, S.R., and Coldiron, B.M. (2015). Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. population, 2012. *JAMA Dermatol.* 151, 1081–1086.
- Ryerson, A.B., Ehemann, C.R., Altekruze, S.F., Ward, J.W., Jemal, A., Sherman, R.L., Henley, S.J., Holtzman, D., Lake, A., Noone, A.M., et al. (2016). Annual report to the nation on the status of cancer, 1975–2012, featuring the increasing incidence of liver cancer. *Cancer* 122, 1312–1337.
- Takemori, H., Katoh, Y., Horike, N., Doi, J., and Okamoto, M. (2002). ACTH-induced nucleocytoplasmic translocation of salt-inducible kinase. Implication in the protein kinase A-activated gene transcription in mouse adrenocortical tumor cells. *J. Biol. Chem.* 277, 42334–42343.
- Tsatmalia, M., Wakamatsu, K., Graham, A.J., and Thody, A.J. (1999). Skin POMC peptides. Their binding affinities and activation of the human MC1 receptor. *Ann. N Y Acad. Sci.* 885, 466–469.

- Valverde, P., Healy, E., Jackson, I., Rees, J.L., and Thody, A.J. (1995). Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nat. Genet.* *11*, 328–330.
- Wakamatsu, K., and Ito, S. (2002). Advanced chemical methods in melanin determination. *Pigment Cell Res.* *15*, 174–183.
- Watson, M., Geller, A.C., Tucker, M.A., Guy, G.P., Jr., and Weinstock, M.A. (2016). Melanoma burden and recent trends among non-Hispanic whites aged 15–49 years, United States. *Prev. Med.* *91*, 294–298.
- Wu, S., Han, J., Vleugels, R.A., Puett, R., Laden, F., Hunter, D.J., and Qureshi, A.A. (2014). Cumulative ultraviolet radiation flux in adulthood and risk of incident skin cancers in women. *Br. J. Cancer* *110*, 1855–1861.
- Yokoyama, S., Woods, S.L., Boyle, G.M., Aoude, L.G., MacGregor, S., Zismann, V., Gartside, M., Cust, A.E., Haq, R., Harland, M., et al. (2011). A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. *Nature* *480*, 99–103.

# Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells

Yiyuan Liu,<sup>1,5</sup> Bobbie Pelham-Webb,<sup>1,2,5</sup> Dafne Campigli Di Giammartino,<sup>1,5</sup> Jiexi Li,<sup>1</sup> Daleum Kim,<sup>1</sup> Katsuhiko Kita,<sup>1</sup> Nestor Saiz,<sup>3</sup> Vidur Garg,<sup>3</sup> Ashley Doane,<sup>4</sup> Paraskevi Giannakakou,<sup>1</sup> Anna-Katerina Hadjantonakis,<sup>3</sup> Olivier Elemento,<sup>4</sup> and Effie Apostolou<sup>1,6,\*</sup>

<sup>1</sup>Joan & Sanford I. Weill Department of Medicine, Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, NY 10021, USA

<sup>2</sup>Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program, New York, NY 10021, USA

<sup>3</sup>Developmental Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

<sup>4</sup>Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY 10021, USA

<sup>5</sup>These authors contributed equally

<sup>6</sup>Lead Contact

\*Correspondence: efa2001@med.cornell.edu

<http://dx.doi.org/10.1016/j.celrep.2017.04.067>

## SUMMARY

During mitosis, transcription is halted and many chromatin features are lost, posing a challenge for the continuity of cell identity, particularly in fast cycling stem cells, which constantly balance self-renewal with differentiation. Here we show that, in pluripotent stem cells, certain histone marks and stem cell regulators remain associated with specific genomic regions of mitotic chromatin, a phenomenon known as mitotic bookmarking. Enhancers of stem cell-related genes are bookmarked by both H3K27ac and the master regulators OCT4, SOX2, and KLF4, while promoters of housekeeping genes retain high levels of mitotic H3K27ac in a cell-type invariant manner. Temporal degradation of OCT4 during mitotic exit compromises its ability both to maintain and induce pluripotency, suggesting that its regulatory function partly depends on its bookmarking activity. Together, our data document a widespread yet specific bookmarking by histone modifications and transcription factors promoting faithful and efficient propagation of stemness after cell division.

## INTRODUCTION

Cell identity is determined by characteristic gene expression programs and chromatin landscapes, which are constantly supervised by key transcription factors (TFs), known as master regulators (Natoli, 2010; Young, 2011). This tightly controlled system is temporarily destabilized during mitosis, when dramatic molecular changes occur, such as global transcriptional shut down, alterations in the abundance of histone modifications, and dissociation of most of the transcription factors and cofactors from condensed mitotic chromatin (Martínez-Balbás et al., 1995; Gottesfeld and Forbes, 1997; Wang and Higgins, 2013). How cell-type-specific programs are faithfully restored in daughter cells

constitutes a fundamental—and yet unanswered—question in biology. Studies in different somatic cell types proposed that the heritability of defined gene expression programs might rely on the mitotic persistence of either epigenetic marks (histone and DNA modifications) and/or TFs on the chromatin, a phenomenon known as mitotic bookmarking. There are examples supporting at least partial retention of specific active or repressive histone marks, which then enables rapid recruitment of the respective histone readers and writers upon G1 entry, ensuring self-perpetuation of gene expression states (Kouskouti and Talianidis, 2005; Margueron and Reinberg, 2010; Valls et al., 2005; Zaidi et al., 2010). There are also a number of studies documenting that selected TFs remain associated with mitotic chromatin, facilitating rapid reactivation of critical genes for the respective cell identity (Caravaca et al., 2013; Kadauke and Blobel, 2013; Kadauke et al., 2012; Young et al., 2007). Interestingly, previously reported bookmarking TFs are either master regulators of cell identity, such as GATA1 in hematopoietic cells (Kadauke et al., 2012), and/or “pioneer” factors—able to bind otherwise “inaccessible” nucleosomal regions (Zaret and Carroll, 2011), such as FoxA1 in liver progenitors (Caravaca et al., 2013). Collectively, these suggest that both epigenetic and TF bookmarking mechanisms contribute to the faithful propagation of cell identity after cell division. Whether similar mechanisms are also important for stem cell fate heritability remains understudied.

Pluripotent stem cells (PSCs) are endowed with the remarkable capacity to self-renew indefinitely, while preserving the potential to differentiate into all somatic cell types in response to developmental cues (Evans, 2011; Tabar and Studer, 2014). In addition, PSCs are characterized by an extremely rapid (10–12 hr) cell cycle, which lacks G0 phase and has an unusually short G1 phase (Coronado et al., 2013; Savatier et al., 2002). This necessitates the presence of very efficient mechanisms for preserving or resetting PSC-specific transcriptional patterns. Notably, there is a well-defined network of TFs that control maintenance and acquisition of stem cell identity (Apostolou and Hochedlinger, 2013; Takahashi and Yamanaka, 2006; Young, 2011). Among them, OCT4, SOX2, ESRRB, and KLF4 have been reported to also function as pioneer factors (Iwafuchi-Doi and Zaret, 2014; Soufi et al., 2012; Soufi et al., 2015), sharing critical

properties with known bookmarking TFs. However, evidence for a potential bookmarking role for some of these TFs has only recently begun to emerge (Festuccia et al., 2016; Teves et al., 2016; Deluz et al., 2016). Similarly, although the functional genomic elements that control stem cell identity and the histone marks that decorate them are well-characterized (Hawkins et al., 2010; Mikkelsen et al., 2007; Whyte et al., 2013), their status during mitosis remains unknown.

Here, combining biochemical and single-cell imaging approaches, we systematically investigated the extent of mitotic retention for selected histone modifications and pluripotency-associated TFs. Our analyses identified in a genome-wide scale the nature and specificity of bookmarking and provided evidence for its biological significance in stem cell identity.

## RESULTS

### Specific Histone Modifications and Pluripotency-Related Transcription Factors Are Maintained on Mitotic Chromatin

To determine candidate bookmarking factors in PSCs, we first tested a number of histone modifications and TFs for their relative ability to remain associated with mitotic chromatin of mouse embryonic stem cells (ESCs). To do so, we performed subcellular fractionation of both asynchronous ESCs and ESCs arrested in mitosis to a high purity (typically >92%), as assayed by fluorescence-activated cell sorting (FACS) analysis for the presence of the mitotic marker histone 3 phospho-serine 10 (H3Ser10p) (Tapi et al., 2006) (Figure 1A). Western blot analysis validated the high purity of the separate fractions, exhibiting complete depletion of cytoplasmic proteins, such as tubulin and GAPDH, from nuclear and chromatin extracts and detection of histone 3 predominantly in the chromatin fraction (Figure S1A). When we focused on the chromatin-bound fractions, we found that architectural factors, such as SMC3 and CTCF, remained associated with chromatin during mitosis, in agreement with previous studies (Burke et al., 2005; Peters et al., 2008), whereas RNA polymerase II phospho-Ser2 (PolII marker for transcriptional elongation) and Mediator (Med12) levels were markedly decreased in accordance with a global transcriptional silencing during mitosis (Gottesfeld and Forbes, 1997) (Figures 1B and 1E). These results demonstrated that our chromatin extracts had the required purity and quality for downstream analyses.

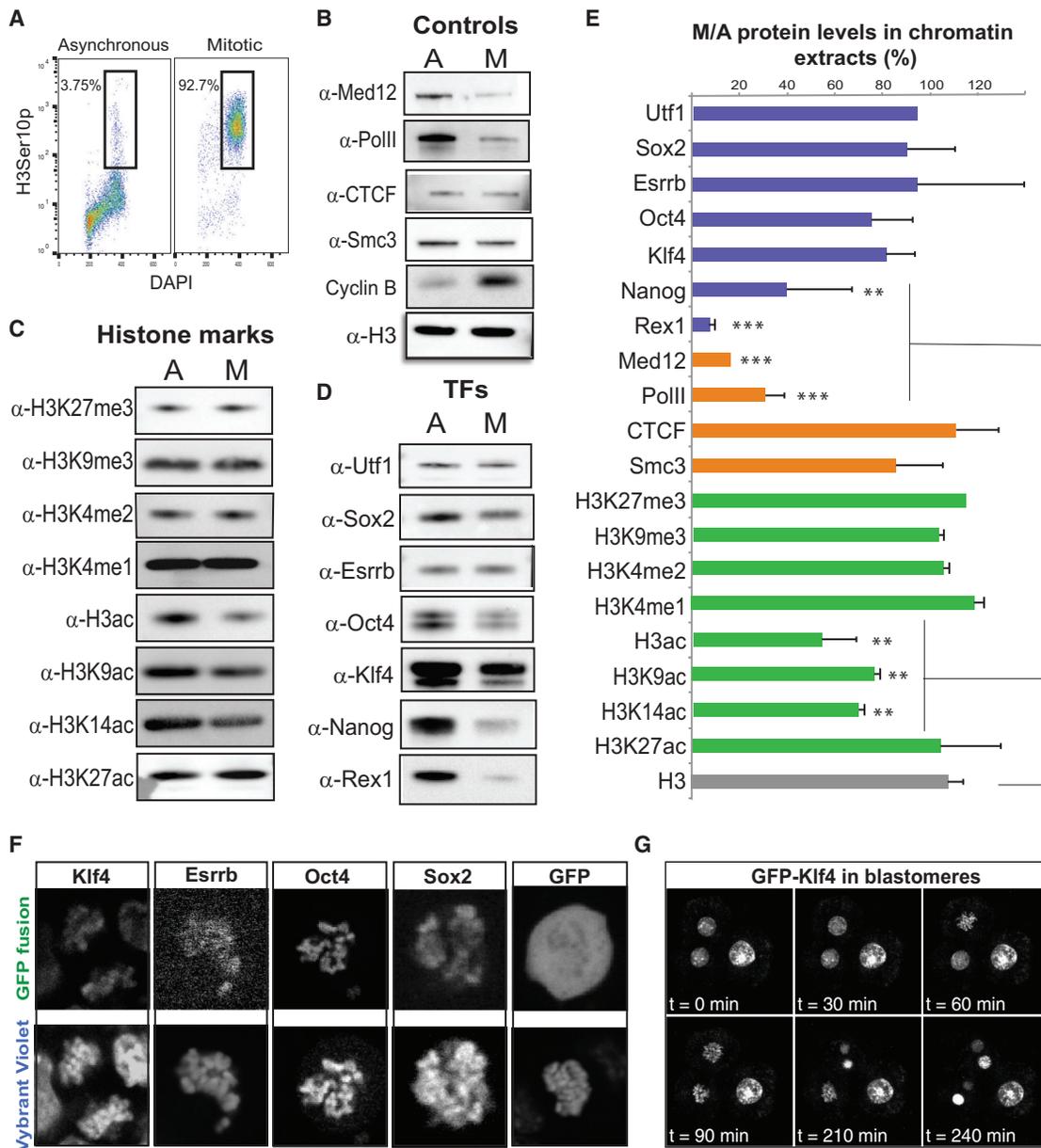
We then tested the relative abundance of common histone modifications on asynchronous and mitotic chromatin. In agreement with previous findings in other cell types (Valls et al., 2005; Wang and Higgins, 2013), all tested histone methylation marks (H3K27me3, H3K9me3, H3K4me1, and H3K4me2) were highly retained during mitosis (Figures 1C and 1E). In contrast, the level of acetylated H3 was decreased, partly due to the reduction of H3K9 and H3K14 acetylation. Interestingly, H3K27 acetylation, which is a hallmark of active regulatory elements and particularly enhancers (Creighton et al., 2010), was detected at similar levels in asynchronous and mitotic ESCs. These results show that—at the global level—some of the key regulatory histone modifications are maintained on mitotic chromatin of ESCs, suggesting that inheritance of these chromatin marks may play a role in faithful propagation of pluripotent cell identity upon cell division.

We also tested the association of known pluripotency-related TFs with mitotic chromatin. UTF1, a chromatin-associated factor with histone-like characteristics (Kooistra et al., 2009; van den Boom et al., 2007), served as a positive control showing almost identical levels in asynchronous and mitotic chromatin extracts (Figures 1D and 1E). In agreement with recent studies (Festuccia et al., 2016; Teves et al., 2016; Deluz et al., 2016), this assay showed that SOX2, ESRRB, and OCT4 as well as KLF4 were also highly retained on mitotic chromatin (ranging between 70%–100% of respective levels in asynchronous cells). By contrast, factors such as NANOG and REX1 showed markedly reduced levels in the chromatin fraction during mitosis (Figures 1D and 1E). Of note, the protein levels of several TFs including NANOG were reduced in total mitotic cell extracts (Figure S1B), suggesting cell-cycle-dependent regulation of protein abundance. However, a clear correlation between total protein levels and chromatin retention was not observed, implying that additional mechanisms such as active dissociation from or preferential association with mitotic chromatin may therefore operate in a TF-specific manner. Taken together, these results reveal that many pluripotency-related TFs remain associated with mitotic chromatin in ESCs.

To independently verify our findings at the single-cell level, we performed immunofluorescence experiments on metaphase spreads of partly synchronized ESCs after fixation with methanol:acetic acid solution (Figures S1C and S1D). OCT4, SOX2, KLF4, and ESRRB were detected both on interphase nuclei and on the condensed chromosomes. In contrast, RNA PolII and NANOG were excluded from the mitotic chromosomes, in concordance with our western blot analyses. Of note, incubation with matched IgG isotypes or only secondary antibodies showed no specific signal on either interphase nuclei or condensed chromosomes. In parallel, live imaging of ESCs overexpressing GFP fusions with OCT4, SOX2, KLF4, or ESRRB (Figure S1E) confirmed a strong enrichment for each of these TFs on mitotic chromosomes, whereas overexpressed GFP protein alone was diffused throughout the cell volume (Figure 1F). Finally, introduction of a KLF4-GFP chimeric protein into 2-cell mouse embryos followed by live imaging demonstrated *in vivo* the ability of this protein to associate with mitotic chromosomes in dividing blastomeres (Figure 1G; Movie S1). Together, these results confirmed at a single-cell level the ability of critical stem cell regulators to remain associated with chromatin during cell division, suggesting a potential bookmarking function.

### H3K27 Acetylation Bookmarks Distinct Sets of Regulatory Elements during Mitosis

Next, we sought to determine in a genome-wide scale the extent and specificity of mitotic retention of H3K27ac, which constitutes an attractive candidate bookmarking feature given its association to active regulatory elements. ChIP-seq experiments in asynchronous and mitotically-arrested ESCs revealed highly overlapping, but distinct, genome-wide occupancy patterns (Figures 2A–2C). There was a large portion of H3K27ac sites that were retained in mitotic cells (bookmarked, B class), whereas distinct sets of H3K27ac peaks were detected only in asynchronous (A class) or only in mitotic cells (M class) (Figure 2C). Examples of ChIP-seq tracks for each category are shown in Figure 2D.



**Figure 1. Histone Modifications and Pluripotency-Related Transcription Factors Are Largely Retained on the Mitotic Chromatin**

(A) Representative FACS plots showing the percentage of asynchronous or nocodazole-treated ESCs (after mitotic shake-off) expressing the mitotic marker H3Ser10p.

(B–D) Western blot analyses showing the relative levels of positive and negative controls (B), selected histone modifications (C), and transcription factors (D) on the chromatin fraction of asynchronous (A) or mitotic (M) ESCs.

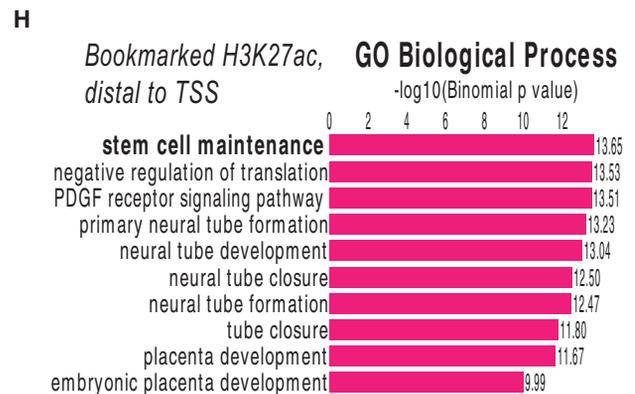
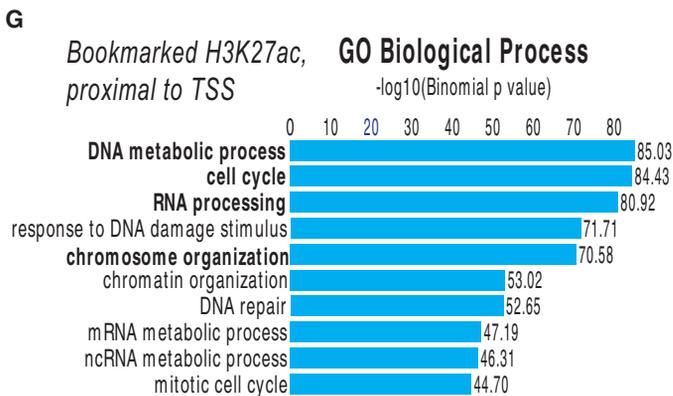
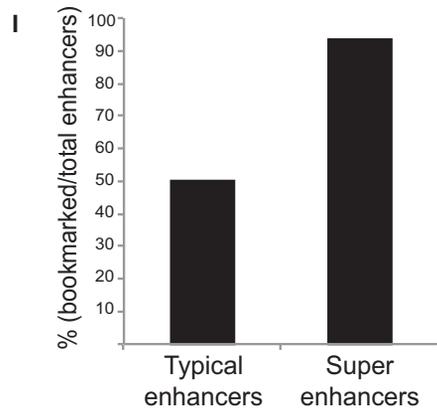
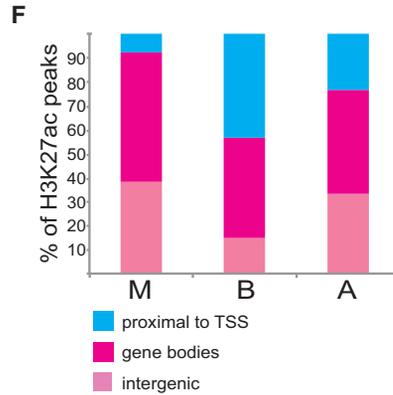
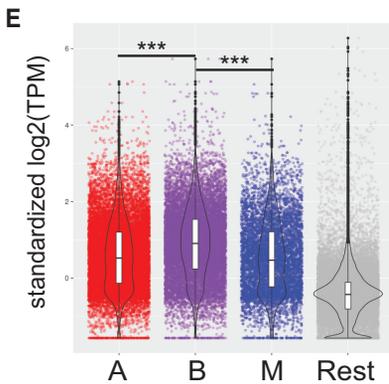
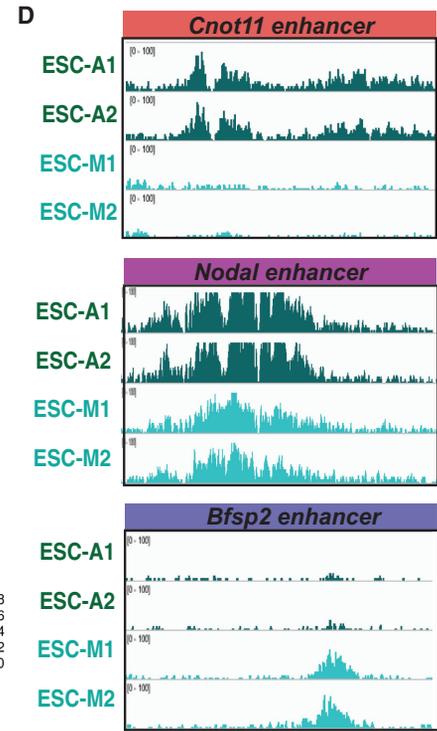
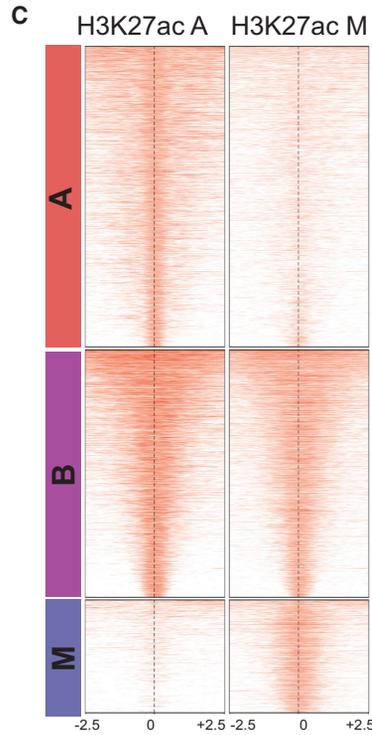
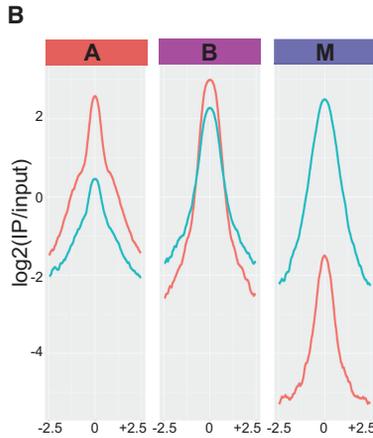
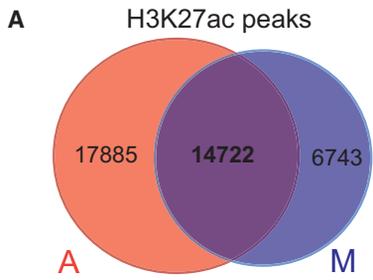
(E) Quantitation of (B)–(D) gel bands using ImageJ. The relative levels of each protein in mitotic chromatin fractions are plotted as a percentage of the respective levels in asynchronous chromatin fractions. Error bars indicate SD based on at least two independent synchronization and fractionation experiments in two different ESC lines (ESC V6.5 and ZHBTc4.1). Asterisks indicate significant downregulation when compared to H3 levels, as calculated by t test (\*\*p < 0.01, \*\*\*p < 0.001).

(F) Representative live-imaging photos of mitotic ESCs expressing ectopic KLF4, ESRRB, OCT4, and SOX2 fused with a GFP reporter. Overexpression of GFP alone is used as a control. The chromosomes were stained using a cell-permeable DNA dye (Vybrant Violet).

(G) Time-lapse images of dividing blastomeres after microinjection of KLF4-GFP mRNA into one out of two blastomeres in 2-cell stage embryos. See also Movie S1.

Correlation of our H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq) data with published RNA sequencing (RNA-seq) data from asynchronous ESCs showed that genes proximal to B-class H3K27ac peaks displayed stronger transcrip-

tional activity compared to the ones that lose or gain H3K27ac during mitosis (Figure 2E). When we analyzed the distribution of H3K27ac peaks that belong in each category across the genome, we noticed a clear overrepresentation of promoter-proximal



(legend on next page)

peaks among B targets, while A and M targets were mostly located >2.5 kb away from transcriptional start sites (TSS), either within gene bodies or in intergenic regions (Figure 2F). Gene ontology analyses using Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010) on the B-class H3K27ac peaks showed enrichment for distinct functional annotations depending on their proximity to the TSS. Specifically, TSS-proximal peaks strongly enriched for genes involved in fundamental cellular processes, such as cell-cycle progression and RNA and DNA metabolism (Figure 2G). On the other hand, distal B-class H3K27ac peaks (>2.5 kb from TSS) were predominantly linked to stem cell maintenance-related genes (Figure 2H), including *Nanog*, *Pou5f1*, and *Sox2*. In fact, we observed that more than 50% of typical ESC-enhancers and more than 90% of the so-called super-enhancers (Whyte et al., 2013), which are linked to genes critical for stem cell identity, remained bookmarked by H3K27ac during mitosis (Figure 2I). Together, these results indicate that H3K27 acetylation during mitosis marks two distinct sets of genomic regions: (1) enhancers of stem cell-associated genes, and (2) promoters of cell-cycle-regulating genes.

To test whether the observed behavior of H3K27 acetylation during mitosis was characteristic of ESCs, we repeated the above analysis using published H3K27ac ChIP-seq dataset from G1E erythroblasts (Hsiung et al., 2016). G1E erythroblasts showed a similar degree of overlap between mitotic and asynchronous H3K27ac peaks as well as a conserved pattern of genomic distribution, with B-class peaks again predominantly located around promoters (not shown). Comparison of the ESC and G1E datasets revealed that a large proportion of bookmarked H3K27ac peaks were common among the different cell types (Figure S2A). GREAT analysis of the common bookmarked peaks showed a strong enrichment around promoters of genes that were mostly involved in cell cycle, RNA processing and ribosome and chromosome organization (Figures S2B and S2C). In contrast, ESC-specific and G1E-specific bookmarked H3K27ac peaks were mostly located away from promoters and were strongly associated with genes characteristic for each specific cell identity (Figures S2B and S2C). Examples are shown in Figure S2D. In agreement with these results, motif scanning analyses using either the ESC or the G1E bookmarked H3K27ac enhancer peaks enriched for binding sites recognized by the master regulators of each cell type, such as OSN (OCT4-SOX2-NANOG) and ESRRB or GATA1, respectively (Figure S2E; Table S2). On the other

hand, bookmarked promoter H3K27ac peaks in both cell types enriched for motifs for common general transcription factors involved in cell-cycle regulation, including MeCP2 and E2F3. Taken together, these results propose that H3K27 acetylation is a previously unappreciated mitotic mark with dual bookmarking behavior on both promoters of cell cycle and homeostasis-related genes and enhancers of genes important for cell identity.

### KLF4, SOX2, and OCT4 Bookmark Critical Stem Cell Regulatory Elements during Mitosis

Previous studies have shown that bookmarking factors may associate with mitotic chromatin in a non-specific manner, coating the compacted chromosomes, or by specific binding at selected genomic sites (Caravaca et al., 2013; Kadauke and Blobel, 2013). To distinguish between these possibilities, we investigated the genome-wide binding patterns of the critical pluripotency regulators KLF4 (K), OCT4 (O), and SOX2 (S) (Apostolou and Hochedlinger, 2013; Young, 2011) in mitotic and asynchronous ESCs. Unbiased clustering of all ChIP-seq samples generated two distinct groups based on the TF they were targeting (KLF4 group and OCT4/SOX2 group), but failed to clearly separate mitotic from asynchronous samples, highlighting their global similarities (Figure S3A). Stringent differential analyses of the asynchronous and mitotic binding sites for each TF revealed a large number of common genomic targets showing that 25%–60% of the asynchronous KOS peaks were maintained during mitosis (bookmarked, B), whereas the rest were markedly decreased (asynchronous-only, A) (Figures 3A). Examples of bookmarked or non-bookmarked genomic regions are shown in Figure 3B. Of note, genomic regions that were found enriched in mitotic samples (mitotic-only, M) had overall lower and more variable signal among replicates and thus were excluded from downstream analyses.

In addition to the large number of KOS bookmarked genomic sites, the strength of mitotic retention (expressed as normalized ChIP-seq signal) in most of these sites was comparable to the respective signal in asynchronous ESCs (Figures 3B and S3B). This finding was surprising given that previously reported bookmarked targets by various TFs, including the recently described ESRRB and SOX2 in ESCs (Festuccia et al., 2016; Deluz et al., 2016), were often characterized by weaker binding during mitosis, suggesting either technical or biological differences between studies. To test the possibility that the strength of the detected signal in our mitotic samples was due to contamination from

### Figure 2. ChIP-Seq Experiments Reveal Distinct but Overlapping Patterns of H3K27 Acetylation on Asynchronous and Mitotic ESCs

(A) Numbers of H3K27ac ChIP-seq peaks in asynchronous and mitotic ESCs and extent of overlap between them.

(B and C) Averaged coverage (B) and enrichment (C) plots of H3K27ac ChIP-seq signals in mitotic and asynchronous ESCs. Bookmarked (B) H3K27ac peaks have comparable signal in both conditions, whereas (A) are preferentially enriched in asynchronous and (M) in mitotic cells. ChIP-seq signals 2.5 kb upstream/downstream of peak centers are shown.

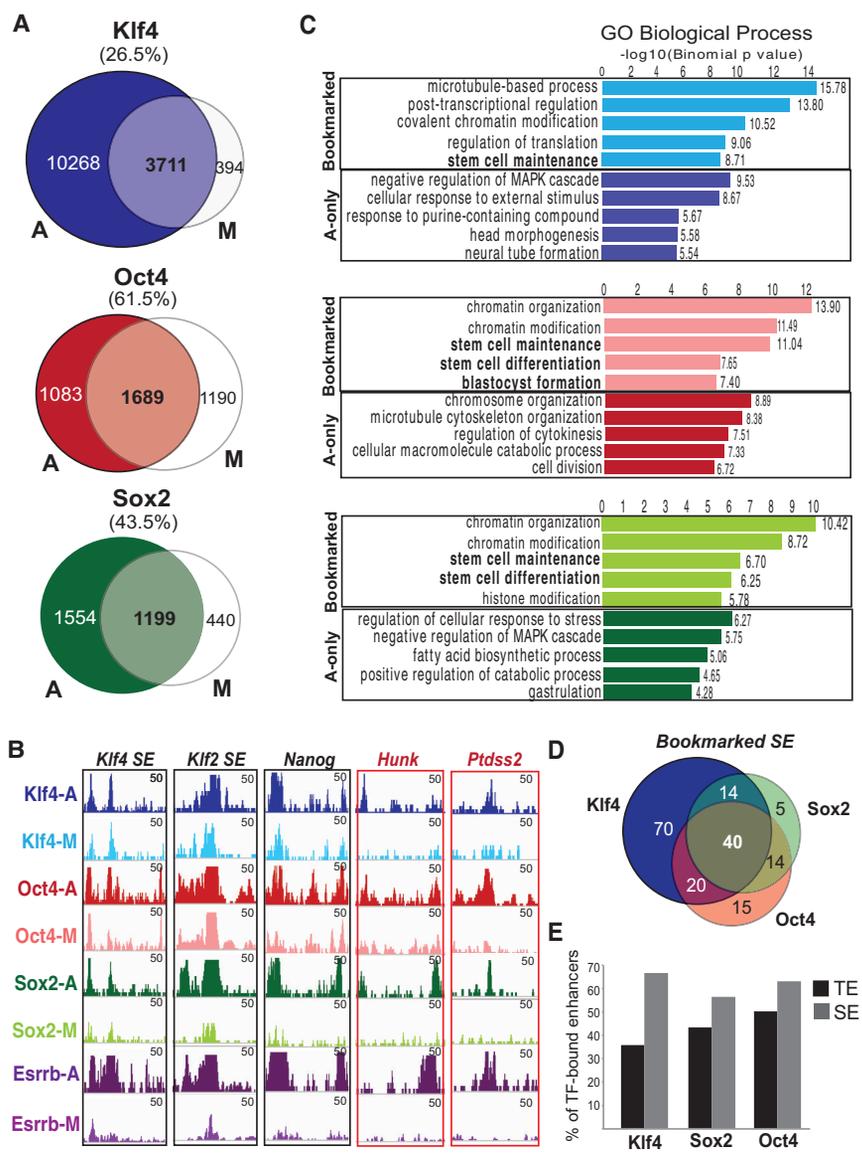
(D) Examples of H3K27ac tracks that belong in each of the A, B, and M categories.

(E) mRNA levels of genes proximal to H3K27ac peaks that belong to each of the A, B, or M categories. Published RNA-seq data from mouse ESCs (Shen et al., 2012) were used and standardized transcripts per million (TPM) were plotted. The most proximal gene to each peak was considered. Genes that were not assigned to any of the ChIP-seq peaks are shown as “Rest.” The expression levels of genes proximal to bookmarked H3K27ac sites were significantly higher compared to the other categories ( $p < 0.001$ ).

(F) Genomic distribution of each category of H3K27ac peaks relative to genes. Genome was partitioned into proximal to TSS (including TSS and 1 kb flanking regions), gene bodies (excluding regions proximal to TSS), and intergenic regions (rest of the genome).

(G and H) Top 10 Gene Ontology (GO) annotations enriched in bookmarked H3K27ac peaks that were either (G) proximal to TSS (<2.5 kb) or (H) located >2.5 kb away from any TSS.

(I) Barplot showing the percentage of typical enhancers and super-enhancers that retain H3K27 acetylation during mitosis (Whyte et al., 2013).



**Figure 3. ChIP-Seq Assays in Mitotic and Asynchronous ESCs Identify the Bookmarked Genomic Targets of KLF4, OCT4, and SOX2**

(A) Venn diagrams showing numbers of unique or common KLF4, OCT4, and SOX2 (KOS)-binding sites in asynchronous (A) and mitotic (M) mESCs. ChIP-seq peaks consistently detected in at least two of three biological replicates were used for the analyses (see the Supplemental Information). The percent of asynchronous peaks that were also detected in mitotic cells (bookmarked) are reported in parentheses.

(B) Examples of genomic regions that retained (black outline) or lost (red outline) KOS binding during mitosis. Note that the extent of KOS retention was similar or higher compared to the previously reported bookmarking factor ESRRB (published datasets from Festuccia et al. [2016]).

(C) Top five gene ontology categories (GO) enriched in KLF4 (blue), OCT4 (red), and SOX2 (green) binding sites that were either common between A and M (bookmarked) or present only in A (A-only). GOs associated with stem cell identity (bold) were overrepresented in bookmarked peaks.

(D) Venn diagram depicting the number of ESC-specific super-enhancers (SE) (Whyte et al., 2013) that remained bookmarked by individual TFs and combinations, highlighting the high frequency of combinatorial bookmarking.

(E) Barplot showing the percentage of KLF4, SOX2, or OCT4-bound typical enhancers (TE) and SE that remained bookmarked during mitosis.

residual G2 cells (not more than 10% in our mitotic samples), we mixed interphase ESCs and neuronal progenitor cells (NPCs), which do not express KLF4, in a ratio of 10:90 and performed KLF4 ChIP-seq and ChIP-qPCR. Figures S3B and S3C confirmed that the signal from mitotic samples around bookmarked binding sites was much higher than the one from the 10% Interphase sample (ESC:NPC mixture) and was in fact similar to the respective signal in asynchronous ESCs. These results validated that KOS remain strongly bound on selected genomic sites during mitosis.

To investigate whether lost (A) or bookmarked (B) binding sites were linked to genes with distinct biological functions, we performed GREAT analysis. Genomic sites from which the TFs dissociated during mitosis mostly enriched for signaling, cell cycle, and differentiation-associated genes. In contrast, genomic regions bookmarked by these TFs were proximal to genes involved in chromatin modification as well as stem cell maintenance and function (Figure 3C). Among them, 40%–50% of typical stem

TFs, their specific DNA-binding motif was present in 90%–95% of their respective bookmarked sites, suggesting that mitotic retention requires specific and direct binding (Table S3). Of note, this percentage was only slightly lower (82%–90%) among the rest of the target sites, indicating that solely the presence of the specific motif is not predictive of mitotic binding. In addition to OCT4, SOX2, and KLF4 motifs, a large number of other DNA sequence motifs scored as highly significant over background among bookmarked sites. After filtering out TFs that are minimally or not expressed in ESCs, we found a number of potentially relevant TF motifs that were preferentially enriched among the KLF4 bookmarked targets, such as sequences for E2F7, KLF3, and TFDP1 binding (Figure S3D). A distinct set of motifs was overrepresented among OCT4 and SOX2 bookmarked binding sites, including the composite OCT4/SOX2/NANOG (OSN) motif. Interestingly, all targets bookmarked by one or more of KOS had a higher frequency of ESRRB motif compared to the

asynchronous- or mitotic-only binding sites. These results suggest that KOS binding during mitosis may be dependent on or facilitated by additional TFs. Comparison of our results with published ChIP-seq datasets demonstrated that the frequency of bookmarked target sites for OCT4 and SOX2 that coincide with binding of two or more additional TFs (ESRRB, NANOG, SOX2, and OCT4) was dramatically higher compared to the A-only sites (Figure S3E), further suggesting that TF synergy increases the likelihood and/or strength of binding during mitosis.

Finally, to gain insights into the interconnection between TFs and H3K27ac on mitotic bookmarking, we performed enrichment analysis of the observed-over-expected overlap of bookmarked genomic regions (Figure S3F). This analysis revealed that H3K27ac-bookmarked enhancers preferentially enriched for cooperative mitotic binding by combinations of TFs, such as KOS and KS, whereas H3K27ac-bookmarked promoters preferentially enriched for KLF4 mitotic retention. These results provide evidence that distinct regulatory regions remain bookmarked by unique combinations of TFs and/or H3K27ac.

### Temporal Degradation of OCT4 during Mitotic Exit Compromises Its Ability to Maintain and Induce Pluripotency

Our findings support that ESC master regulators are largely retained on selective genomic regions during mitosis, suggesting a potential bookmarking function important for stem cell identity. To test this hypothesis, we focused on the master regulator OCT4 and established a system that has been previously shown to enable conditional protein degradation during mitotic exit (Kadauke et al., 2012). Specifically, we engineered a chimeric GFP-OCT4 protein (GFP-DB-OCT4 or GDO) that carries a destruction box (DB) of the mouse cyclin B (Glutzer et al., 1991; Holloway et al., 1993), which is targeted for degradation by the anaphase promoting complex (APC) during M-to-G1 transition (Figure 4A). As a control, we also generated a non-degradable version (GFP-DBmut-OCT4 or GMO) that carries a mutated DB sequence (R42A, DB<sup>mut</sup>), which has been shown to escape targeting by APC (Kadauke et al., 2012). By FACS analysis as well as live imaging, we confirmed that GMO-infected cells displayed stable GFP expression throughout the cell cycle, whereas in GDO-infected cells, GFP was rapidly lost during M-to-G1 transition, reappeared about 1 hr later, and gradually accumulated at later stages of the cell cycle (Figures S4A–S4C; Movies S2 and S3).

We then tested the relative ability of GDO and GMO chimeric proteins to preserve stemness when ectopically expressed in ZHBTc4.1 ESCs that rapidly silence endogenous OCT4 expression upon doxycycline treatment (Niwa et al., 2000). Western blot analyses of total protein extracts validated that doxycycline induced efficient depletion of endogenous OCT4, while expression of the exogenous GDO and GMO chimeric proteins were not affected (Figure S4D). In agreement with previous reports, silencing of endogenous OCT4 resulted in abrupt differentiation of uninfected ZHBTc4.1 cells, as shown by the loss of critical stem cell regulators such as NANOG, SOX2, and ESRRB (Figure S4D) and by their inability to give rise to alkaline phosphatase (AP)-positive colonies (Figure S4E). Although both GDO and GMO expression were able to ameliorate these phenotypes, GDO-infected ESCs were impaired in maintaining sufficient

mRNA and protein expression of stem cell regulators upon endogenous OCT4 depletion (Figures 4B, S4D, and S4F). In agreement, scoring of ESC colonies based either on their morphology or NANOG expression showed that GDO-expressing cells formed a significantly higher percentage of partially or completely differentiated colonies upon doxycycline treatment compared to the GMO-expressing cells (Figures 4C and 4D). Taken together, these results suggest that temporal degradation of OCT4 during mitotic exit compromises its ability to maintain pluripotency.

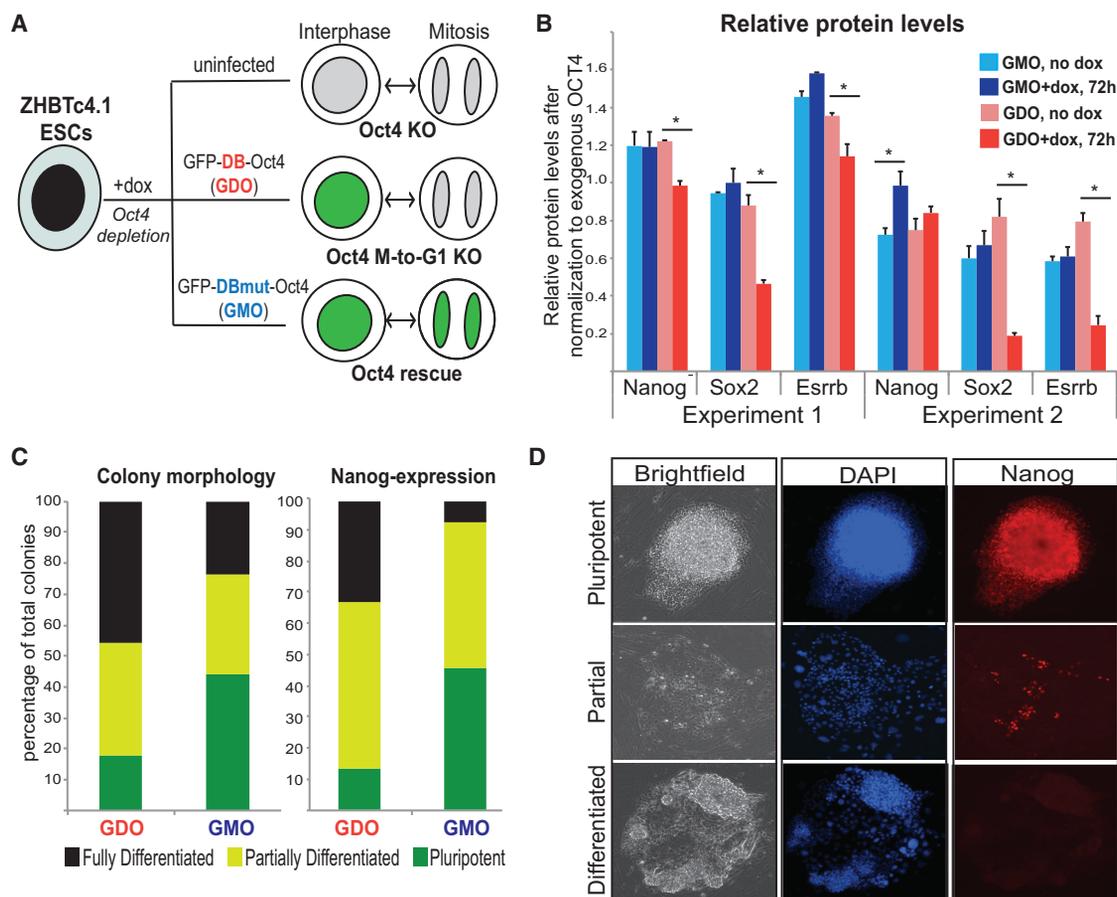
We next investigated the role of OCT4 mitotic bookmarking on reprogramming somatic cells to induced pluripotent stem cells (iPSCs) by overexpressing GDO or GMO along with wild-type SOX2, KLF4, and cMYC (Sommer et al., 2009) in mouse embryonic fibroblasts (MEFs) (Figures 5A and S5A). Cells expressing GDO protein were deficient in upregulating the early pluripotency marker SSEA-1 compared to the GMO-expressing ones (Figure S5B). Consistently, GDO-KSM reprogrammed MEFs generated substantially reduced numbers of iPSC colonies, as assessed by ESC-like morphology (Figure 5B), alkaline phosphatase staining, and NANOG expression (Figures S5C and S5D), compared to the ones expressing GMO or wild-type OCT4. Of note, immunofluorescence experiments on metaphase chromosome spreads from OKSM MEFs undergoing reprogramming revealed persistent binding of the ectopic OCT4, SOX2, and KLF4 factors on the condensed chromosomes and exclusion of RNA PolII (Figure 5C). Collectively, these results demonstrate the capacity of OCT4, SOX2, and KLF4 to bind on mitotic chromatin of both ESCs and somatic cells and provide evidence that bookmarking is also important for induction of pluripotency.

## DISCUSSION

Our study revealed a widespread yet specific retention of selected histone modifications and stem cell regulators on mitotic chromatin of ESCs, arguing that extensive chromatin and TF bookmarking might ensure faithful propagation of stem cell identity.

Among the highly retained histone modifications that we observed, H3K27ac was particularly intriguing, because histone acetylation levels were overall decreased during mitosis, in concordance with reduced transcriptional activity. Given that H3K27ac is strongly linked to enhancer activity and cell-type specificity, we considered it an attractive bookmarking candidate for propagation of cell identity. Interestingly, analysis of H3K27ac ChIP-seq data from ESCs as well as erythroblasts revealed a bookmarking preference for both cell-type-specific enhancers and housekeeping promoters. Whether the preferential mitotic retention of H3K27ac represents a residual mark of high transcriptional activity during interphase or an actual bookmarking mechanism that ensures rapid transcriptional reactivation of essential genes during mitotic exit remains to be determined. In support of the latter hypothesis, presence of H3K27ac in mitotic erythroblasts has been correlated with increased transcriptional activity during early G1 (Hsiung et al., 2016). However, more definitive studies will be required to interrogate the significance of this mark in transcriptional memory.

In addition to the H3K27ac-bookmarked targets, we also identified distinct subsets of genomic regions that lost or gained this



**Figure 4. Temporal Degradation of OCT4 during Mitotic Exit Compromises Its Ability to Maintain Pluripotency**

(A) Graphic illustration of our experimental strategy to test the effect of mitotic degradation of OCT4 on stem cell identity. The GFP-DB-OCT4 (GDO) chimeric protein carrying the destruction box (DB) of Cyclin B is degraded specifically during mitotic exit. In contrast, GFP-DBmut-OCT4 (GMO) carrying an R42A mutation (DBmut) remains unaffected. See also Figures S1A–S1C and Movies S2 and S3.

(B) Quantitation of western blot experiments (Figure S4C) testing protein levels of selected stem cell regulators in ZHBTc4.1 cells infected with GDO or GMO and either untreated or treated with doxycycline (dox). Gel bands quantitated by ImageJ were normalized to the uninfected no dox control for each protein and then normalized to the level of exogenous OCT4 for each condition. Error bars indicate SD of two technical replicates, and results from two independent experiments are shown. \*p value < 0.05, as calculated by t test.

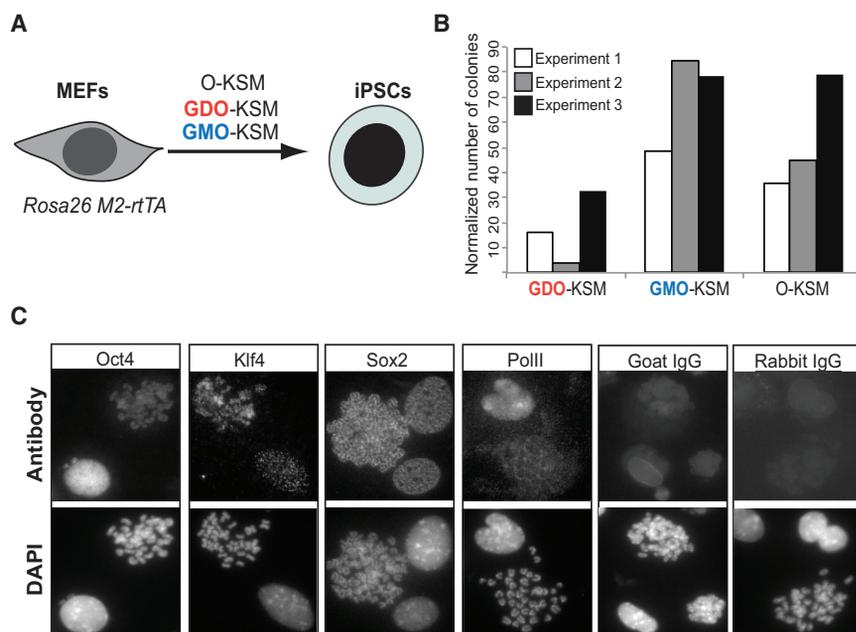
(C) Barplot showing the quantity and quality of colonies formed by ZHBTc4.1 ESCs expressing either GMO or GDO in the presence of doxycycline. GFP positive cells were sorted and plated on feeders and colonies were scored 7–10 days later based either on their morphology or NANOG expression. The results of three independent experiments are summarized; t test showed significantly (p value < 0.05) higher percentage of pluripotent-like colonies in the GMO expressing cells. Of note, uninfected ZHBTc4.1 fully differentiated in the presence of dox.

(D) Examples of pluripotent, partially differentiated, and fully differentiated colonies based on their morphology (Bright field) and the presence of NANOG-positive cells (red) as detected by immunofluorescence.

histone mark during mitosis, suggesting that mechanisms for deposition and removal of this histone modification may be at least partially active during mitosis or mitotic entry. HDACs have been reported to play crucial roles in cell-cycle progression through mitosis (Li et al., 2006; Cimini et al., 2003), resulting in decreased histone acetylation levels, and thus the loss of H3K27ac peaks was not unexpected. In contrast, the de novo acetylation of distinct genomic sites during mitosis is puzzling. A potential explanation would be that these mitotic-specific enhancer elements play a role in maintaining or inducing transcriptional activity of selected genes during mitosis, a possibility that could be addressed by nascent RNA sequencing experiments. Of note, although global transcriptional shut down during mitosis is well-

documented, an increasing number of studies report subsets of genes that “escape” this silencing (Halley-Stott et al., 2014; Liang et al., 2015; Liu et al., 2017; Sciortino et al., 2001).

Previous studies in somatic cells have reported that most TFs and cofactors dissociate from mitotic chromatin, except for a few bookmarking factors (Martínez-Balbás et al., 1995; Kadauke and Blobel, 2013). Here, using multiple independent experimental approaches, we report that a large number of stem cell regulators, including the recently described ESRRB and SOX2 (Festuccia et al., 2016; Deluz et al., 2016), were largely retained on mitotic chromatin of ESCs. Interestingly, in contrast with these studies, our data revealed a strong, widespread-yet-specific binding of KOS during mitosis. Biological and/or technical differences,



**Figure 5. Temporal Degradation of OCT4 during Mitotic Exit Compromises Its Ability to Induce Pluripotency**

(A) Experimental strategy to test the relative ability of GDO and GMO to induce pluripotency when overexpressed together with SOX2, KLF4, and cMYC in mouse embryonic fibroblasts (MEFs). Wild-type OCT4 was used as a positive control. (B) Quantitation of ESC-like colonies that were generated upon reprogramming with the respective construct. The number of colonies was normalized to the infection efficiency as described in the Supplemental Experimental Procedures. Three independent experiments are shown. (C) Representative immunofluorescence photos of metaphase chromosomes from MEFs expressing OKSM after doxycycline induction. OCT4, SOX2, and KLF4 were efficiently detected on the mitotic chromosomes, whereas RNA polymerase II (PolII) was excluded from the condensed chromosomes. Matched IgG isotypes (goat for KOS and rabbit for PolII) were used as negative controls.

such as synchronization and immunoprecipitation efficiencies, may explain these discrepancies. Notably, our findings are in agreement with a recently proposed model that TF retention on mitotic chromatin of ESCs may be a much broader phenomenon than previously appreciated (Teves et al., 2016). Whether this phenomenon is characteristic of pluripotent stem cell chromatin or is also present in other fast-cycling progenitor or even differentiated cells remains to be tested. Global proteomics assays in various cell types could enable determination of the relative chromatin composition of mitotic and interphase cells and give a better understanding of the key players and principles that govern this time window.

Importantly, despite the widespread mitotic retention we observed, specific TFs, such as NANOG and REX1, were mostly dissociated from mitotic chromatin. The biological significance and underlying mechanisms of retaining or excluding specific TFs from mitotic chromatin warrant further investigation. Interestingly, all of the TFs that remained bound on mitotic chromatin have been reported to either have histone-like properties (UTF1) or function as pioneer factors (OCT4, SOX2, KLF4, and ESRRB) (Soufi et al., 2012, 2015), suggesting that intrinsic DNA-binding properties—in addition to TF cooperativity—may determine the bookmarking potential of each protein. In agreement, our experiments showed that ectopically expressed KLF4 in blastomeres and reprogramming TFs in MEFs were able to access mitotic chromatin. However, the bookmarking capacity and specificity of defined TFs in different cellular contexts still need to be determined.

The high degree and specificity of mitotic retention of KLF4, OCT4, and SOX2 we observed in ESCs strongly supports a potential bookmarking role for faithful inheritance of stemness in daughter cells. Indeed, our functional assays confirmed that temporal degradation of OCT4 during G1 entry induced destabilization of stem cell identity and unscheduled differentiation. Although we cannot exclude that the slow recovery of OCT4 protein levels

(3 hr until full recovery, Figure S4C) may also account for the observed phenotypes, our results highlight the importance of this time window for OCT4 to safeguard pluripotency. Moreover, we showed that OCT4 presence during mitotic exit was also critical for efficient reprogramming of somatic cells into iPSCs. Elegant experiments in the past have shown that cell division is required for successful TF-mediated reprogramming, proposing that transient silencing of the somatic-specific transcriptional program and dissociation of the lineage-specifying factors during mitosis provide a unique chance for a new cell fate to arise (Egli et al., 2008, 2011; Koche et al., 2011). Our findings suggest that active mitotic binding of reprogramming factors may facilitate overwriting somatic cell identity.

Mitosis is believed to represent a temporal crisis of cell identity and a window of opportunity for cell fate transitions. Multiple studies have shown that during mitotic exit and G1 entry ESCs are the most susceptible to environmental cues that induce lineage specification (Pauklin and Vallier, 2013; Dalton, 2015; Pauklin et al., 2016). Therefore, understanding the mechanisms that govern this critical cell-cycle window will enable the identification of molecular levers that could be used to shift the balance of self-renewal/differentiation to promote a certain lineage. Our results constitute an important step toward this direction by revealing patterns and basic principles of mitotic occupancy for critical chromatin marks and stem cell regulators.

## EXPERIMENTAL PROCEDURES

See the Supplemental Information for details.

### Cell Culture

V6.5, ZHBTc4.1 (Niwa et al., 2000), and H2B-mCherry expressing ESCs were cultured on irradiated feeder cells in KO-DMEM (Invitrogen) supplemented with GlutaMAX, pen-strep, nonessential amino acids,  $\beta$ -mercaptoethanol, 1,000 U/mL LIF, and 15% heat-inactivated fetal bovine serum.

### Mitotic Arrest and Subcellular Fractionation

ESCs were passaged on gelatinized plates the day before synchronization. Nocodazole (200  $\mu\text{g}/\text{mL}$ ) was added to the medium for 7 hr prior to collection by mitotic shake-off for western blot and ChIP-seq assays. Synchronization efficiency was determined by FACS analysis on ethanol-fixed cells stained with H3Ser10p and DAPI. Chromatin-associated extracts were prepared using a subcellular protein fractionation kit (ThermoScientific, #78840) following manufacturer's instructions.

### ChIP-Seq Assays

ChIP was performed as described previously (Apostolou et al., 2013). Related antibodies and primers are listed in Tables S1 and S4. Library construction and ChIP-seq analysis are described in the Supplemental Information.

### DNA Constructs, Infections, and Functional Assays

Turbo GFP followed by either the Cyclin B destruction box (DB) (13–90 aa) or a mutated R24A version (DBmut) was cloned in frame with Oct4 into: (1) the pHAGE-Ef1a vector to generate GDO or GMO constructs, or (2) the pHAGE-tetO-Stemcca (Sommer et al., 2009) vector to generate the GDO-KSM and GMO-KSM constructs. Virus production and transduction was performed as described previously (Apostolou et al., 2013).

ZHBTc4.1 ESCs either uninfected or infected with GDO or GMO were plated on gelatin without feeders in presence or absence of dox (2  $\mu\text{g}/\text{mL}$ ) in regular ESC medium (+LIF) or in the absence of LIF (–LIF). Cells were then collected for western blot analyses, qRT-PCR, or colony formation and scoring as described in the Supplemental Information.

Mouse embryonic fibroblasts (MEFs) carrying the Rosa26-rtTA allele were infected with GDO-KSM, GMO-KSM, and then reprogrammed as described before (Stadtfield et al., 2012).

### Statistical Analyses

Two-sample two-sided t test or Wilcoxon test were used to calculate the reported p values in Figures 1E and 4B and S4F or 2E, respectively.

### ACCESSION NUMBERS

The accession number for the ChIP-seq data reported in this paper is GEO: GSE92846.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, four tables, and three movies and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.04.067>.

### AUTHOR CONTRIBUTIONS

E.A. conceived, designed, and supervised the study, performed experiments, and wrote the manuscript with help from all authors. Y.L. performed all the bioinformatics analyses under the supervision of O.E. and E.A. and with input from A.D., B.P.-W., and D.K. performed fractionation and western blot analyses. B.P.-W. performed IF on metaphase spreads. D.C.D.G. performed ChIP-seq experiments and validations. B.P.-W. and J.L. generated the GDO and GMO constructs and performed the functional analyses for the maintenance and induction of pluripotency, respectively. D.C., D.G., and K.K. performed live imaging of GDO and GMO expressing ESCs with guidance from P.G., while N.S. and V.G. performed the microinjections and live imaging of blastomeres with guidance from A.-K.H.

### ACKNOWLEDGMENTS

We are grateful to Todd Evans, Ritu Kumar, Matthias Stadtfield, and Konrad Hochedlinger for critical reading of the manuscript. We also want to thank Ari Melnick and Anja Fischer for antibodies and Minjung Kang for help with confocal imaging. B.P. is supported by a Medical Scientist Training Program grant from the National Institute of General Medical Sciences of the NIH under award num-

ber T32GM007739 to the Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program. D.C.D.G. is supported by the New York Stem Cell Foundation (NYSCF-D-F46) and the Family-Friendly Postdoctoral Initiative at Weill Cornell Medicine. This work was funded by the NIH Director's New Innovator Award (DP2DA043813), the Edward Jr. Mallinckrodt Foundation, and the Tri-Institutional Stem Cell Initiative by the Starr Foundation (2016-018).

Received: January 4, 2017

Revised: March 17, 2017

Accepted: April 24, 2017

Published: May 16, 2017

### REFERENCES

- Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. *Nature* 502, 462–471.
- Apostolou, E., Ferrari, F., Walsh, R.M., Bar-Nur, O., Stadtfield, M., Cheloufi, S., Stuart, H.T., Polo, J.M., Ohsumi, T.K., Borowsky, M.L., et al. (2013). Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell* 12, 699–712.
- Burke, L.J., Zhang, R., Bartkuhn, M., Tiwari, V.K., Tavoosidana, G., Kurukuti, S., Weth, C., Leers, J., Galjart, N., Ohlsson, R., and Renkawitz, R. (2005). CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. *EMBO J.* 24, 3291–3300.
- Caravaca, J.M., Donahue, G., Becker, J.S., He, X., Vinson, C., and Zaret, K.S. (2013). Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes Dev.* 27, 251–260.
- Cimini, D., Mattiuzzo, M., Torosantucci, L., and Degrossi, F. (2003). Histone hyperacetylation in mitosis prevents sister chromatid separation and produces chromosome segregation defects. *Mol. Biol. Cell* 14, 3821–3833.
- Coronado, D., Godet, M., Bourillot, P.Y., Taponnier, Y., Bernat, A., Petit, M., Afanassieff, M., Markossian, S., Malashicheva, A., Iacone, R., et al. (2013). A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. *Stem Cell Res. (Amst.)* 10, 118–131.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* 107, 21931–21936.
- Dalton, S. (2015). Linking the cell cycle to cell fate decisions. *Trends Cell Biol.* 25, 592–600.
- Deluz, C., Friman, E.T., Strebinger, D., Benke, A., Raccaud, M., Callegari, A., Leleu, M., Manley, S., and Suter, D.M. (2016). A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. *Genes Dev.* 30, 2538–2550.
- Egli, D., Birkhoff, G., and Eggan, K. (2008). Mediators of reprogramming: transcription factors and transitions through mitosis. *Nat. Rev. Mol. Cell Biol.* 9, 505–516.
- Egli, D., Chen, A.E., Saphier, G., Ichida, J., Fitzgerald, C., Go, K.J., Acevedo, N., Patel, J., Baetscher, M., Kearns, W.G., et al. (2011). Reprogramming within hours following nuclear transfer into mouse but not human zygotes. *Nat. Commun.* 2, 488.
- Evans, M. (2011). Discovering pluripotency: 30 years of mouse embryonic stem cells. *Nat. Rev. Mol. Cell Biol.* 12, 680–686.
- Festuccia, N., Dubois, A., Vandormael-Pourmin, S., Gallego Tejada, E., Mouren, A., Bessonard, S., Mueller, F., Proux, C., Cohen-Tannoudji, M., and Navarro, P. (2016). Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol.* 18, 1139–1148.
- Glotzer, M., Murray, A.W., and Kirschner, M.W. (1991). Cyclin is degraded by the ubiquitin pathway. *Nature* 349, 132–138.
- Gottesfeld, J.M., and Forbes, D.J. (1997). Mitotic repression of the transcriptional machinery. *Trends Biochem. Sci.* 22, 197–202.
- Halley-Stott, R.P., Jullien, J., Pasque, V., and Gurdon, J. (2014). Mitosis gives a brief window of opportunity for a change in gene transcription. *PLoS Biol.* 12, e1001914.

- Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., et al. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* 6, 479–491.
- Holloway, S.L., Glotzer, M., King, R.W., and Murray, A.W. (1993). Anaphase is initiated by proteolysis rather than by the inactivation of maturation-promoting factor. *Cell* 73, 1393–1402.
- Hsiung, C.C., Bartman, C.R., Huang, P., Ginart, P., Stonestrom, A.J., Keller, C.A., Face, C., Jahn, K.S., Evans, P., Sankaranarayanan, L., et al. (2016). A hyperactive transcriptional state marks genome reactivation at the mitosis-G1 transition. *Genes Dev.* 30, 1423–1439.
- Iwafuchi-Doi, M., and Zaret, K.S. (2014). Pioneer transcription factors in cell reprogramming. *Genes Dev.* 28, 2679–2692.
- Kadauke, S., and Blobel, G.A. (2013). Mitotic bookmarking by transcription factors. *Epigenetics Chromatin* 6, 6.
- Kadauke, S., Udegama, M.I., Pawlicki, J.M., Achtman, J.C., Jain, D.P., Cheng, Y., Hardison, R.C., and Blobel, G.A. (2012). Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell* 150, 725–737.
- Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., Bernstein, B.E., and Meissner, A. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* 8, 96–105.
- Kooistra, S.M., Thummer, R.P., and Eggen, B.J. (2009). Characterization of human UTF1, a chromatin-associated protein with repressor activity expressed in pluripotent cells. *Stem Cell Res. (Amst.)* 2, 211–218.
- Koukouti, A., and Talianidis, I. (2005). Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *EMBO J.* 24, 347–357.
- Li, Y., Kao, G.D., Garcia, B.A., Shabanowitz, J., Hunt, D.F., Qin, J., Phelan, C., and Lazar, M.A. (2006). A novel histone deacetylase pathway regulates mitosis by modulating Aurora B kinase activity. *Genes Dev.* 20, 2566–2579.
- Liang, K., Woodfin, A.R., Slaughter, B.D., Unruh, J.R., Box, A.C., Rickels, R.A., Gao, X., Haug, J.S., Jaspersen, S.L., and Shilatifard, A. (2015). Mitotic transcriptional activation: clearance of actively engaged Pol II via transcriptional elongation control in mitosis. *Mol. Cell* 60, 435–445.
- Liu, Y., Chen, S., Wang, S., Soares, F., Ficher, M., Meng, F., Du, Z., Lin, C., Meyer, C., DeCaprio, J.A., et al. (2017). Transcriptional landscape of the human cell cycle. *Proc. Natl. Acad. Sci. USA* 114, 3473–3478.
- Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. *Nat. Rev. Genet.* 11, 285–296.
- Martínez-Balbás, M.A., Dey, A., Rabindran, S.K., Ozato, K., and Wu, C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* 83, 29–38.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaaf, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 28, 495–501.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553–560.
- Natoli, G. (2010). Maintaining cell identity through global control of genomic organization. *Immunity* 33, 12–24.
- Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* 24, 372–376.
- Pauklin, S., and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. *Cell* 155, 135–147.
- Pauklin, S., Madrigal, P., Bertero, A., and Vallier, L. (2016). Initiation of stem cell differentiation involves cell cycle-dependent regulation of developmental genes by Cyclin D. *Genes Dev.* 30, 421–433.
- Peters, J.M., Tedeschi, A., and Schmitz, J. (2008). The cohesin complex and its roles in chromosome biology. *Genes Dev.* 22, 3089–3114.
- Savatier, P., Lapillonne, H., Jirmanova, L., Vitelli, L., and Samarut, J. (2002). Analysis of the cell cycle in mouse embryonic stem cells. *Methods Mol. Biol.* 185, 27–33.
- Sciortino, S., Gurtner, A., Manni, I., Fontemaggi, G., Dey, A., Sacchi, A., Ozato, K., and Piaggio, G. (2001). The cyclin B1 gene is actively transcribed during mitosis in HeLa cells. *EMBO Rep.* 2, 1018–1023.
- Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., and Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature* 488, 116–120.
- Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* 27, 543–549.
- Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151, 994–1004.
- Soufi, A., Garcia, M.F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K.S. (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* 161, 555–568.
- Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R.M., Chen, T., Ooi, S.S., Kim, S.Y., Bestor, T.H., Shioda, T., et al. (2012). Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nat. Genet.* 44, 398–405.
- Tabar, V., and Studer, L. (2014). Pluripotent stem cells in regenerative medicine: challenges and recent progress. *Nat. Rev. Genet.* 15, 82–92.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.
- Tapia, C., Kutzner, H., Mentzel, T., Savic, S., Baumhoer, D., and Glatz, K. (2006). Two mitosis-specific antibodies, MPM-2 and phospho-histone H3 (Ser28), allow rapid and precise determination of mitotic activity. *Am. J. Surg. Pathol.* 30, 83–89.
- Teves, S.S., An, L., Hansen, A.S., Xie, L., Darzacq, X., and Tjian, R. (2016). A dynamic mode of mitotic bookmarking by transcription factors. *eLife* 5, e22280.
- Valls, E., Sánchez-Molina, S., and Martínez-Balbás, M.A. (2005). Role of histone modifications in marking and activating genes through mitosis. *J. Biol. Chem.* 280, 42592–42600.
- van den Boom, V., Kooistra, S.M., Boesjes, M., Geverts, B., Houtsmuller, A.B., Monzen, K., Komuro, I., Essers, J., Drenth-Diephuis, L.J., and Eggen, B.J. (2007). UTF1 is a chromatin-associated protein involved in ES cell differentiation. *J. Cell Biol.* 178, 913–924.
- Wang, F., and Higgins, J.M. (2013). Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol.* 23, 175–184.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319.
- Young, R.A. (2011). Control of the embryonic stem cell state. *Cell* 144, 940–954.
- Young, D.W., Hassan, M.Q., Pratap, J., Galindo, M., Zaidi, S.K., Lee, S.H., Yang, X., Xie, R., Javed, A., Underwood, J.M., et al. (2007). Mitotic occupancy and lineage-specific transcriptional control of rRNA genes by Runx2. *Nature* 445, 442–446.
- Zaidi, S.K., Young, D.W., Montecino, M.A., Lian, J.B., van Wijnen, A.J., Stein, J.L., and Stein, G.S. (2010). Mitotic bookmarking of genes: a novel dimension to epigenetic control. *Nat. Rev. Genet.* 11, 583–589.
- Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 25, 2227–2241.

# Whole-Body Profiling of Cancer Metastasis with Single-Cell Resolution

Shimpei I. Kubota,<sup>1,5</sup> Kei Takahashi,<sup>2,5</sup> Jun Nishida,<sup>2</sup> Yasuyuki Morishita,<sup>2</sup> Shogo Ehata,<sup>2</sup> Kazuki Tainaka,<sup>1,3</sup> Kohei Miyazono,<sup>2,\*</sup> and Hiroki R. Ueda<sup>1,4,6,\*</sup>

<sup>1</sup>Department of Systems Pharmacology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

<sup>2</sup>Department of Molecular Pathology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

<sup>3</sup>Brain Research Institute, Niigata University, 1-757 Asahimachidori, Chuo-ku, Niigata 951-8585, Japan

<sup>4</sup>Laboratory for Synthetic Biology, RIKEN Quantitative Biology Center, 1-3 Yamadaoka, Suita, Osaka 565-5241, Japan

<sup>5</sup>These authors contributed equally

<sup>6</sup>Lead Contact

\*Correspondence: miyazono@m.u-tokyo.ac.jp (K.M.), uedah-ky@umin.ac.jp (H.R.U.)

<http://dx.doi.org/10.1016/j.celrep.2017.06.010>

## SUMMARY

Stochastic and proliferative events initiated from a single cell can disrupt homeostatic balance and lead to fatal disease processes such as cancer metastasis. To overcome metastasis, it is necessary to detect and quantify sparsely distributed metastatic cells throughout the body at early stages. Here, we demonstrate that clear, unobstructed brain/body imaging cocktails and computational analysis (CUBIC)-based cancer (CUBIC-cancer) analysis with a refractive index (RI)-optimized protocol enables comprehensive cancer cell profiling of the whole body and organs. We applied CUBIC-cancer analysis to 13 mouse models using nine cancer cell lines and spatiotemporal quantification of metastatic cancer progression at single-cell resolution. CUBIC-cancer analysis suggests that the epithelial-mesenchymal transition promotes not only extravasation but also cell survival at metastatic sites. CUBIC-cancer analysis is also applicable to pharmacotherapeutic profiling of anti-tumor drugs. CUBIC-cancer analysis is compatible with *in vivo* bioluminescence imaging and 2D histology. We suggest that a scalable analytical pipeline with these three modalities may contribute to addressing currently incurable metastatic diseases.

## INTRODUCTION

Rudolf Virchow, the father of modern pathology, declared that the body is a “cell state, in which every cell is a citizen” and disease is “merely a conflict of citizens in this state brought about by the action of external forces” (Nicholls, 1927). His cell theory brought the current concept of cellular pathology based on microscopic histological analysis to the early model of pathol-

ogy, which largely relied on clinical symptoms and gross appearances in patients. Thereafter, comprehensive analysis and decoding of the more than 100 billion cells comprising the mammalian body have become one of the ultimate goals in biology and medicine. Although recent advances in a series of live-imaging systems have provided important breakthroughs to examine dynamics at the whole-body scale, it is still difficult to clarify the generation and progression of diseases with stochastic and proliferative processes, such as autoimmune and malignant neoplastic diseases (Saadatpour et al., 2017; Takahashi et al., 2015). In particular, tumor metastasis can be initiated by a few cancer cells, and use of current imaging systems is a challenge due to their limited spatial resolution. Additionally, according to metastasis images from these imaging systems, 3D reconstitution of possibly affected organs by 2D histology is labor intensive. To bridge the spatial resolution gap between these approaches, a fundamentally novel analytical platform is required. To detect and quantify sparsely distributed metastatic cells at early stages of cancer metastasis, it is necessary to detect metastatic cells throughout the whole body and visualize individual tumor microenvironments.

Tissue-clearing-based 3D imaging is a most promising strategy for the visualization of the entire mouse body/organs at single-cell resolution (Susaki and Ueda, 2016). Recently, various kinds of tissue-clearing protocols have been developed, including organic solvent-based methods (Dodt et al., 2007; Ertürk et al., 2012; Renier et al., 2014), hydrophilic chemical-based methods (Cuccarese et al., 2017; Hama et al., 2011; Ke et al., 2013), and hydrogel-based methods (Chung et al., 2013). Spalteholz (1914) first established the basic concept of human tissue clearing based on organic solvents. Although tissue-clearing protocols have been available for a century, it is still difficult to comprehensively visualize sparsely distributed pathogenic cells throughout the mouse body using current tissue-clearing protocols.

In this study, we first investigate optimal refractive indices (RIs) ranging from 1.44 to 1.52 against each organ, and we demonstrate that an RI of 1.52 has the highest clearing efficiency among major organs. Then, we describe a highly efficient clearing



protocol for the whole adult mouse body using a more efficient delipidation cocktail, clear unobstructed brain/body imaging cocktails and computational analysis (CUBIC)-L, and an optimized RI-matching medium, CUBIC-R. The specimens treated with this clearing protocol are clear enough to use for 3D imaging of the entire mouse body. The combination of light-sheet fluorescent microscopy (LSFM) and confocal laser-scanning microscopy (CLSM) enables the detection of cancer micrometastases in various kinds of mouse models. Our CUBIC-cancer analysis is compatible with the *in vivo* bioluminescence imaging system and conventional 2D histology, successfully bridging the spatial resolution gap between these approaches. CUBIC-cancer analysis is applied here to 13 mouse models using nine cancer cell lines, and it distinguishes different hematogenous metastatic patterns, such as angiogenic growth and co-optive growth. Moreover, CUBIC-cancer analysis enables the spatiotemporal quantification of metastatic cells even in the early stages of metastasis formation. In addition, we investigated the *in vivo* role of transforming growth factor  $\beta$  (TGF- $\beta$ ) signaling in lung metastasis, and we applied an *in vivo* system to evaluate the therapeutic effect of anti-cancer drugs on cancer metastasis. CUBIC-cancer analysis provides a fundamental platform for the analysis of the metastatic growth and progression of cancer throughout the body.

## RESULTS

### Optimized RI Cocktail Enables Whole-Body and Whole-Organ Clearing in Adult Mice

Previously, we achieved the whole-body imaging of postnatal day 1 mice at single-cell resolution by whole-body clearing based on a ScaleCUBIC-1 reagent (Tainaka et al., 2014). However, it remains difficult to sensitively and comprehensively detect the distribution of cancer cells that have metastasized to organs throughout the body of adult mice because of an insufficient delipidation efficiency and the unoptimized RI value of the ScaleCUBIC-1 reagent. To address these issues, we have performed comprehensive chemical profiling of hydrophilic chemicals to characterize highly effective delipidation reagents and water-soluble high-RI reagents (T.C. Murakami and K. Tainaka, unpublished data). Through a further combinatorial analysis, we identified the optimized cocktail of 10 w%/10 w% *N*-butyldiethanolamine/Triton X-100 (termed CUBIC-L) for delipidation. CUBIC-L enabled the complete delipidation of isolated paraformaldehyde (PFA)-fixed organs within 2 to 5 days, depending on the organ without pre-treatment with CUBIC-perfusion and quenching of proteins (Figure 1A, left). We found that CUBIC-L could also permeabilize skin-detached body samples of adult

mice using the CUBIC-perfusion protocol and subsequent delipidation for 7 days (Figure 1A, right; Supplemental Experimental Procedures).

As for the RI-matching step, there is an ongoing controversy about the most appropriate RI for permeabilized tissues (Tainaka et al., 2016). The RI for each intact biological tissue may depend on the contents and densities of lipids, proteins, and other constituents (Johnsen and Widder, 1999; Tuchin, 2015). Since hydrophilic RI media with an RI over 1.52 were likely to precipitate at room temperature, we evaluated RI-matching cocktails composed of antipyrine and nicotinamide with RIs ranging from 1.44 to 1.52 for clearing permeabilized major organs, considering practical use. After a 2-day treatment, all major organs appeared to be most efficiently transparentized by the RI-matching cocktail with an RI of 1.52 (Figures 1B and S1A), termed CUBIC-R, which is composed of 45 w%/30 w% antipyrine/nicotinamide. Indeed, the transmittance of the chemically treated organs increased as the RI value increased (Figures 1C and S1B). We noted that all organs treated with CUBIC-L and CUBIC-R showed significantly high transmittance in the 450 to 750 nm range, probably due to the high decoloring ability of *N*-butyldiethanolamine in CUBIC-L (Figure S1C). These results prompted us to render the whole-body clearing of adult mice using CUBIC-R (RI = 1.52) as a final RI-matching step. The transmission images of adult mice demonstrated that the new protocol markedly transparentized adult mice to visualize caudal vertebrae in ventral to dorsal (V-D) images (Figure 1D).

### The New CUBIC Protocol Enables the Whole-Body and Whole-Organ Imaging of Adult Mice by 3D Nuclear Counterstaining and Immunohistochemistry

Simultaneous visualization of cancer cells, overall organ structures, and related molecular markers would facilitate our understanding of the underlying molecular mechanisms of cancer metastasis. For the comprehensive analysis of cancer metastasis, we initially carried out nuclear counterstaining with RedDot2 or propidium iodide (PI) to acquire whole-body and whole-organ structural images using LSFM according to our previous staining protocols (Figure 1E) (Susaki et al., 2014; Tatsuki et al., 2016). The resulting 3D-reconstituted image of each organ enabled the visualization of anatomical structural images even in deeper regions (Figure S1D). The new CUBIC protocol allowed the whole-body imaging of adult mice, clearly visualizing the internal structures of cardiovascular and abdominal organs (Figures 1F and S1E).

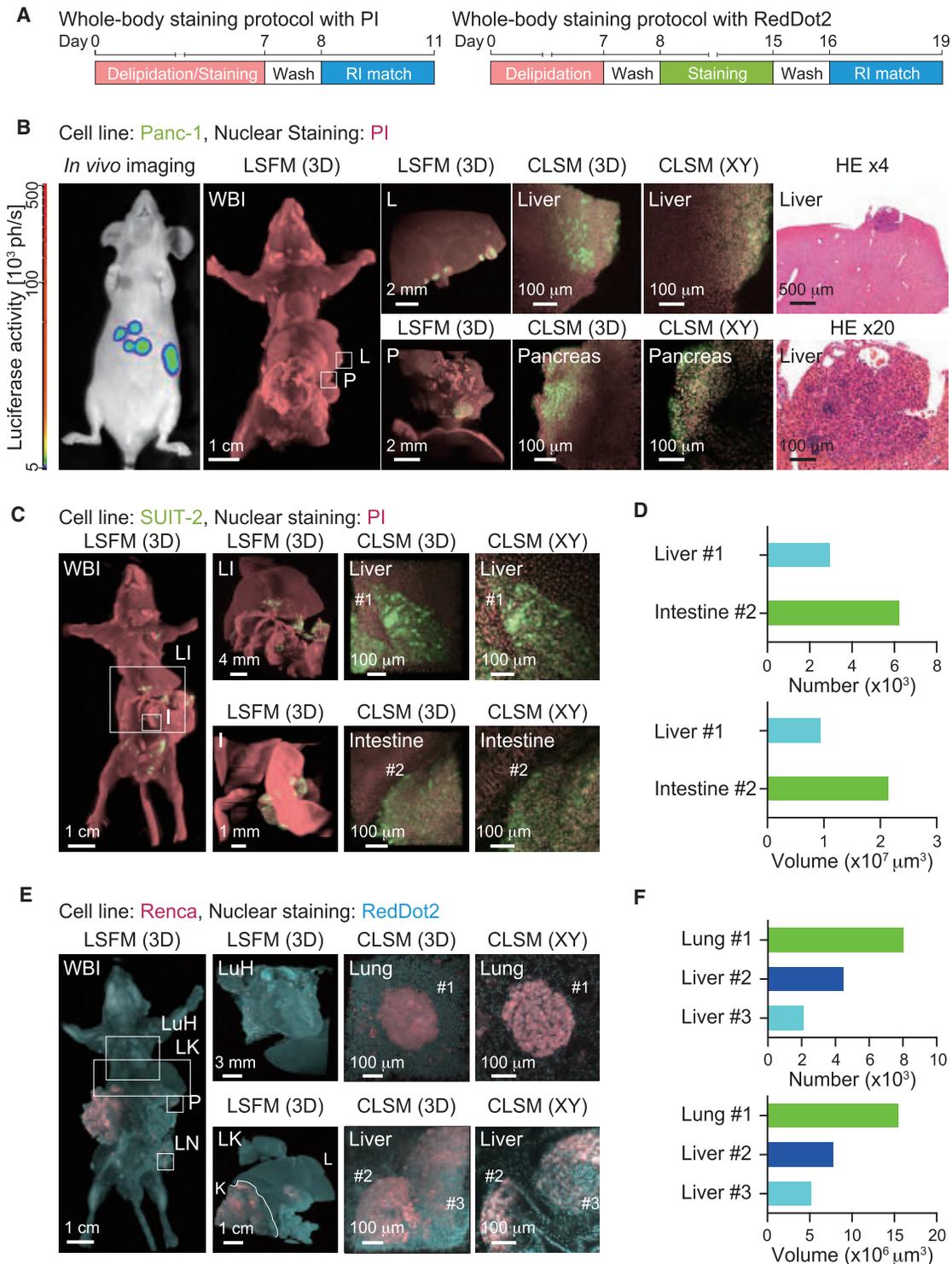
Next, we investigated whether the new CUBIC protocol is compatible with 3D immunohistochemistry. The brains of adult mice were co-stained with RedDot2 and anti- $\alpha$ -smooth muscle

(D) Bright-field images of whole body of C57BL/6N mouse (6 weeks old, male) treated with CUBIC-R. After delipidation and decolorization, adult mouse bodies were treated with PBS buffer (upper) or CUBIC-R (lower).

(E) Whole-body and whole-organ staining protocols. The whole-body staining protocol with PI (left) and the whole-organ staining protocol with RedDot2 or fluorescently labeled primary antibody (right) are shown.

(F) The 3D-reconstituted (3D) body images of PI-stained adult C57BL/6N mouse (10 weeks old, male). Whole-body imaging (WBI), thoracic organ imaging, and abdominal organ imaging are shown.

(G) The 3D-immunostained organ images. The brain from C57BL/6N mouse (15 weeks old, male) was stained with RedDot2 and fluorescein isothiocyanate (FITC)-conjugated anti- $\alpha$ -SMA antibody. 3D (upper) and x-y plane (XY) (lower) images are shown. The white inset was magnified next to the XY image. See also Figure S1.



**Figure 2. Whole-Body Imaging of Cancer Metastasis at Single-Cell Resolution**

(A) Whole-body and whole-organ staining protocols. The whole-body staining protocol with PI (left) and the whole-body staining protocol with RedDot2 (right) are shown.

(B) Whole-body imaging of the experimental liver metastasis model by splenic injection with Panc-1 cells in BALB/c-*nu/nu* mice. The bioluminescence image (leftmost), WBI, images around the metastatic sites in the liver (L) and pancreas (P) with CUBIC-cancer analysis (second and third from left), magnified 3D and XY images (fourth and fifth from left) (Panc-1: GFP, nuclei: PI), and representative images of HE staining after CUBIC-cancer analysis (rightmost) are shown.

(legend continued on next page)

actin ( $\alpha$ -SMA) antibody according to the staining protocol (Figures 1E and 1G). Spatial distributions of smooth muscle cells and lymphatic endothelial cells were simultaneously visualized in the lung by co-staining of  $\alpha$ -SMA and vascular endothelial growth factor receptor 3 (VEGFR3) (Figure S1F). Hence, our CUBIC protocol was also applicable to 3D multiple immunohistochemical labeling. These results indicate that the new CUBIC protocol is a powerful tool for comprehensive pathological analysis.

### CUBIC-Cancer Analysis Enables the Whole-Body Imaging of Cancer Metastasis at Single-Cell Resolution

To bridge the resolution gap between *in vivo* bioluminescence imaging and 2D histology with our scalable CUBIC-based cancer (CUBIC-cancer) analysis, we established cancer cells stably expressing firefly luciferase and fluorescence proteins, and we used them in mouse models (Figure 2A). First, we employed an experimental liver metastasis model by splenic injection of human pancreatic cancer cells (Panc-1). Nine weeks after injection, bioluminescence signals were detected in the abdomens of the mice (Figure 2B, left panel). We applied the CUBIC-cancer analysis and PI-staining protocol to this mouse model, and then we acquired whole-body images with LSFM (Figure 2A). As a result, fluorescence signals were successfully visualized in several organs, suggesting that the CUBIC-cancer analysis is applicable to the fluorescent proteins (Figure 2B, middle panels). In addition, magnified 3D images of these organs were obtained with CLSM in the CUBIC-cancer analysis. We noted that the resolution of these 3D images was high enough to discriminate individual cancer cells and comparable to that of 2D hematoxylin and eosin (HE) slice images, which were obtained after CUBIC-cancer analysis (Figure 2B, right panels).

As another pathway of cancer metastasis, peritoneal dissemination model mice were generated by the intraperitoneal injection of human pancreatic cancer cells (SUIT-2). Twenty-two days after injection, mCherry signals from SUIT-2 cells were observed in the liver and small intestine by CUBIC-cancer analysis (Figure 2C), roughly corresponding to those of bioluminescence signals from *in vivo* bioluminescence imaging (Figure S2A, left panel). Combined with counterstained confocal images, we succeeded in the direct quantification of the cell numbers of cancer colonies (Figures 2C and 2D). The CUBIC-cancer analysis also recapitulated cancer dissemination in other mice with SUIT-2 cells (Figures S2B and S2C), indicating the wide applicability of the CUBIC-cancer analysis protocol.

To prepare a syngeneic tumor model, we orthotopically injected murine renal cancer cells (Renca) into wild-type BALB/c

mice. Although the formation of a primary renal tumor was visualized 2 weeks after injection, it was difficult to detect cancer metastasis to other organs by *in vivo* bioluminescence imaging (Figure S2A, right panel). In contrast, CUBIC-cancer analysis revealed that numerous metastatic colonies had widely spread to the lung, liver, pancreas, and mesenteric lymph nodes (Figures 2E and S2D). Using CUBIC-cancer analysis combined with LSFM and CLSM, we also succeeded in the direct quantification of the cell numbers of these dispersed metastatic colonies (Figures 2F and S2E). Taken together, CUBIC-cancer analysis makes it possible to not only detect the dispersion of metastatic cancer cells throughout the entire body with high sensitivity but also determine the spatial distribution of individual cancer cells at single-cell resolution. In addition, CUBIC-cancer analysis is compatible with *in vivo* bioluminescence imaging and 2D histology.

### Whole-Organ CUBIC-Cancer Analysis Provides a Systematic, Robust, Widely Applicable, and Immunolabeling-Compatible Analytical Platform for Cancer Metastasis

After the comprehensive detection of cancer metastasis throughout the body, the whole-organ CUBIC-cancer analysis would globally delineate the metastatic landscape of each cancer metastasis model. To this end, whole-organ clearing and staining were conducted according to Figure 3A. Initially, we quantified the growth of primary tumors in orthotopic models with SUIT-2 cells and OS-RC-2 cells (Figure S3A). CUBIC-cancer analysis also permitted us to visualize metastatic colonies in whole organs, including the lung, liver, and intestine (Figure 3B). The colony number and volume of the metastasis in each organ were calculated (Figure S3B). We also exemplified CUBIC-cancer analysis with a Panc-1 cell liver metastasis model (Figure S3C). This highly reproducible analysis would overcome enduring issues in our understanding of *in vivo* metastatic progression in various cancers. Thus, CUBIC-cancer analysis enabled the complete quantification of metastatic colonies in an individual tumor-bearing mouse.

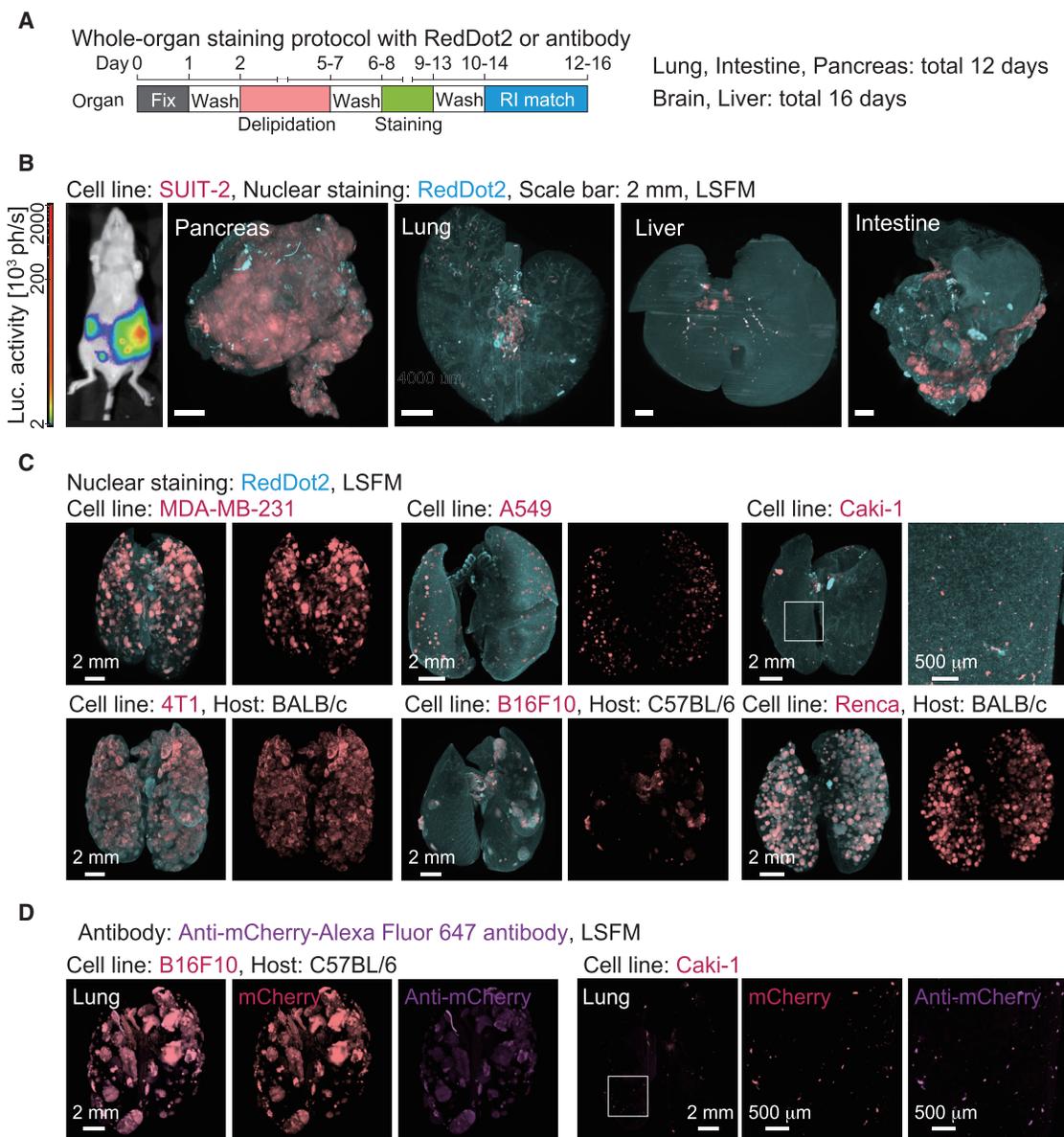
To verify CUBIC-cancer analysis as a general purpose platform, several lung metastasis models derived from various types of cancer cells were examined (Figure 3C). The detected red-color signals were shown to have originated from mCherry, using immunohistochemical staining with anti-mCherry antibody (Figures 3D and S3D). In the experimental metastasis models, all mouse cancer cells, *i.e.*, 4T1, B16F10, and Renca cells, tended to grow to large colonies, while the morphologies of these colonies were all different. In case of experimental metastasis

(C) Whole-body imaging of the experimental peritoneal dissemination model by intraperitoneal injection with SUIT-2 cells in BALB/c-*nu/nu* mice. WBI and abdominal organ images in liver (L) and intestine (I) (leftmost and second from left) and magnified 3D and XY images (third from left and rightmost) are shown (SUIT-2, GFP; nuclei, PI).

(D) Quantification of the cell number and the volume of metastatic colonies. Spot analysis and surface analysis were applied to the magnified images in (C).

(E) Whole-body imaging of the spontaneous metastasis model with orthotopic injection of Renca cells in BALB/c mice. WBI and organ images in lung (Lu), heart (H), liver (L), kidney (K), pancreas (P), and lymph nodes (LN) (leftmost and second from left) and magnified 3D and XY images (third from left and rightmost) are shown (Renca, mCherry; nuclei, RedDot2).

(F) Quantification of the cell number and the volume of metastatic colonies. Spot analysis and surface analysis were applied to the magnified images in (E). See also Figure S2 and Table S1.



**Figure 3. Whole-Organ Imaging of Various Cancer Metastasis Models**

(A) Whole-organ staining protocols with RedDot2 or antibody.

(B) Whole-organ imaging of the spontaneous metastasis model with orthotopic injection of SUIT-2 cells in BALB/c-*nu/nu* mice. The bioluminescence image (leftmost) and primary tumors in the pancreas and the metastatic colonization in the distant organs (second, third, and fourth from left and rightmost) are shown (SUIT-2, mCherry; nuclei, RedDot2).

(C) Whole-lung imaging of several metastasis models. In the experimental lung metastasis model, cancer cells (MDA-MB-231, A549, 4T1, B16F10, and Renca) were intravenously injected in mice. In case of spontaneous metastasis, Caki-1 cells were orthotopically injected into the kidney (cancer cell, mCherry; nuclei, RedDot2).

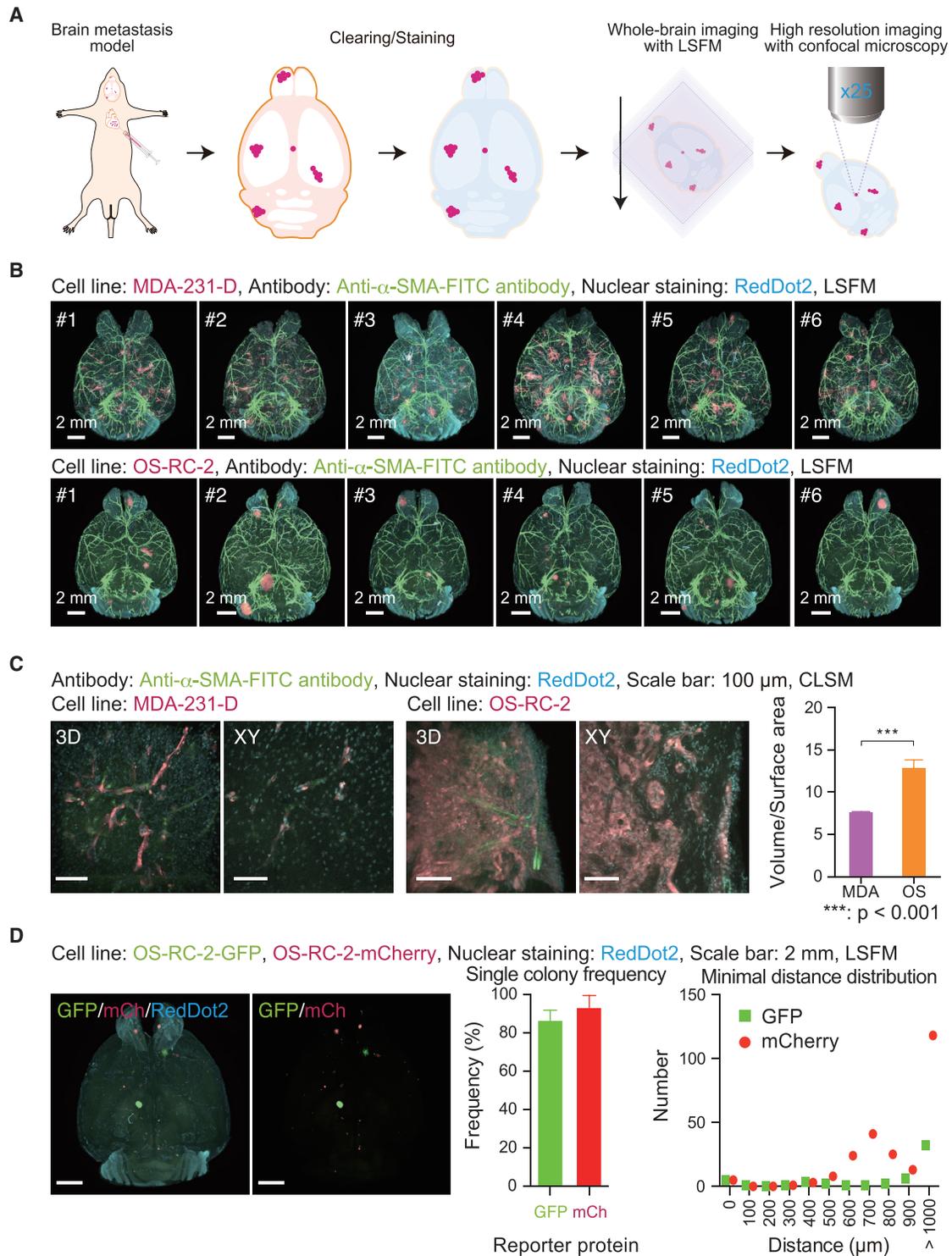
(D) Whole-lung 3D immunohistochemistry with anti-mCherry antibody (cancer cell, mCherry; anti-mCherry, Alexa Fluor 647). White insets were magnified next to the whole-lung images in (C) and (D).

See also Figure S3.

with human cancer cells, MDA-MB-231 cells also formed large colonies throughout the lungs compared to A549 cells. In contrast, in spontaneous lung metastasis with Caki-1 cells, metastatic colonies were observed throughout the lungs, but their volumes were small. Overall, we succeeded in applying

CUBIC-cancer analysis to 13 different mouse models with nine cancer cell lines (Table S1).

These results demonstrate that whole-organ CUBIC-cancer analysis is a systematic, robust, widely applicable, and immunolabeling-compatible analytical platform for cancer metastasis.



**Figure 4. 3D Analysis of the Metastatic Patterns in Experimental Brain Metastasis Models**

(A) Scheme of the experimental procedure.

(B) Whole-brain imaging of the experimental brain metastasis models with MDA-231-D cells or OS-RC-2 cells in BALB/c-*nu/nu* mice. The 3D images of the brain samples are shown (upper, MDA-231-D at day 28; lower, OS-RC-2 at day 40; cancer cell, mCherry;  $\alpha$ -SMA, FITC; nuclei, RedDot2). Animal number of each group is  $n = 6$ .

(legend continued on next page)

### CUBIC-Cancer Analysis Enables Monitoring of the Different Patterns of Brain Metastasis Formation with Blood Vessels in 3D

Brain metastasis is frequently observed in melanoma and some other cancers, and it results in a high risk of mortality (Gavrilovic and Posner, 2005; Nussbaum et al., 1996). Recently, two-photon microscopy-based analytical methods have clarified the correlation between brain metastasis and the cerebral blood vasculature (Carbonell et al., 2009; Kienast et al., 2010). Nevertheless, quantitative and statistical analyses of the spatial distribution of all metastatic colonies throughout the brain are still challenging. Thus, we investigated whether our CUBIC-cancer analysis can be used to classify the potentially different types of brain metastasis by quantitatively and statistically analyzing metastatic colonies in the entire organ (Figure 4A). The brain metastasis model of MDA-231-D cells was compared with another model using OS-RC-2 cells. Although there appeared to be no distinct difference in metastatic features between these models by *in vivo* bioluminescence imaging (Figure S4A), CUBIC-cancer analysis clearly showed that metastatic colonies in these models exhibited different histological patterns (Figure 4B).

Many colonies of MDA-231-D cells were localized with  $\alpha$ -SMA-positive vascular smooth muscle cells. On the other hand, most colonies of OS-RC-2 cells were distant from  $\alpha$ -SMA-positive vascular smooth muscle cells (Figure 4B). In addition, these metastatic morphological features were also identified by the histological analysis based on HE staining with or without CUBIC-cancer analysis (Figure S4B). To quantify the morphological features, we used the index “volume per surface area,” which was positively correlated with the roundness of the colony, and we found a significant difference between the two metastatic models (Figure 4C). Furthermore, the colonies of MDA-231-D cells tended to be smaller and closer to the blood vessels than those of OS-RC-2 cells (Figure S4C). These results suggest that the colony formation of OS-RC-2 cells in the brain may be dependent on cancer cell-mediated neovascularization (angiogenic growth), whereas MDA-231-D cells may be able to migrate toward pre-existing vasculatures and hijack them during the process of metastasis (co-optive growth).

For elucidating the mechanism of cancer metastasis, it is important to determine whether brain metastatic foci are generated by a single cell. To test this, equal numbers of mCherry- or GFP-expressing OS-RC-2 cells were mixed and used for intracardiac injection into the same mouse. CUBIC-cancer analysis revealed that most mCherry-positive cells in the brain were separate from the colonies of GFP-positive cells (Figure 4D), thus revealing the metastatic properties of individual cancer cells.

### CUBIC-Cancer Analysis Is Applicable to Statistical Spatiotemporal Analysis during the Initial Steps of Metastatic Progression

The statistical spatiotemporal analysis of cancer metastasis in an entire organ has the potential to be a powerful tool for cancer research. To visualize time-dependent cancer metastasis, an experimental lung metastasis model with MDA-231-D cells was used in combination with CUBIC-cancer analysis. We then carried out whole-lung imaging of metastatic foci or colonies (Figure 5A). To ensure the non-biased exhaustive detection of size-variable metastatic foci, the average volume of nuclei was adopted as the threshold for focus detection (Figure S5A). Immunodeficient BALB/*c-nu/nu* mice and immunocompetent BALB/*c* mice were intravenously injected with MDA-231-D cells, and metastatic foci were detected by CUBIC-cancer analysis 1 hr after injection (Figures 5B and 5C). In total, three to four mice at each time point were analyzed (Figures S5B and S5C).

In BALB/*c-nu/nu* mice, numerous metastatic foci spread throughout the lung were drastically eliminated within 1 day after injection (Figure 5D). Although the number of foci continued to decrease after day 1, the average volume of the foci markedly increased from day 3, suggesting metastatic colonization. On the other hand, in BALB/*c* mice, metastatic foci were thoroughly excluded without any colonization (Figures 5E and S5C). Confocal images were used to quantify the emerged metastatic colonies on day 14 in BALB/*c-nu/nu* mice and the accumulated cell aggregations in BALB/*c* mice on day 3 (Figures 5F and 5G). Since the mCherry signals were often surrounded by those of the cell aggregations, the aggregations likely reflected the immune response of the host (Figure 5H). Overall, CUBIC-cancer analysis enabled not only a statistical time course analysis of cancer metastasis but also direct observations of the stochastic immune-mediated elimination of metastatic foci.

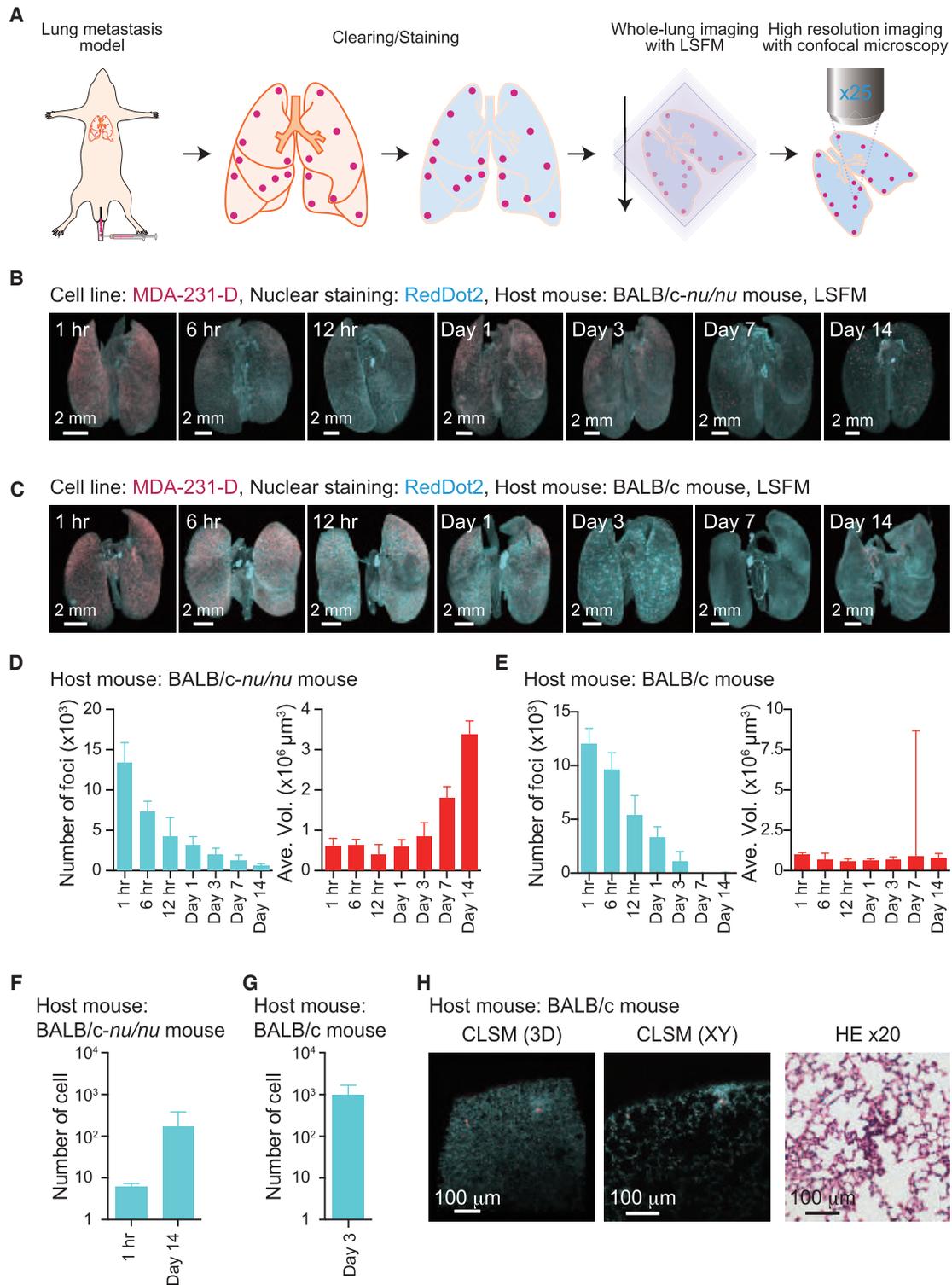
### CUBIC-Cancer Analysis Reveals that the Epithelial-Mesenchymal Transition Might Promote Not Only the Extravasation but Also the Survival of Cancer Cells at Metastatic Sites

The epithelial-mesenchymal transition (EMT) is known to play a pivotal role in cancer metastasis. Although roles of EMT in the intravasation of cancer cells have been demonstrated (Nieto et al., 2016; Tsai and Yang, 2013), it has not been clearly determined whether the EMT is involved in the extravasation of cancer cells (Nieto et al., 2016; Tsai and Yang, 2013). Thus, we aimed to elucidate the EMT involvement in the extravasation of cancer metastasis *in vivo* by CUBIC-cancer analysis. In accordance with a previous report (Saito et al., 2009), induction of the EMT in A549 cells by TGF- $\beta$  stimulation was confirmed by gene profiling and morphological examination (Figures S6A and S6B).

(C) Quantification of the morphological features in the two metastatic models. Magnified 3D and XY images of typical metastatic foci are shown. Surface analysis was applied to the 3D images in (B). As a characteristic indicator of different metastatic patterns, the volume of the foci divided by its surface area was calculated. Data represent mean  $\pm$  SEM (\*\*\*) $p < 0.001$ .

(D) Lineage analysis using renal cancer cell lines expressing GFP or mCherry. The 3D images of the brains stained with RedDot2 (left), the frequency of the metastatic colonies consisting of a single color (middle), and the minimal distance between metastatic colonies expressing GFP or mCherry (right) are shown. Data represent mean  $\pm$  SD.

See also Figure S4.



**Figure 5. Spatiotemporal Dynamics of Metastatic Progression in the Experimental Lung Metastasis Model**

(A) Scheme of the experimental procedure.

(B and C) Time course images of the experimental lung metastasis model with MDA-231-D cells in BALB/c-*nu/nu* mice (B) or BALB/c mice (C). BALB/c-*nu/nu* mice or BALB/c mice were intravenously injected with MDA-231-D cells. The 3D images of the lung samples at the indicated time points are shown (MDA-231-D, mCherry; nuclei, RedDot2).

(legend continued on next page)

To investigate the EMT roles in cancer cell arrest in microvessels of distant organs and extravasation, BALB/c-*nu/nu* mice were intravenously injected with A549 cells pre-treated with or without TGF- $\beta$  (Figure 6A). Metastatic foci of A549 cells were clearly detected throughout the lung, and the colonization of the metastatic foci gradually progressed over 14 days after the early intensive elimination (Figure 6B). Confocal images allowed the visualization of metastatic foci around alveoli 1 hr after injection, and they showed the appearance of spheres regardless of TGF- $\beta$  stimulation 14 days after injection. Interestingly, statistical analysis revealed that TGF- $\beta$  stimulation significantly increased the number of metastatic foci from 1 to 14 days (Figure 6C).

In addition, the histological investigation after CUBIC-cancer analysis revealed that there was no clear difference between unstimulated and TGF- $\beta$ -pre-stimulated cancer cells in the expression of E-cadherin in metastatic colonies 14 days after injection (Figure 6D). CUBIC-cancer analysis may thus be of value in elucidating unresolved issues regarding the *in vivo* role of the EMT in tumor metastasis.

### CUBIC-Cancer Analysis Enables the Pharmacotherapeutic Profiling of Anti-tumor Drugs against Cancer Metastasis

Finally, we attempted to establish a therapeutic evaluation system for metastatic whole-lung cancers by statistically profiling metastatic cells. To assess chemotherapeutic responses in a syngeneic mouse tumor model, wild-type BALB/c mice were injected with 4T1 cells intravenously and treated with the anti-tumor drugs doxorubicin (Dox), 5-FU, and cyclophosphamide (CPA) (Figure 7A). Eight days after 4T1 injection, the proliferation of metastatic colonies tended to be suppressed in the drug-treated groups (Figures 7B and 7C). In particular, CPA significantly reduced both the total volume and number of colonies. The daily administration of 5-FU resulted in the drastic reduction of metastatic colonies (Figure 7D). Importantly, many metastatic foci were still detected in the lungs of 5-FU-treated mice, indicating the high sensitivity of CUBIC-cancer analysis. Some foci were composed of just a single cancer cell, which might be dormant or resistant to anti-tumor drugs (Figures 7D and 7E). Taken together, CUBIC-cancer analysis can be used to evaluate the *in vivo* effects of anti-tumor drugs at the single-cell level.

## DISCUSSION

### RI-Optimized CUBIC Protocol Enables the Scalable Whole-Body Cell Profiling of Various Cancer Models

The RI values of tissue and cell structures are highly diverse, ranging from 1.34 to 1.59, because the RI for each intact biological tissue depends on the contents and densities of lipids, pro-

teins, and other constituents (Tuchin, 2015). Therefore, optimal RI values were likely to be dependent on each organ. The results of the transmittance evaluation of major organs mounted on various RI media indicated that the transparency of the organs tended to increase as the RI values increased (Figures 1C and S1B). Interestingly, this tendency was observed regardless of the organ, suggesting that these organs mostly contain the constituents with relatively similar RI values. We noted that the RI value-dependent increase in the transparency of most of the organs did not reach a plateau, even at an RI of 1.52. It would be very difficult to prepare hydrophilic RI-matching media with higher RI values (RI > 1.52) using current chemical candidates, because of their water solubility or low RI value per unit mass. Future development of such RI-matching media (RI > 1.52) may be favorable to make these organs transparent more efficiently. Another technical challenge is to develop whole-body-clearing protocols compatible with bone tissue clearing. Mild, effective, and highly permeable decalcification methods compatible with our CUBIC protocol would facilitate our understanding of the mechanisms of cancer metastases, including bone metastasis. We noted that perfusion-assisted agent release *in situ* (PARS) and ultimate 3D imaging of solvent-cleared organs (uDISCO) protocols might be also applicable to whole-body profiling of cancer metastasis with a single-cell resolution (Pan et al., 2016; Susaki et al., 2015; Tanaka et al., 2014; Yang et al., 2014). Therefore, comparison of the characteristics of the proposed CUBIC-cancer protocol in this study with those of these two methods would be important in the future.

### CUBIC-Cancer Analysis Provides a Global Analytical Pipeline of Cancer Metastasis by Bridging the Resolution Gap between Conventional *In Vivo* Bioluminescence Imaging and 2D Histology

To obtain an overview of cancer metastasis, it would be important to create a global analytical pipeline visualizing the overall dynamics of metastatic progression, spatiotemporal distribution of metastatic colonies at single-cell resolution, and their surrounding microenvironments. In this study, we attempted to bridge the resolution gap between *in vivo* bioluminescence imaging and 2D histology by our scalable CUBIC-cancer analysis, using cancer cell lines stably expressing luciferase and fluorescent proteins.

We determined whether fluorescent signals in whole-body and whole-organ images actually originated from cancer cells or indirectly originated from some other cells, such as phagocytes that digested labeled cancer cells. To compare the *in vivo* bioluminescence imaging and CUBIC-cancer analysis, we performed *in vivo* bioluminescence imaging before tissue clearing.

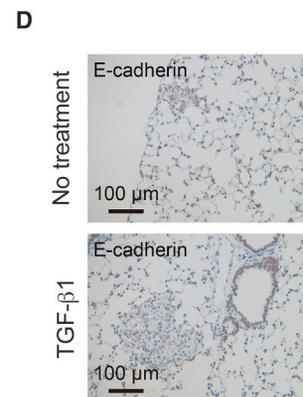
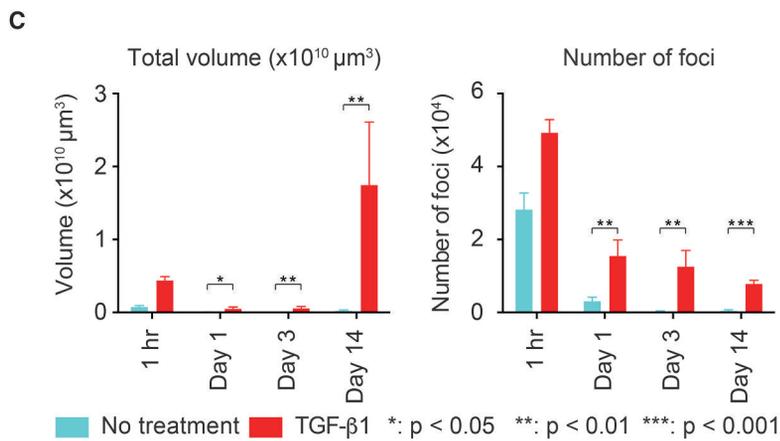
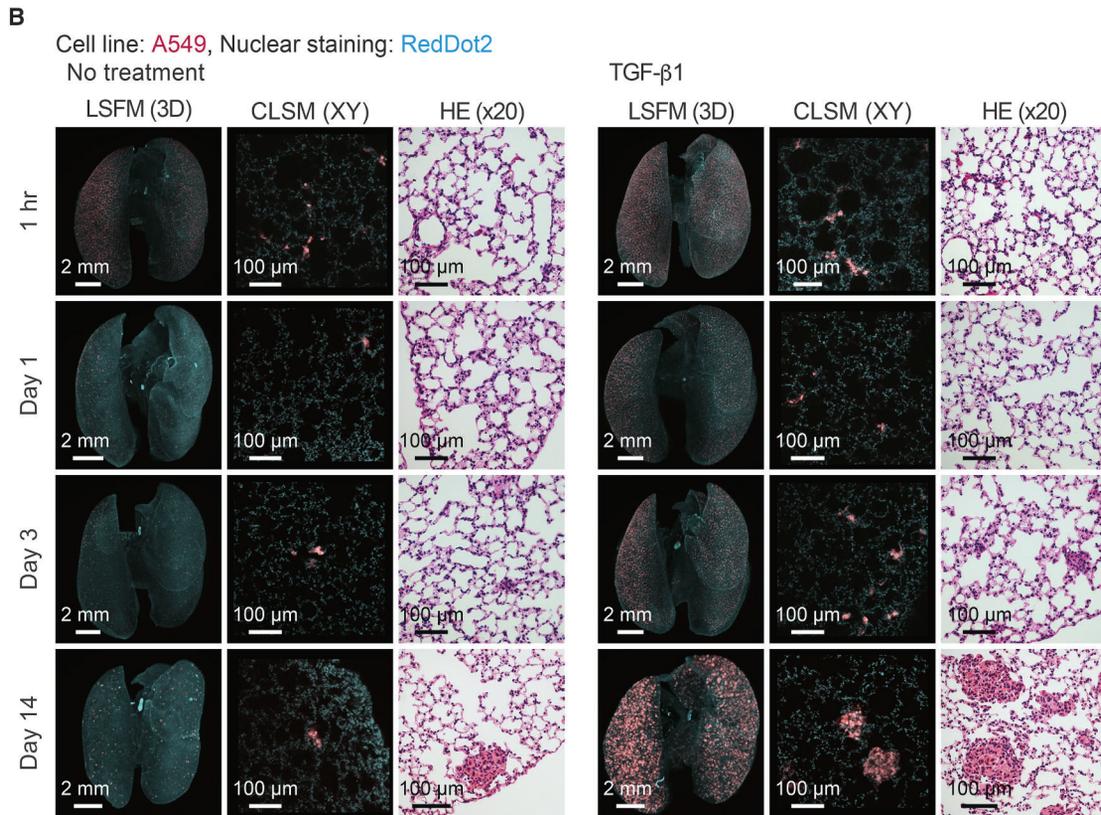
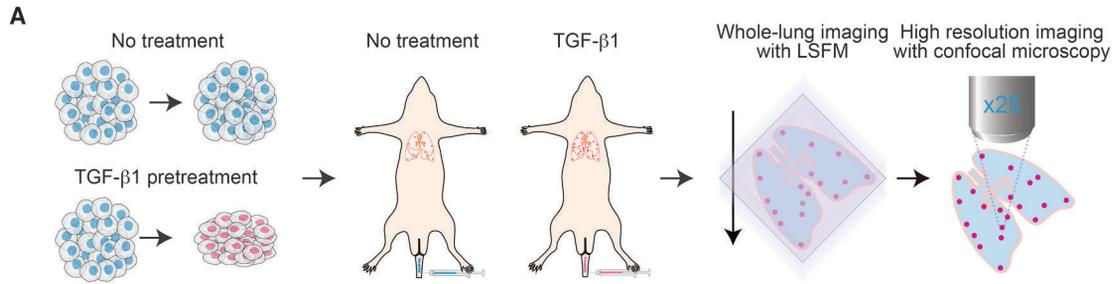
(D and E) Quantification of the metastatic foci in BALB/c-*nu/nu* mice (D) or BALB/c mice (E). The number and the average volume of foci from (B) or (C) are shown. Animal number at each time point is  $n = 3-4$ . Data represent mean  $\pm$  SD.

(F) Quantification of the cell number of metastatic colonies. Spot analysis and surface analysis were applied to the 3D images in (B). Data represent mean  $\pm$  SD.

(G) Quantification of the cell number of cell aggregates observed in BALB/c mice. Spot analysis and surface analysis were applied to the 3D images of day 3 in (C). Data represent mean  $\pm$  SD.

(H) Representative magnified 3D and XY images of the lungs containing the cell aggregates at day 3 after injection in BALB/c mice (left and middle) (MDA-231-D, mCherry; nuclei, RedDot2). The typical HE staining pattern after CUBIC-cancer analysis including the cell aggregates (right most) is shown.

See also Figure S5.



(legend on next page)

Bioluminescence images from luciferase activity roughly corresponded to the fluorescent signals (Figures 2B, 3B, S2C, and S3C), despite apparent differences in multiple organ metastases by Renca cells (Figures 2E and S2A) and brain metastasis by MDA-231-D cells (Figures 4B and S4A). In the former case, bioluminescence signals in metastatic organs may not have been detected with this dynamic range because the signals were too intense at the primary lesion. In the latter case, bioluminescence signals may have been shielded by the skull, and they seemed to leak out to the eyes and neck (Figure S4A). In addition, because we could perform HE staining of CUBIC-treated organs, we compared the CUBIC-cancer analysis results with the data obtained by the conventional histological method. We found that the basic tissue pathology was well preserved in HE staining, and we could detect the tumors in the CUBIC-treated samples (Figures 2B and 6B). However, the histology of the metastatic tumors of OS-RC-2 cells, which are clear cell renal carcinoma and contain high amounts of glycogen, was distinct between the intact samples and samples after CUBIC-cancer analysis, possibly because of the delipidation process (Figure S4B). These observations indicated that certain CUBIC-treated samples can be used for immunohistochemical analysis, although the cell membrane in the samples may be damaged due to delipidation. Together, CUBIC-cancer analysis could successfully bridge the resolution gap between conventional *in vivo* bioluminescence imaging and 2D histology.

### EMT Significantly Promotes the Extravasation or Survival of Cancer Cells *In Vivo*

The EMT contributes to the metastatic process of intravasation in cancer metastasis (Nieto et al., 2016; Polyak and Weinberg, 2009). Cancer cells that have undergone EMT participate in extravasation by enhancing the attachment of cancer cells to metastatic sites through the induction of matrix metalloproteinases or formation of filopodia-like protrusions (Labelle et al., 2011; Shibue et al., 2012). In this study, we investigated the relevance of the EMT in lung metastasis by statistical and temporal analysis of metastatic colonies in an experimental extravasation model (Figure 6). The nature of the *in vivo* dynamics of metastatic extravasation is an ongoing controversy. While the metastatic extravasation was postulated to proceed over 1–2 days (Tsai and Yang, 2013), another group demonstrated using an *in vitro* extravasation assay that cancer cells could extravasate from microvessels within 3 hr (Spiegel et al., 2016). Our results showed that the number of foci at 1 hr after injection was not significantly different, suggesting that the EMT appears to be not involved in the arrest of cells in microvessels. To our surprise,

the number of metastatic foci markedly increased on day 1 after the injection of cells pre-treated with TGF- $\beta$ 1, and this number was approximately sustained up to day 14 (Figure 6C). These findings suggest that the fate of cancer cells in metastatic sites is determined in the very early phase (within a day) after arriving at distant organs, and that the EMT might be associated with the survival of cancer cells at metastatic sites. The EMT is thought to confer stem cell-like features that allow cells to disseminate and gain resistance to apoptosis or anti-tumor therapies (Polyak and Weinberg, 2009). Hence, the significant promotion of metastasis by EMT induction may reflect the transformation of differentiated cells into a stem cell-like phenotype, including the induction of extracellular matrix production and gaining of the resistance to apoptosis.

Plasticity between the epithelial and mesenchymal phenotypes also plays a pivotal role in the process of metastasis (Ye and Weinberg, 2015). In clinical studies, intravasated cancer cells circulating in the blood have been shown to exhibit the mesenchymal phenotype, while colonized cancer cells in organs were found to exhibit the epithelial phenotype (Kallergi et al., 2011; Kowalski et al., 2003). Because of the relatively short-term stimulation of A549 cells by TGF- $\beta$ 1 in the present study, we assumed that the plasticity of the EMT was maintained in this experiment. In fact, our results suggested that A549 cells that colonized the lung were restored to the epithelial phenotype with the restoration of E-cadherin expression. Taken together, our CUBIC-cancer analysis clarified the *in vivo* dynamics of the plasticity between the epithelial and mesenchymal phenotypes during the progression of cancer metastasis through the statistical analysis of metastatic foci at the whole-organ level.

### CUBIC-Cancer Analysis Provides an *In Vivo* Therapeutic Evaluation System to Aid in the Treatment of Metastatic Tumors

To quantitatively evaluate the therapeutic effects of anti-tumor drugs, it is necessary to establish an ultra-sensitive detection system that can determine whether metastatic colonies are completely eliminated or whether single cancer cells resistant to anti-tumor drugs still remain. This distinction is particularly important because the emergence of chemoresistant cancer cells eventually leads to the proliferation of cancer cells and the development of incurable tumors. It is difficult to detect metastasized cells at single-cell resolution by *in vivo* bioluminescence imaging, and visualizing whole-organ images by 2D histology is labor intensive. Notably, our CUBIC-cancer analysis of the entire lung successfully detected tiny foci consisting of even a single cancer cell (Figures 7D and 7E). Therefore, our

#### Figure 6. *In Vivo* Role of the EMT Induced by TGF- $\beta$ Pre-treatment in the Progression of Lung Metastasis

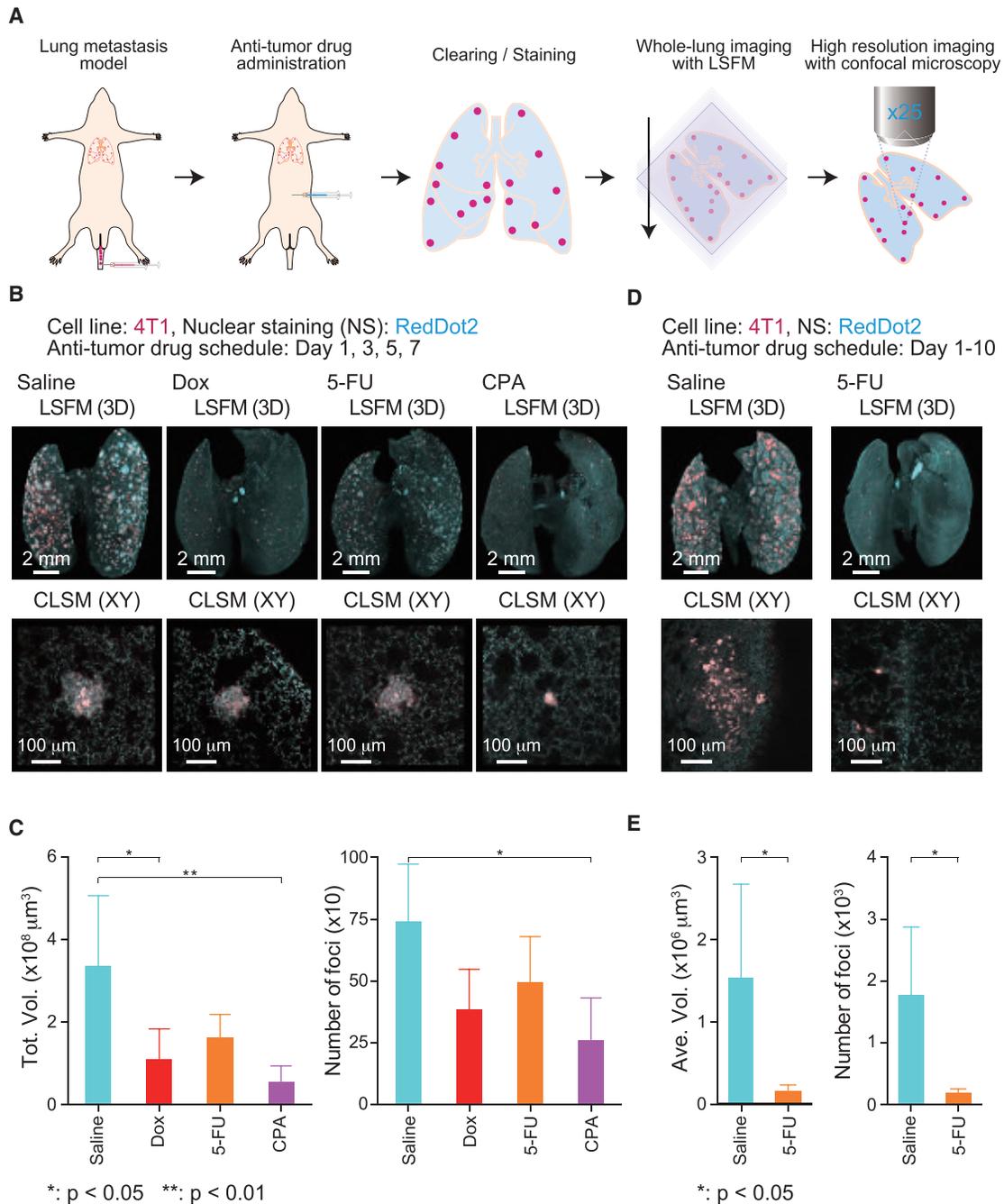
(A) Scheme of the experimental procedure.

(B) Time course images of the experimental lung metastasis model with A549 cells. A549 cells pre-treated without (left) or with (right) TGF- $\beta$ 1 were intravenously injected in BALB/c-*nu/nu* mice. The 3D images of the lung samples at the indicated time points are shown (A549, mCherry; nuclei, RedDot2). Representative images of HE staining after CUBIC-cancer analysis are also shown.

(C) Quantification of the volume and the number of foci of metastatic colonies pre-treated with or without TGF- $\beta$ 1. Surface analysis was applied to the 3D images in (B). Animal number at each time point is  $n = 4$ . Data represent mean  $\pm$  SD.

(D) Immunohistochemistry of E-cadherin. Samples on day 14 after CUBIC-cancer analysis were subjected to immunohistochemistry with anti-E-cadherin antibody.

See also Figure S6.



**Figure 7. Quantitative Evaluation of Therapeutic Effects of Anti-tumor Drugs in an Experimental Lung Metastatic Model**

(A) Scheme of the experimental procedure.

(B) In vivo therapeutic efficacy of anti-tumor drugs. The 3D images of the lung samples are shown (4T1, mCherry; nuclei, RedDot2).

(C) Quantification of pharmacotherapeutic effects of anti-tumor drugs. Animal number of each group is  $n = 4$ . Data represent mean  $\pm$  SD.

(D) In vivo therapeutic evaluation of once daily treatment with 5-FU. The 3D images of the lung samples are shown (4T1, mCherry; nuclei, RedDot2).

(E) Quantification of pharmacotherapeutic effect of once daily treatment with 5-FU. Animal number of each group is  $n = 4$  (5-FU) and  $n = 5$  (saline). Data represent mean  $\pm$  SD.

CUBIC-cancer analysis may provide critical information for the development of a curative treatment for metastasis.

Furthermore, we wish to emphasize another important advantage of CUBIC-cancer analysis: it allows a seamless connection

between in vivo live imaging and 2D histology, complementing the shortcomings of each modality. The global analytical pipeline consisting of these methods is quite robust and reliable, because the experimental data could be thoroughly verified

by the three modalities. An *in vivo* therapeutic evaluation system is urgently needed not only for malignant neoplastic diseases but also for autoimmune diseases, which remain a challenge clinically, and emerging induced pluripotent stem cell (iPSC)-based regenerative therapies. This analytical pipeline would have great potential for becoming a *de facto* standard for an *in vivo* therapeutic evaluation system for complex systemic diseases.

## EXPERIMENTAL PROCEDURES

Details are also supplied in the Supplemental Experimental Procedures.

### Preparation of Clearing Solutions

Clearing solutions were composed of four chemicals, which were selected from chemical screening (T.C. Murakami and K. Tainaka, unpublished data). CUBIC-L for decolorization and delipidation was prepared as a mixture of 10 w% polyethylene glycol mono-*p*-isooctylphenyl ether/Triton X-100 (12967-45, Nacalai Tesque) and 10 w% *N*-butyl-diethanolamine (B0725, Tokyo Chemical Industry). CUBIC-R for RI matching was prepared as a mixture of 45 w% 2,3-dimethyl-1-phenyl-5-pyrazolone/antipyrene (D1876, Tokyo Chemical Industry) and 30 w% nicotinamide (N0078, Tokyo Chemical Industry) (Table S2). See also the Supplemental Experimental Procedures.

### Microscopy

Whole-body and whole-organ images were acquired with a custom-build LSMF (developed by Olympus). High-resolution images for cell profiling were acquired with CLSM (FLUOVIEW FV1200, Olympus). The RI-matched sample was immersed in a mixture of silicon oil HIVAC-F4 (Shin-Etsu Chemical) and mineral oil (RI = 1.467, M8410, Sigma-Aldrich) during image acquisition. See also the Supplemental Experimental Procedures.

### Experimental and Spontaneous Mouse Metastatic Tumor Models

Inbred wild-type BALB/c mice, C57BL/6 mice, and BALB/*c-nu/nu* mice were purchased from Sankyo Labo Service. All experiments were approved and carried out according to the Animal Care and Use Committee of the Graduate School of Medicine, The University of Tokyo. Each metastasis model is described briefly. For experimental liver metastasis, BALB/*c-nu/nu* mice (5 weeks old, female) were injected with Panc-1 cells by open injection in the spleen. For peritoneal dissemination, SUIT-2 cells were injected into pancreas orthotopically or intraperitoneally in BALB/*c-nu/nu* mice (5 weeks old female). For experimental lung metastasis by intravenous injection, BALB/*c-nu/nu* (5 weeks old, female), BALB/c mice (5 weeks old, female or male), or C57BL/6 mice (5 weeks old, female) were injected with each cell line (MDA-MB-231, MDA-231-D, A549, 4T1, Renca, and B16F10 cells). For spontaneous lung metastasis by renal subcapsule injection, BALB/c (5 weeks old, male) or BALB/*c-nu/nu* (5 weeks old, male) mice were injected with Renca, OS-RC-2, or Caki-1 cells orthotopically. For experimental brain metastasis by intracardiac injection, BALB/*c-nu/nu* (4 weeks old, female MDA-231-D, male OS-RC-2) mice were injected with MDA-231-D or OS-RC-2 cells by puncture into the left ventricle of heart. For the stimulation with TGF- $\beta$ , A549 cells were pre-treated with or without TGF- $\beta$ 1 (5 ng/mL) for 72 hr in culture and subjected to *in vivo* experiment. See also the Supplemental Experimental Procedures.

### Statistical Analysis

An unpaired *t* test was used to compare the index of volume per surface area of brain metastasis and to compare the pharmacotherapeutic effects of 5-FU for statistical significance. Multiple *t* tests were used to compare the effect of TGF- $\beta$ 1 stimulation for statistical significance. Dunnett's multiple comparisons test was used to examine the pharmacotherapeutic effects of anti-tumor drugs for statistical significance. All statistical analyses were performed with GraphPad Prism6 software.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and four tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.06.010>.

## AUTHOR CONTRIBUTIONS

H.R.U., K.M., S.I.K., K. Takahashi, K. Tainaka, and S.E. designed the study. S.I.K. performed most of the imaging and analysis of metastasis model mice. K. Takahashi and J.N. performed most of the establishment of cancer cell lines for CUBIC-cancer analyses and preparation of metastasis model mice. Y.M. contributed to the histological experiments. H.R.U., K.M., S.I.K., K. Takahashi, K. Tainaka, and S.E. wrote the manuscript. All authors discussed the results and commented on the manuscript text.

## ACKNOWLEDGMENTS

We thank all lab members at The University of Tokyo and RIKEN QBiC, in particular, A. Millius for editing, T.C. Murakami and C. Shimizu for chemical screening for tissue clearing, S. Kitamura and K. Nakamura for developing optimized delipidation medium, M. Fujiwara for developing RI-optimized medium, E.A. Susaki and A. Kuno for developing 3D staining methods, S. Shoi for helping with statistical analysis, R. Ohno for housing mice, and Bitplane for instruction on Imaris 8.1.2. We also thank Y. Katsuno and F. Murai for plasmid preparation and H. Miyoshi (RIKEN) for providing the lentiviral vectors. This work was supported by grants from AMED-CREST (Japan Agency for Medical Research and Development [AMED]/Ministry of Education, Culture, Sports, Science and Technology, Japan [MEXT], H.R.U.); CREST (JST/MEXT, H.R.U.); Brain/MINDS (AMED/MEXT, H.R.U.); Basic Science and Platform Technology Program for Innovative Biological Medicine (AMED/MEXT, H.R.U.); Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (JSPS) (KAKENHI grant 14J01180 to S.I.K., grant 17J30007 to S.I.K., grant 13J10991 to K. Takahashi, and grant 16J05993 to J.N.), Grant-in-Aid for Challenging Exploratory Research (JSPS KAKENHI grant 16K15124 to K. Tainaka), Grants-in-Aid for Scientific Research (S) (JSPS KAKENHI grants 25221004 to H.R.U. and 15H05774 to K.M.), and Grants-in-Aid for Scientific Research on Innovative Areas (JSPS KAKENHI grant 23115006 to H.R.U.); and Brain Protein Aging and Dementia Control (17H05688 to K. Tainaka) from MEXT; a grant for Practical Research for Innovative Cancer Control (16ck0106193h0001) from AMED (K.M.); and a Specific Research Grant from The Tokyo Society of Medical Sciences (S.E.). T.C. Murakami, K. Tainaka, and H.R.U. have filed a patent application for CUBIC-reagents. A part of this study was done in collaboration with the Olympus Corporation.

Received: March 24, 2017

Revised: May 2, 2017

Accepted: May 29, 2017

Published: July 5, 2017

## REFERENCES

- Carbonell, W.S., Ansoorge, O., Sibson, N., and Muschel, R. (2009). The vascular basement membrane as "soil" in brain metastasis. *PLoS ONE* 4, e5857.
- Chung, K., Wallace, J., Kim, S.Y., Kalyanasundaram, S., Andalman, A.S., Davidson, T.J., Mirzabekov, J.J., Zalocusky, K.A., Mattis, J., Denisin, A.K., et al. (2013). Structural and molecular interrogation of intact biological systems. *Nature* 497, 332–337.
- Cuccarese, M.F., Dubach, J.M., Pfirschke, C., Engblom, C., Garris, C., Miller, M.A., Pittet, M.J., and Weissleder, R. (2017). Heterogeneity of macrophage infiltration and therapeutic response in lung carcinoma revealed by 3D organ imaging. *Nat. Commun.* 8, 14293.
- Dotd, H.U., Leischner, U., Schierloh, A., Jährling, N., Mauch, C.P., Deininger, K., Deussing, J.M., Eder, M., Ziegglänsberger, W., and Becker, K. (2007).

- Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain. *Nat. Methods* 4, 331–336.
- Ertürk, A., Becker, K., Jährling, N., Mauch, C.P., Hojer, C.D., Egen, J.G., Hellal, F., Bradke, F., Sheng, M., and Dodt, H.U. (2012). Three-dimensional imaging of solvent-cleared organs using 3DISCO. *Nat. Protoc.* 7, 1983–1995.
- Gavrillovic, I.T., and Posner, J.B. (2005). Brain metastases: epidemiology and pathophysiology. *J. Neurooncol.* 75, 5–14.
- Hama, H., Kurokawa, H., Kawano, H., Ando, R., Shimogori, T., Noda, H., Fukami, K., Sakae-Sawano, A., and Miyawaki, A. (2011). Scale: a chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. *Nat. Neurosci.* 14, 1481–1488.
- Johnsen, S., and Widder, E.A. (1999). The physical basis of transparency in biological tissue: ultrastructure and the minimization of light scattering. *J. Theor. Biol.* 199, 181–198.
- Kallergi, G., Papadaki, M.A., Politaki, E., Mavroudis, D., Georgoulas, V., and Agelaki, S. (2011). Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients. *Breast Cancer Res.* 13, R59.
- Ke, M.T., Fujimoto, S., and Imai, T. (2013). SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nat. Neurosci.* 16, 1154–1161.
- Kienast, Y., von Baumgarten, L., Fuhrmann, M., Klinkert, W.E., Goldbrunner, R., Herms, J., and Winkler, F. (2010). Real-time imaging reveals the single steps of brain metastasis formation. *Nat. Med.* 16, 116–122.
- Kowalski, P.J., Rubin, M.A., and Kleer, C.G. (2003). E-cadherin expression in primary carcinomas of the breast and its distant metastases. *Breast Cancer Res.* 5, R217–R222.
- Labelle, M., Begum, S., and Hynes, R.O. (2011). Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* 20, 576–590.
- Nicholls, A.G. (1927). The fourth chapter in the history of pathology. *Can. Med. Assoc. J.* 17, 843–845.
- Nieto, M.A., Huang, R.Y., Jackson, R.A., and Thiery, J.P. (2016). EMT: 2016. *Cell* 166, 21–45.
- Nussbaum, E.S., Djalilian, H.R., Cho, K.H., and Hall, W.A. (1996). Brain metastases. Histology, multiplicity, surgery, and survival. *Cancer* 78, 1781–1788.
- Pan, C., Cai, R., Quacquarelli, F.P., Ghasemigharagoz, A., Loubopoulos, A., Matryba, P., Plesnila, N., Dichgans, M., Hellal, F., and Ertürk, A. (2016). Shrinkage-mediated imaging of entire organs and organisms using uDISCO. *Nat. Methods* 13, 859–867.
- Polyak, K., and Weinberg, R.A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat. Rev. Cancer* 9, 265–273.
- Renier, N., Wu, Z., Simon, D.J., Yang, J., Ariel, P., and Tessier-Lavigne, M. (2014). iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* 159, 896–910.
- Saadatpour, Z., Rezaei, A., Ebrahimnejad, H., Baghaei, B., Bjorklund, G., Chartrand, M., Sahebkar, A., Morovati, H., Mirzaei, H.R., and Mirzaei, H. (2017). Imaging techniques: new avenues in cancer gene and cell therapy. *Cancer Gene Ther.* 24, 1–5.
- Saito, R.A., Watabe, T., Horiguchi, K., Kohyama, T., Saitoh, M., Nagase, T., and Miyazono, K. (2009). Thyroid transcription factor-1 inhibits transforming growth factor-beta-mediated epithelial-to-mesenchymal transition in lung adenocarcinoma cells. *Cancer Res.* 69, 2783–2791.
- Shibue, T., Brooks, M.W., Inan, M.F., Reinhardt, F., and Weinberg, R.A. (2012). The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. *Cancer Discov.* 2, 706–721.
- Spalteholz, W. (1914). Über das Durchsichtigmachen von menschlichen und tierischen Präparaten und seine theoretischen Bedingungen, nebst Anhang (Leipzig: S Hierzel), pp. 1–93.
- Spiegel, A., Brooks, M.W., Houshyar, S., Reinhardt, F., Ardolino, M., Fessler, E., Chen, M.B., Krall, J.A., DeCock, J., Zervantonakis, I.K., et al. (2016). Neutrophils suppress intraluminal NK cell-mediated tumor cell clearance and enhance extravasation of disseminated carcinoma cells. *Cancer Discov.* 6, 630–649.
- Susaki, E.A., and Ueda, H.R. (2016). Whole-body and whole-organ clearing and imaging techniques with single-cell resolution: Toward organism-level systems biology in mammals. *Cell Chem. Biol.* 23, 137–157.
- Susaki, E.A., Tainaka, K., Perrin, D., Kishino, F., Tawara, T., Watanabe, T.M., Yokoyama, C., Onoe, H., Eguchi, M., Yamaguchi, S., et al. (2014). Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* 157, 726–739.
- Susaki, E.A., Tainaka, K., Perrin, D., Yukinaga, H., Kuno, A., and Ueda, H.R. (2015). Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nat. Protoc.* 10, 1709–1727.
- Tainaka, K., Kubota, S.I., Suyama, T.Q., Susaki, E.A., Perrin, D., Ukai-Tadenuma, M., Ukai, H., and Ueda, H.R. (2014). Whole-body imaging with single-cell resolution by tissue decolorization. *Cell* 159, 911–924.
- Tainaka, K., Kuno, A., Kubota, S.I., Murakami, T., and Ueda, H.R. (2016). Chemical principles in tissue clearing and staining protocols for whole-body cell profiling. *Annu. Rev. Cell Dev. Biol.* 32, 713–741.
- Takahashi, K., Nagai, N., Ogura, K., Tsuneyama, K., Saiki, I., Irimura, T., and Hayakawa, Y. (2015). Mammary tissue microenvironment determines T cell-dependent breast cancer-associated inflammation. *Cancer Sci.* 106, 867–874.
- Tatsuki, F., Sunagawa, G.A., Shi, S., Susaki, E.A., Yukinaga, H., Perrin, D., Sumiyama, K., Ukai-Tadenuma, M., Fujishima, H., Ohno, R., et al. (2016). Involvement of Ca(2+)-dependent hyperpolarization in sleep duration in mammals. *Neuron* 90, 70–85.
- Tsai, J.H., and Yang, J. (2013). Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes Dev.* 27, 2192–2206.
- Tuchin, V.V. (2015). Tissue optics and photonics: Light-tissue interaction. *J. Biomed. Photonics Eng.* 1, 98–135.
- Yang, B., Treweek, J.B., Kulkarni, R.P., Deverman, B.E., Chen, C.K., Lubeck, E., Shah, S., Cai, L., and Gradinaru, V. (2014). Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* 158, 945–958.
- Ye, X., and Weinberg, R.A. (2015). Epithelial-mesenchymal plasticity: A central regulator of cancer progression. *Trends Cell Biol.* 25, 675–686.

# Major Shifts in Glial Regional Identity Are a Transcriptional Hallmark of Human Brain Aging

Lilach Soreq,<sup>1,2</sup> UK Brain Expression Consortium, North American Brain Expression Consortium, Jamie Rose,<sup>3</sup> Eyal Soreq,<sup>4</sup> John Hardy,<sup>1,5</sup> Daniah Trabzuni,<sup>1,6</sup> Mark R. Cookson,<sup>7</sup> Colin Smith,<sup>3</sup> Mina Ryten,<sup>1,9</sup> Rickie Patani,<sup>1,2,5,8,10,\*</sup> and Jernej Ule<sup>1,2,11,\*</sup>

<sup>1</sup>Institute of Neurology, University College London, London WC1N 3BG, UK

<sup>2</sup>The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

<sup>3</sup>MRC Edinburgh Brain Bank, Academic Neuropathology, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh EH16 4SB, UK

<sup>4</sup>The Computational, Cognitive and Clinical NeuroImaging Laboratory, Division of Brain Sciences, Imperial College, London SW7 2AZ, UK

<sup>5</sup>Reta Lila Weston Institute of Neurological Studies, UCL ION, 1 Wakefield Street, London WC1N 1PJ, UK

<sup>6</sup>Departments of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh 12713, Saudi Arabia

<sup>7</sup>Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda, MD 20892, USA

<sup>8</sup>Euan MacDonald Centre for MND, University of Edinburgh, Edinburgh EH8 9YL, UK

<sup>9</sup>Department of Medical and Molecular Genetics, King's College London, Guy's Hospital, Great Maze Pond, London SE1 9RT, UK

<sup>10</sup>Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 1TN, UK

<sup>11</sup>Lead Contact

\*Correspondence: rickie.patani@ucl.ac.uk (R.P.), j.ule@ucl.ac.uk (J.U.)

<http://dx.doi.org/10.1016/j.celrep.2016.12.011>

## SUMMARY

Gene expression studies suggest that aging of the human brain is determined by a complex interplay of molecular events, although both its region- and cell-type-specific consequences remain poorly understood. Here, we extensively characterized aging-altered gene expression changes across ten human brain regions from 480 individuals ranging in age from 16 to 106 years. We show that astrocyte- and oligodendrocyte-specific genes, but not neuron-specific genes, shift their regional expression patterns upon aging, particularly in the hippocampus and substantia nigra, while the expression of microglia- and endothelial-specific genes increase in all brain regions. In line with these changes, high-resolution immunohistochemistry demonstrated decreased numbers of oligodendrocytes and of neuronal subpopulations in the aging brain cortex. Finally, glial-specific genes predict age with greater precision than neuron-specific genes, thus highlighting the need for greater mechanistic understanding of neuron-glia interactions in aging and late-life diseases.

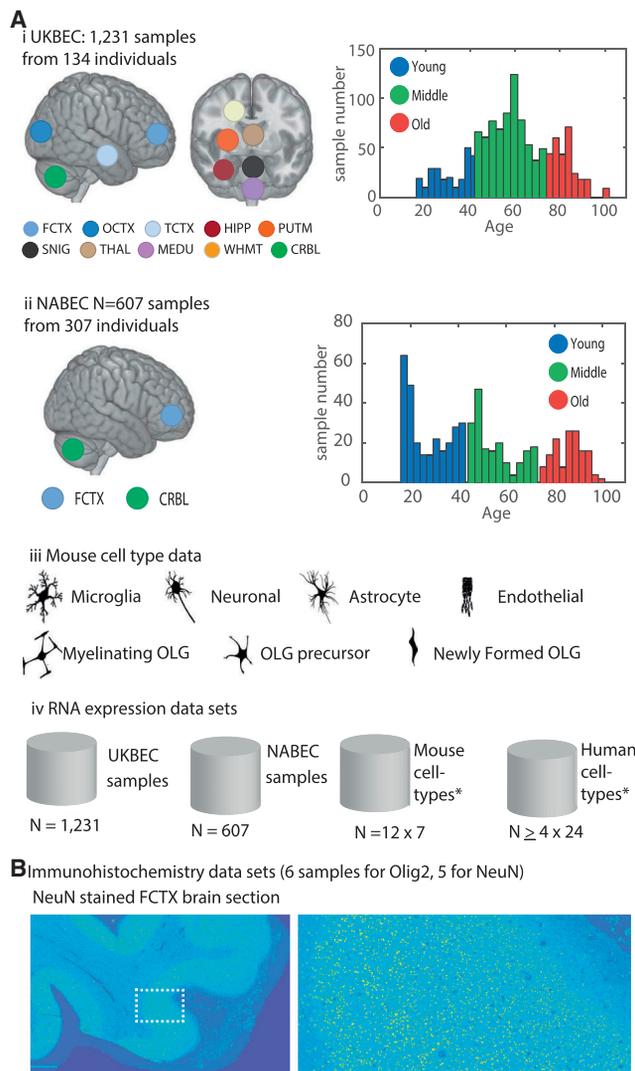
## INTRODUCTION

Aging, an inevitable time-dependent functional decline, is present in all living organisms. The intimate relationship between aging and neurodegeneration raises the possibility of shared transcriptional and post-transcriptional gene regulation programs; however, we still lack a comprehensive transcriptome-

wide picture of the effects of aging across different human brain regions and cell types (De Strooper and Karran, 2016). RNA expression profiling of the aging brain has been studied historically using a limited number of brain regions in animal models or human post-mortem tissues. A major unrealized goal therefore remains a comprehensive characterization of the transcriptional landscape across multiple human brain regions in a physiological age range, which may provide insights into the cellular architecture and molecular pathways of aging.

The unparalleled complexity of the human brain is a function of its structural and functional cellular diversity, which arises from tightly regulated transcriptional programs. Limited availability to human post-mortem samples has hampered comprehensive transcriptomic analysis of the brain, particularly of region- and cell-type-specific diversity. However, through international collaboration, a comprehensive atlas of the brain's transcriptome based on samples from two individuals (the Allen Brain Atlas) has been achieved. This study illustrated how transcripts of genes involved in different pathways are expressed across the brain, but the potential effect of age on the regional differences was not examined.

By current consensus, astrocyte (AC) and neuronal numbers appear generally preserved in aging (Fabricius et al., 2013; Matarin et al., 2015; Pelvig et al., 2008). It is clear, however, that Alzheimer's disease (AD) and other neurodegenerative diseases for which age is a major risk factor are associated with inflammatory changes mediated by microglia (MG) (Cribbs et al., 2012; Frank et al., 2008). Brain aging includes accumulation of senescent MG, altered signaling, and pro-inflammatory phenotypes (Mosher and Wyss-Coray, 2014), and it was shown that MG display regional sensitivity to aging (Streit and Xue, 2010). Immune-related changes were also strongly associated with aging in mouse models of amyloid pathology (Matarin et al., 2015). Nevertheless, animal models and human tissue have reported



**Figure 1. Analyzed Samples and Datasets**

(A) The samples of the UKBEC and NABEC datasets were divided into three age groups each (young: 16–44, middle: 45–74, old:  $\geq 75$ ). (i) The main analyzed dataset (UKBEC) is composed of 1,231 brain samples interrogated by exon microarrays, from brain samples of 134 individuals from 16 to 102 years old and up to ten brain regions each. The brain regions included both cortical and sub-cortical regions, specifically: the frontal cortex (FCTX), temporal cortex (TCTX), occipital cortex (OCTX), intralobular white matter (WHMT), cerebellum (CRBL), substantia nigra (SNIG), putamen (PUTM), thalamus (THAL), hippocampus (HIPPO), and medulla (MEDU) for UKBEC and the FCTX and CRBL for NABEC. (ii) The independent (NABEC) dataset of brain samples from FCTX and CRBL 307 individuals (16–101 years old). (iii) In addition, seven cell types were identified based on analysis of available RNA-seq data from mice cortex ([http://web.stanford.edu/group/barres\\_lab/brain\\_maseq.html](http://web.stanford.edu/group/barres_lab/brain_maseq.html)). (iv) A summary of all expression data used in this study. The total number of samples described in (i)–(iii) is listed, as well as the human RNA-seq analysis of 24 CNS human cell types (Table S7) ([http://web.stanford.edu/group/barres\\_lab/brainseqMariko/brainseq2.html](http://web.stanford.edu/group/barres_lab/brainseqMariko/brainseq2.html)).

(B) High-resolution immunohistochemical imaging dataset was produced from samples of young and three old FCTX from the UKBEC cohort, following staining by OLIG2 antibody and computational analysis for the quantification of the OLG cell population. Staining by NeuN of FCTX sections from the same brain samples followed by targeted computational analysis was conducted for

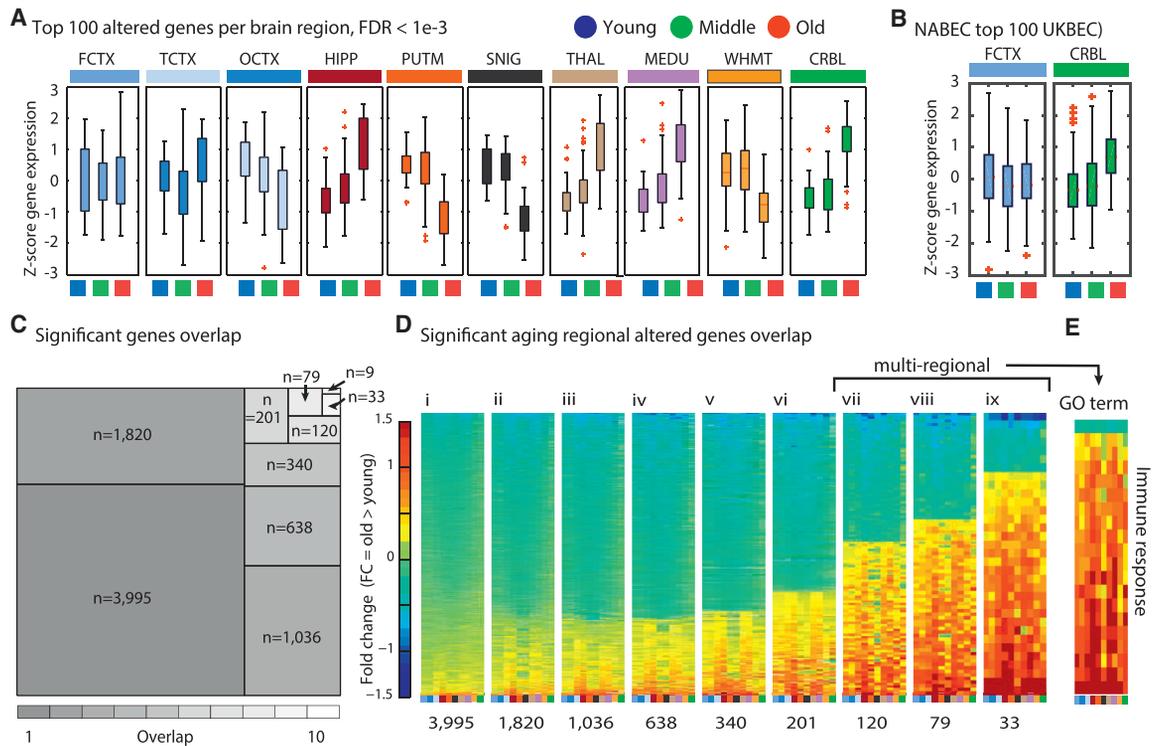
variable and apparently contrasting alterations in ACs (reactivity or atrophy) and MG (MHC class II antigen increase or atrophy) (Cerbai et al., 2012; Streit and Xue, 2010; Tremblay et al., 2012). Accumulation of oligodendrocytes (OLGs) was previously reported in aging monkey cortex (Peters and Sethares, 2004), while stereological quantification of glia in neocortical regions of old brains has suggested a reduction in the number of OLGs, as evident by a >3-fold greater atrophy of the sub-cortical white matter (WHMT) compared to cortical regions and an age-determined loss of myelin (Head et al., 2004; Vernooij et al., 2008). Furthermore, MG-mediated neuroinflammation has been described as a common hallmark of both AD and Parkinson’s disease (PD) and is believed to be mechanistically important in driving pathogenesis (Orre et al., 2013; Perry and Teeling, 2013). Collectively, these findings suggest that the field stands to benefit from systematic and comprehensive analysis of aging-related changes in the cellular and molecular composition of the human brain.

Apart from the study of region-dependent microglial response to aging, the importance of both region- and cell-type-specific changes in the aging brain remains poorly understood. Studies have been hampered by the limited availability of cross-regional post-mortem tissue across a range of ages. To overcome these limitations, we analyzed gene expression patterns in ten brain regions (including cortical and sub-cortical areas) using more than 1,800 brain samples from two large independent cohorts, representing the most comprehensive human aging brain gene expression analysis to date. We report striking changes in cell-type-specific expression patterns across different brain regions, which revealed major shifts in glial regional identity upon aging in the human brain.

## RESULTS

In this study, we examined two extensive gene expression datasets from post-mortem human samples and sampled multiple (up to ten) brain regions per individual. The primary dataset was produced by the UK Brain Expression Consortium (UKBEC) and included 1,231 tissue samples collected from 134 adult individuals between 16 and 102 years old, with each contributing post-mortem samples of up to ten brain regions (Figure 1Ai). The brain regions included both cortical and sub-cortical regions, specifically the frontal cortex (FCTX), temporal cortex (TCTX), occipital cortex (OCTX), intralobular white matter (WHMT), cerebellum (CRBL), substantia nigra (SNIG), putamen (PUTM), thalamus (THAL), hippocampus (HIPPO), and medulla (MEDU). The second dataset, which allowed independent external cross-validation, was produced by the North American Brain Expression Consortium (NABEC) (Gibbs et al., 2010; Kumar et al., 2013), including 307 samples from two brain regions (age range: 16 to 101 years old; Figure 1Aii). The third dataset was also used for validation including samples with an age range of 27 to 106 years old (Lu et al., 2014). None of the brain samples had neuropathological evidence of diagnosable degenerative

quantification of the neuronal cell population (an example of one of the NeuN stained sections is shown on the right, in the zoomed-in view of the area marked on the left-hand side). OLG, oligodendrocyte.



**Figure 2. Multi-regional Aging-Altered Genes Are Mainly Upregulated**

(A) The direction of expression change of the top 100 genes detected as significantly differentially expressed upon aging in each of the studied expression datasets from ten UKBEC brain regions (ANOVA test significance threshold:  $FDR < 1e-3$ ; the test compared the three defined age groups). (B) Age-group based separation of 607 FCTX and CRBL samples (the NABEC cohort) was based on measured expression of the nine cross-regional genes. (C) A tree map of the number of genes that were altered upon aging, dependent on the number of brain regions where the change is observed. (D) Fold change of the genes that were altered upon aging, separated into heatmaps dependent on the number of brain regions where the change is observed. (E) Fold change of the multi-regional genes that were enriched in the Gene Ontology term immune response (standardized Z score; range is as shown for the heatmaps on the left). Brain region abbreviations are explained in the legend to Figure 1A. See also Figure S1A for the total number of aging-altered genes per region.

diseases (Table S1). To detect differentially expressed genes, we assigned each sample to one of three age groups (young: 16–44, middle: 45–74, old:  $\geq 75$  years old) and applied a collection of tailored data-mining computational approaches (Figure 1). We excluded gender-based sample separation to specifically identify the effects of age on gene expression profiles.

### Region-Specific and Global Transcriptional “Signatures” of the Aging Human Brain

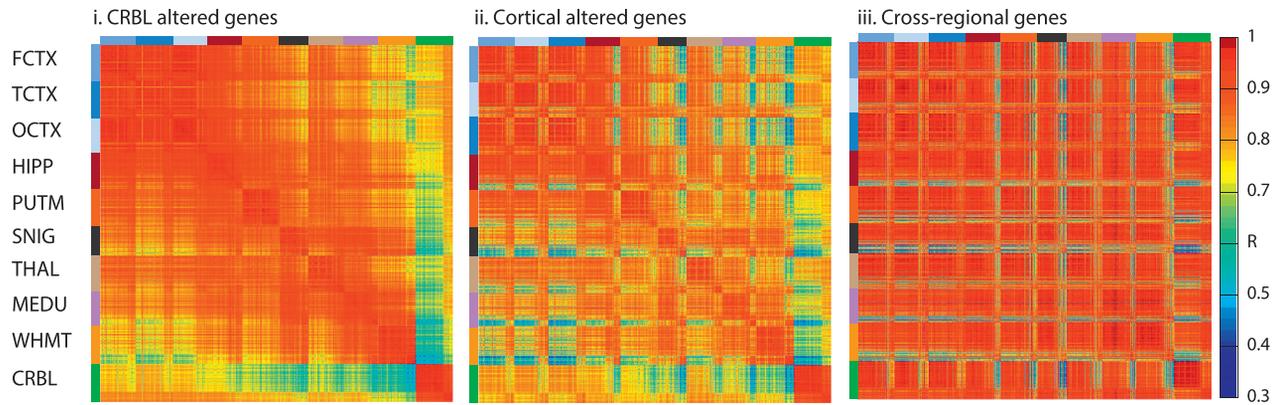
We first sought to address whether region-specific differences in gene expression patterns occur within the brain upon aging. Both the number of differentially expressed genes (threshold: false discovery rate [FDR]  $< 1e-3$ ) (Figure S1) and the direction of expression change varied in a region-specific manner (Figure 2A). The general directions of gene expression change were preserved in the independent NABEC dataset (Figure 2B). We applied a stringent threshold to enable isolation of global changes across the UKBEC brain regions. Most changes were specific for one region (hereafter referred to as “region specific”) or a few regions (hereafter referred to as “region selective,” including genes altered in two to seven regions), while some genes were altered in eight or more brain regions (hereafter

referred to as “multi-regional”), and nine genes were found to be significantly altered in all ten brain regions upon aging ( $FDR < 1e-3$ , hereafter referred to as “cross-regional”) (Figure 2C; Figure S2). The rates and number of overlapping age-altered genes varied between pairs of brain regions (Figure S2). Multi-regional genes predominantly exhibited increased expression levels upon aging (Figures 2Dviii–2Dix). This group of genes was enriched in the Gene Ontology (GO) term “immune response”, which had the general trend to be upregulated in aging (Figure 2E).

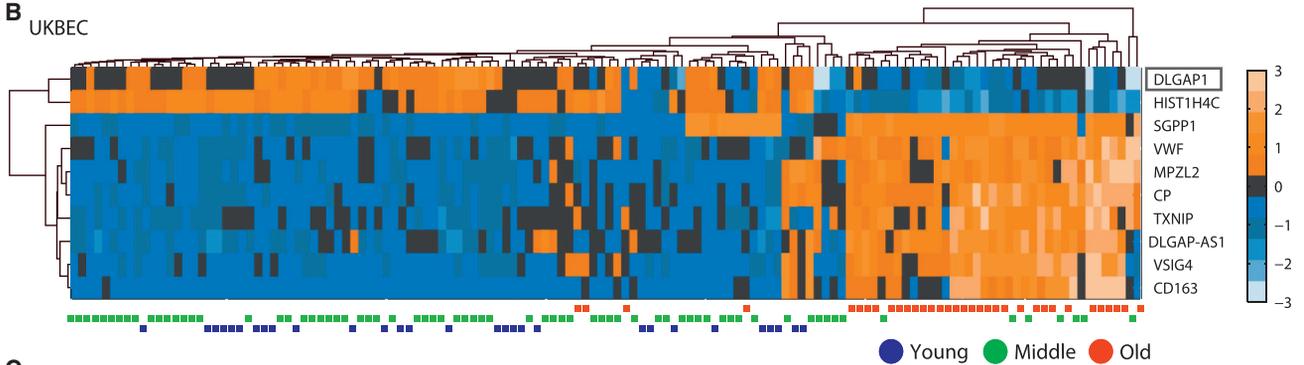
### Nine Cross-Regional Aging-Altered Genes Accurately Predict the Age Categories

To assess the ability of different groups of genes to classify the brain samples in both cohorts by brain region and age group criteria, we next used a non-linear dimension reduction classification method. The expression of regional-selective genes separated the samples well based on their regional identity (Figure S2). In addition, genes with aging-altered expression patterns in the CRBL, WHMT, or cortical samples generally have distinctive regional expression, as evident by sample-to-sample correlation scores that were computed among each of the 1,231

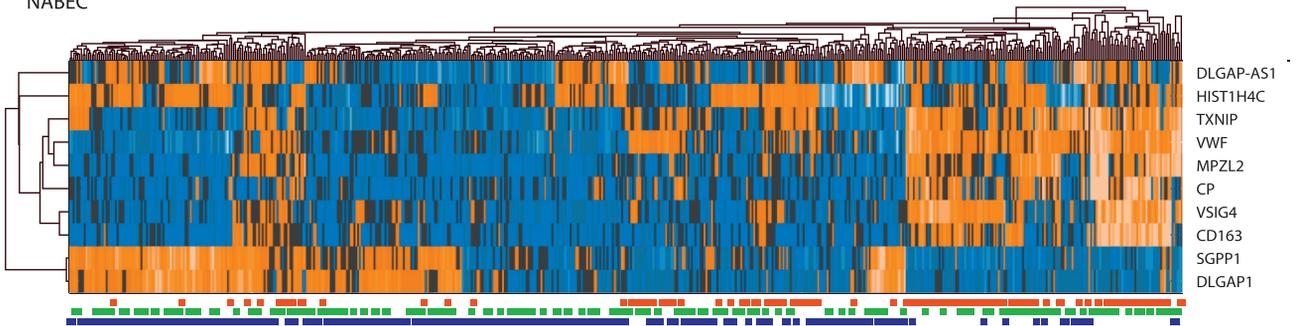
**A** Sample to sample cross correlation



**B** UKBEC

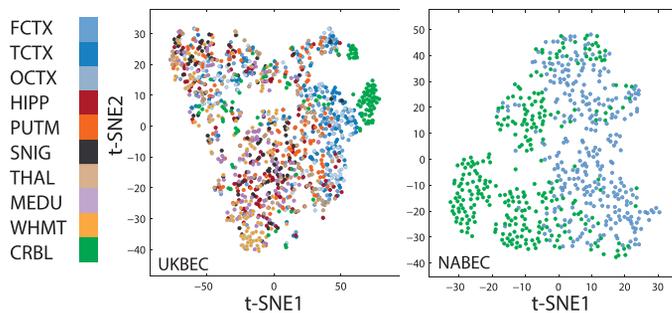


**C** NABEC

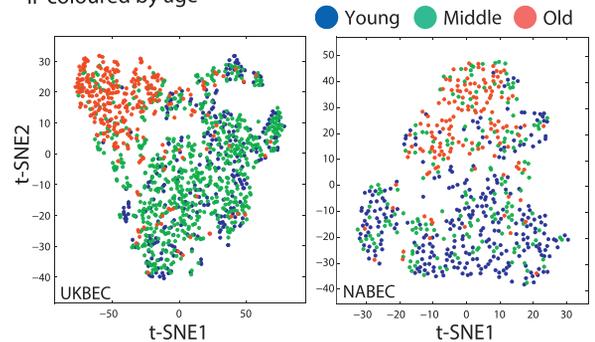


**D**

i coloured by tissue



ii coloured by age



(legend on next page)

brain samples based on their expression signals (Figures 3Ai and Aii; Figure S1B). The 642 CRBL-altered genes showed high inter-regional correlation in expression patterns with the cortical regions (including the HIPP), the 265 cortical aging-altered genes revealed high correlation among the cortical regions and HIPP samples, and the 801 WHMT-altered genes showed increased correlation specifically within CRBL and WHMT.

In contrast to region-specific aging-altered genes, expression patterns of these cross-regional genes were correlated among samples only based on age, rather than brain region (Figure 3Aiii). Eight of the pan-regional genes were upregulated, and one (*HIST1H4C*) was downregulated upon aging (Figure 3B; Table S5). The cross-regional genes successfully discriminated the samples based on age group in most UKBEC cases (130 of 134) (Figure 3B). One of these genes was a non-coding RNA (*DLGAP1-AS1*) that is antisense to the protein-coding gene (*DLGAP1*). We identified a robust, reciprocal relationship in expression between *DLGAP1* and *DLGAP1-AS1* upon aging. *DLGAP1* is highly brain-specific, while *DLGAP1-AS1* and *DLGAP1-AS2* are normally expressed in internal organs and the bone marrow (Gene Cards database) (Harel et al., 2009). We find that in contrast to upregulation of *DLGAP1-AS*, expression of *DLGAP1* shows an aging-altered decrease across brain regions, despite not reaching statistical significance in each region (Figure 3B). This demonstrates coupling between the age-dependent decrease in the expression of the brain-specific protein-coding gene *DLGAP1* and the increase of its antisense RNA, which is otherwise only expressed outside of the brain.

As a sign of the validity of the cross-regional genes, they were efficient in classifying samples of the independent NABEC dataset based on age, even though the NABEC data were not used to identify these genes (Figure 3C). Moreover, a non-linear classification based on these genes separated samples belonging to young, middle-age, and old groups in the UKBEC, the NABEC, and an additional independent FCTX brain expression cohort (Lu et al., 2014) (Figure 3D; Figure S3C). This model also verified the prediction of age group based on expression levels in the NABEC cohort (Figure S3A). To exclude the effects of other variables, we show that the UKBEC samples were not classified by gender (Figure S3B). The cross-regional genes also correctly classified samples by age group in an additional independent dataset of cortical samples (Lu et al., 2014; Figure S3) and did not classify the UKBEC dataset by gender (Figure S3).

### Major Shifts in Region-Specific Expression Profiles of Glia-Specific Genes in the Aging Brain

To investigate the biological relevance of the age-related gene expression changes, we first examined the expression profiles of the cross-regional genes in recently produced RNA sequencing (RNA-seq) data from seven purified mouse brain cell types (Zhang et al., 2014). All cross-regional genes were expressed in a cell-type-specific manner, in particular within glial cells and mainly in MG and OLGs (Figure S4A). We therefore further examined the cell-type-specific expression patterns of all aging-altered genes. For this, we calculated genome-wide expression scores to identify genes specific for each cell type (Table S3). We then examined whether expression of cell-type-specific genes was altered upon aging. For three cell types (neurons, ACs, and OLGs), enrichment data previously generated by microarrays was also available (Cahoy et al., 2008). We selected genes that were demonstrated to be specific by both our defined cell-specific lists found by analysis of RNA-seq data (Zhang et al., 2014) and the published microarray cell-specific lists from mice (Cahoy et al., 2008), in addition to being altered in aging. In agreement with the great diversity of neuronal cell types across different brain regions, neuron-specific genes were most enriched among the regional aging-altered genes (Figure S4B). In contrast, glia-specific genes were most enriched among the multi- and cross-regional altered genes, and this was most pronounced for OLG precursors and MG (Figure S4B).

We first sought to investigate the changes in expression of MG-specific genes, because these cells have been most extensively linked to aging so far (Erraji-Benchekroun et al., 2005; Sibille, 2013). Consistent with previous studies (Erraji-Benchekroun et al., 2005; Sibille, 2013), most MG-specific aging-altered genes had low expression in brain samples from the young group but strongly increased their expression in all regions in the old group samples (Figure S5A). Our study extends these findings to a multi-regional phenomenon. Moreover, we find that a small number of MG-specific genes have high absolute expression in the young group but decreased expression upon aging in all regions. This suggests that the change in expression not only reflects an increased number of MG but most likely also includes a dramatic change in the MG gene expression program. Although most brain regions were not strongly separated by the MG-specific aging-altered genes, the CRBL was distinct (Figure 4A). This complies with a report of a distinct MG expression profile in the CRBL of young mice (Grabert et al., 2016).

### Figure 3. The Nine Cross-Regional Genes Discriminate Samples Based on Age

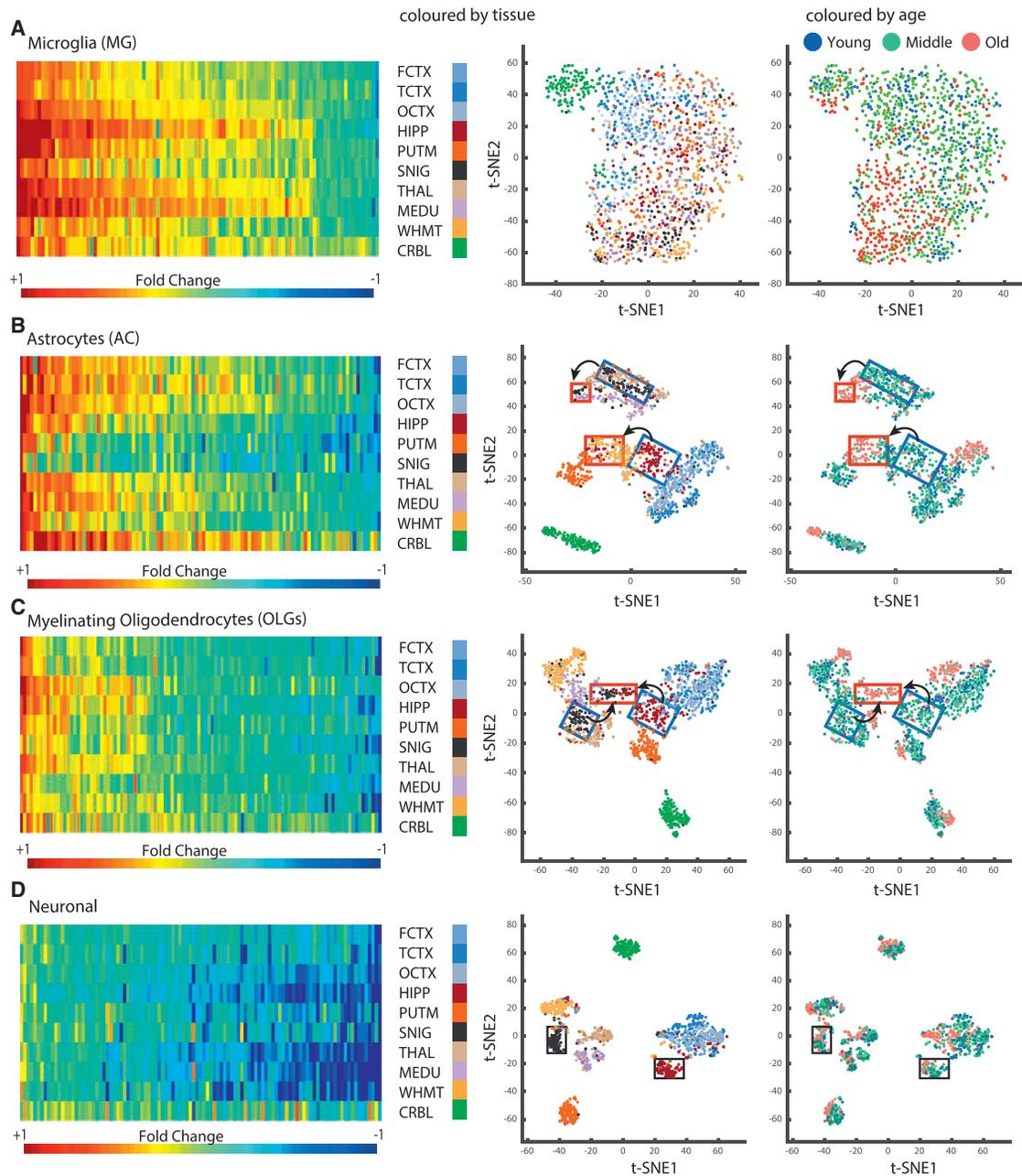
(A) Correlation scores were calculated between each pair of brain samples among the UKBEC samples based on different lists of aging-altered genes using Spearman correlation. (i) Correlation scores based on CRBL aging-altered genes. (ii) Correlation scores based on cortical aging-altered genes. (iii) Correlation scores based on cross-regional aging-altered genes. See Figure S1B for correlation based on WHMT-altered genes.

(B) Hierarchical classification of the UKBEC cohort based on the expression signals of the cross-regional altered genes and the *DLGAP* antisense. Rows, genes; columns, samples. Age group is denoted in blue (young: 16–44 years), green (middle-age: 45–74 years), or red (old:  $\geq 75$  years). The dendrograms show the Euclidian distance measured for both rows and columns. Right color bar, standardized fold change (*Z* score; range:  $-3$  to  $3$ ; orange represents increased expression in aging, and blue denotes decrease).

(C) Hierarchical classification of the NABEC expression dataset based on the profiles of the cross-regional genes and the *DLGAP* antisense.

(D) Non-linear dimensionality reduction by t-distributed stochastic neighbor embedding (t-SNE) is based on the expression of the nine cross-regional genes, with the x axis showing t-SNE1 and the y axis showing t-SNE2. Either the ten UKBEC brain regions or the two NABEC brain regions (FCTX and CRBL) are classified, as marked in the plots. (i) Each sample is colored based on its corresponding tissue (colors are marked on the left). (ii) The same samples are colored based on their age group (colors are marked on the top).

Brain region abbreviations are explained in the legend to Figure 1A.



**Figure 4. Gliia-Specific Genes Show Major Shifts in Regional Identity upon Aging**

On the left, heatmaps show the fold change between old and young groups in the expression of the top 100 aging-altered cell-type-specific genes across regions (the color bar corresponds to the standardized Z score, with blue corresponding to decrease and red to increase in gene expression; range:  $-1$  to  $+1$ ). On the right, non-linear dimensionality reduction by t-distributed stochastic neighbor embedding (t-SNE) is used to classify a sample of the ten UKBEC brain regions based on the expression of the top 20 aging-altered cell-specific genes, with the x axis showing t-SNE1 and the y axis showing t-SNE2. In the first plot, each sample is colored based on its corresponding tissue (colors are marked on the left of the plot), and in the second plot, the same samples are colored based on their age group (colors are marked on the top of the plots).

(A) Sample classification based on the aging-altered MG-specific genes.

(B) Sample classification based on the aging-altered AC-specific genes.

(C) Sample classification based on aging-altered myelinating OLG-specific genes.

(D) Sample classification based on the expression signals of aging-altered neuron-specific genes. The SNIG and PUTM samples are marked by rectangles as an example of the loss of region-specific expression upon aging for OLG- and AC-specific genes.

Brain region abbreviations are explained in the legend to Figure 1A. MG, microglia; AC, astrocyte; OLG, oligodendrocyte. See Figure S2 for sample classification based on region-specific genes compared with multi-regional genes and Figure S6 for heatmaps and classification plots based on the three cell-type microarray gene markers.

Apart from the CRBL samples that formed a separate cluster, samples from the old group clustered together for all other regions and separately from younger samples, indicating that MG-specific gene expression is more defined by age than by regional identity (Figure 4A). The aging-altered endothelial-specific genes showed a similar pattern of changes as the MG-specific genes, with a general upregulation across all brain regions and a notable age group separation and lack of clear regional identity (Figure S5D).

Next, we examined the expression profiles of genes specific for either ACs or OLGs. In young samples, we observed much higher absolute expression of AC-specific genes in the midbrain regions compared to cortex and HIPP (Figure S5). However, AC-specific genes increase their expression within cortical regions and exhibit decreased expression in basal ganglia (BG) upon aging (Figure 4B); therefore, their absolute expression signals become more similar across regions in the aging brain (Figure S5A). SNIG and THAL, which show the highest expression of AC-specific genes in the young brain, have a generic decrease of AC-specific genes upon aging (Figure 4B). In contrast, the AC-specific genes with the lowest expression in the young group increase their expression upon aging in all regions except SNIG and THAL. This leads to remarkable shifts in the regional patterns of AC-specific gene expression. Although expression of AC-specific genes clustered most brain regions separately for the young group, only four regional clusters remained in the old group, two of which were the CRBL and the cortical regions (Figure 4B). The most pronounced change is seen for HIPP and SNIG. For example, HIPP clusters close to cortex in the young group but shifts toward the WHMT and PUTM in the samples from the old group (Figure 4B).

The aging-altered genes that are specific for all stages of OLG differentiation, including OLG precursors, newly formed OLGs, and myelinating OLG generally show a trend toward decreased expression in all regions upon aging (Figures S5A–S5C). Moreover, OLG-specific genes show a shift of region-specific gene expression upon aging, with the strongest change of regional identity seen in HIPP and SNIG. In the samples from the young group, HIPP clusters close to cortical samples and SNIG clusters close to MEDU and THAL, whereas in the samples from the old group, HIPP and SNIG cluster closer to each other (Figure 4C).

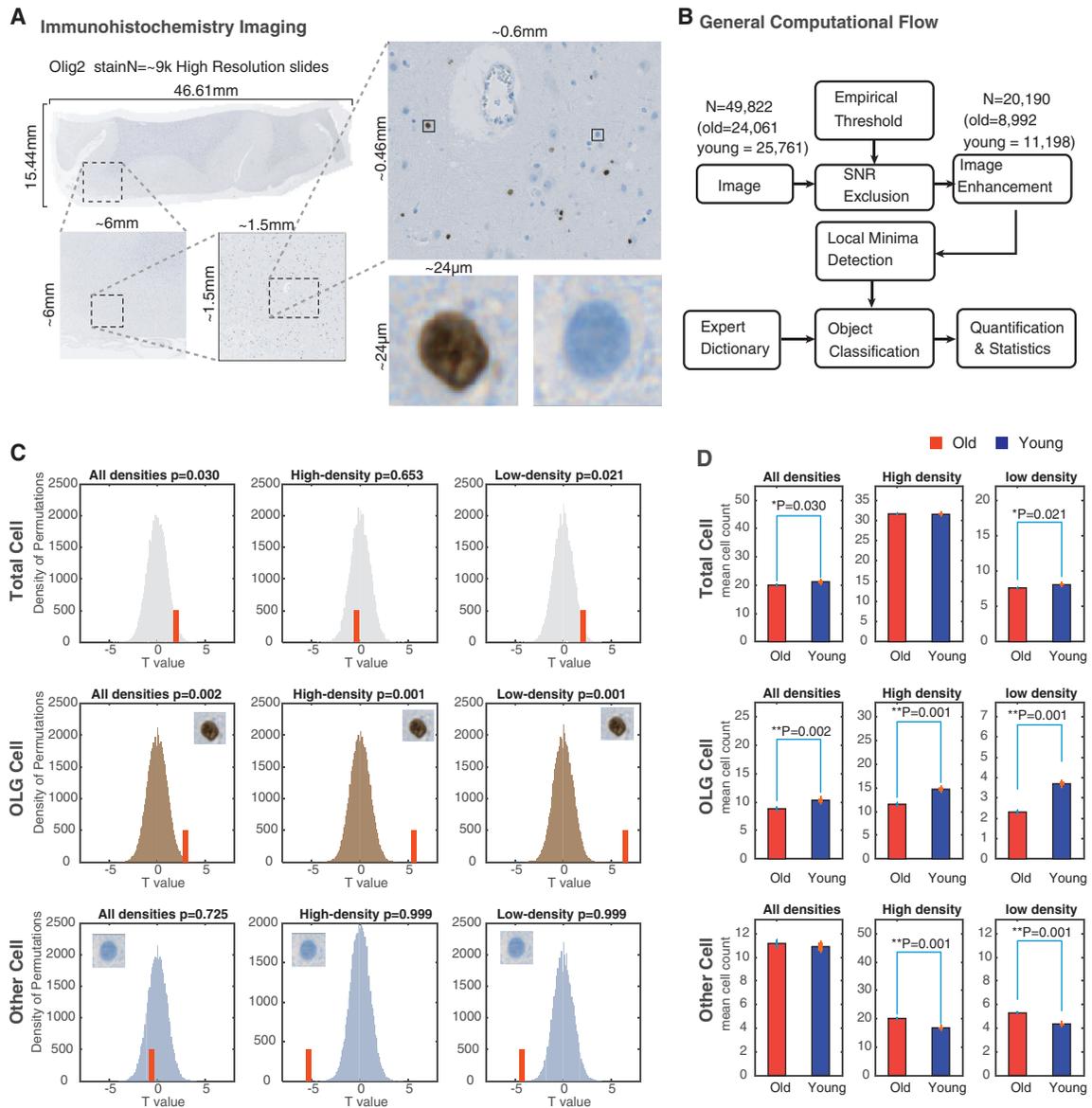
Similar to the OLG-specific genes, the aging-altered neuron-specific genes showed predominant downregulation in all brain regions upon aging (Figure D; Figure S5B), in agreement with previous studies that observed decreased expression of neuron-specific genes in the cortex (Erraji-Benchekroun et al., 2005), but with the added insight that this occurs in a brain-wide manner. Classification based on aging-altered neuron-specific genes yielded a striking separation of samples based on their regional identity (Figure 4D), and the old group samples remained clustered closest to the young samples of the same brain region (Figure 4D). Thus, neuron-specific gene expression is more defined by regional identity than by age. This agrees with the finding that the downregulated genes, which are often neuron specific, are also generally region-specific, while the upregulated genes, which are often MG specific, are generally multi-regional (Figure 2D; Figure S5B).

### Specific Neuronal Subpopulations and Oligodendrocytes Are Decreased in the Aging Brain

To examine how gene expression changes may relate to changes in brain cell populations, we developed an efficient pipeline for analysis of high-resolution image tiles of immunolabeled sections of FCTX. We developed a targeted computational pipeline for detection and quantification of the stained cells based on the scanned images, which consists of big data detection, segmentation, and quantification pipeline using thresholding, filtering, and object detection.

Noting the trend for decreased expression of OLG-specific genes in the frontal cortex (Figure 4D), we examined serial sections immunolabeled with OLIG2 antibody from the tissue blocks from the same brain samples that were used for the microarray study. We selected three young and three old cases based on their microarray profiles, such that it was representative for their age (see Supplemental Experimental Procedures). Approximately 50,000 image tiles were analyzed from the three young and three old FCTX sections (Figure 5A). We counted the number of OLIG2-positive cells compared to the total number of cell nuclei in each tile. Statistics was calculated on two classes of tile density, likely corresponding to local variations in the proportion of white matter (low density of nuclei) and gray matter (high density of nuclei), in addition to all densities combined (all tiles). The number of OLIG2-positive cells decreased in all classes of tiles in the aging FCTX, with the largest decrease in the low-density tiles in old compared to young cases (Figure 5C, middle panel). In contrast, the number of other cells significantly increased low-density tiles (Figure 5C, lower panel), in agreement with the slight increase in the total number of cells in the same tiles (Figure 5C, upper panel). This analysis demonstrates that the decreased expression of OLG-specific genes might partly reflect a decrease in three cortical OLG cell population.

The aging RNA expression signatures also revealed downregulation of neuron-specific genes (Figure 4B); therefore, we analyzed high-resolution images produced from the three young and three old FCTX sections stained with NeuN antibody to mark the neurons. This antibody detects the neuron-specific RNA-binding protein RBFOX3, which is predominantly nuclear, but is also present in the cytoplasm of the cell body (Kim et al., 2009) (Figure 6A). We used the tissue samples from the same cases as were used for OLIG2 quantification, thus allowing direct comparison of the two cell types. To capture the large diversity of both shape and size of cell bodies in the neuronal populations, we used a large tile size (10,000 × 10,000 pixels each) (Figure 6Bi). This allowed us to extract information from almost all layers of the neocortex in each slide of gray matter. A preliminary quality control analysis flagged one image from a young individual as a technical outlier, and this sample was therefore omitted from further analysis, although we provide access to its data (<https://figshare.com/s/f2675361af1242f3565f>). We processed 1,044 image tiles using our cell detection pipeline and applied an information exclusion criterion (entropy > 5) to contain the most meaningful slides (n = 641). In an attempt to enrich the regions of gray matter with the highest information content, we further focused on the 184 tiles with the highest density of nuclei. A total of 371,096 neurons were identified. We further segregated cells into four bins of total area of cell



### Figure 5. Decreased Counts of Oligodendrocytes in the Frontal Cortex upon Aging

Six FCTX brain sections were stained and imaged (from three old and three young post-mortem brain samples). Each sample contains thousands of equal-size slides each  $1,600 \times 1,200$  pixels, as captured by a Zeiss AxioScan slide scanner following staining with the Olig2 antibody.

(A) An example of a BA9-Olig2 slide shown in a full-resolution pyramid, with gradual zooming into two typical cells: one stained brown (OLG cells) and one stained blue (other cells).

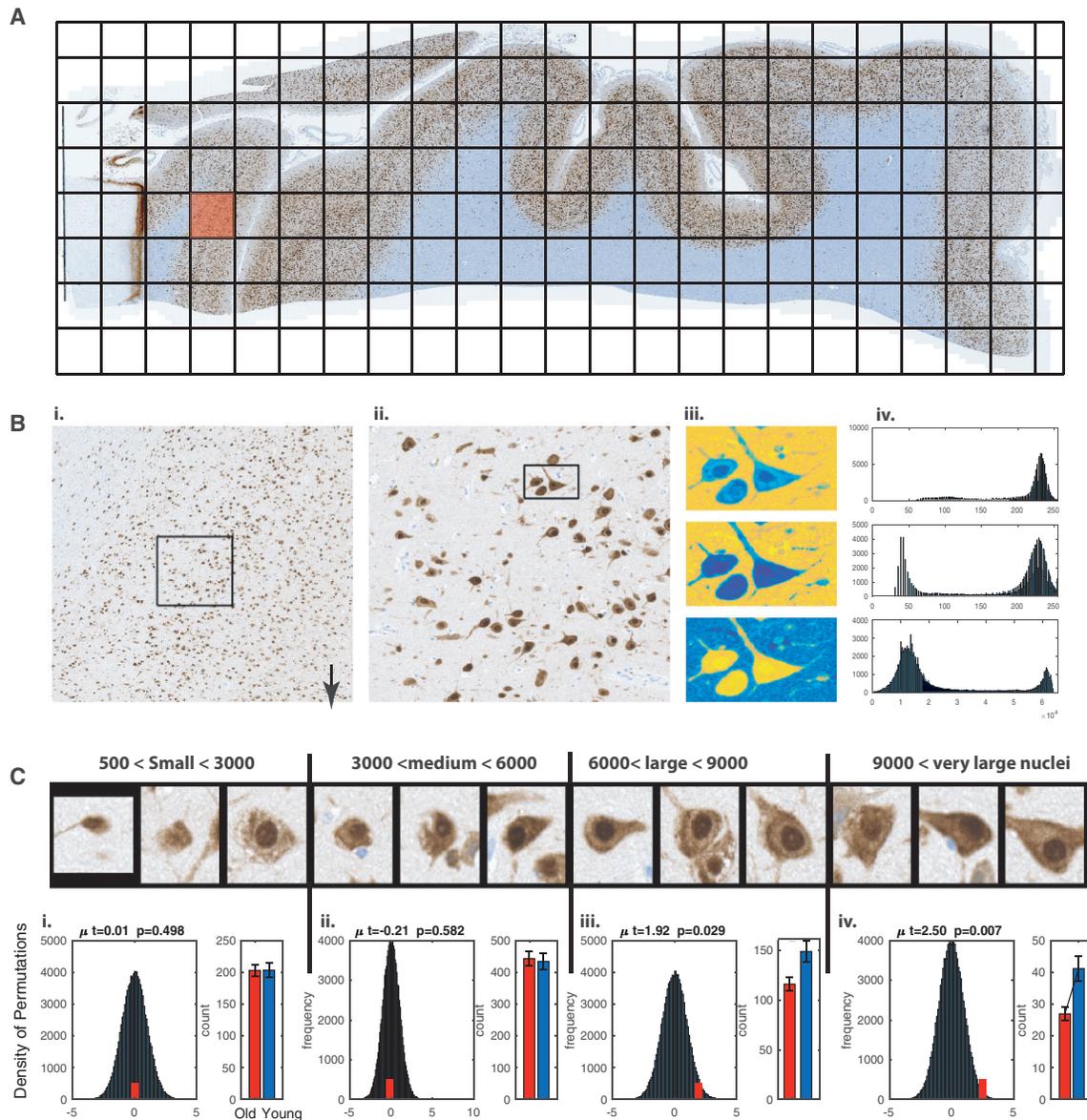
(B) General computational pipeline for the analysis of high-resolution immunohistochemical high-dimensional imaging data allowed us to quantify both OLG and other cells in each FCTX slide.

(C) Comparison of OLG counts that asks if the number of cells of interest is different in young samples compared to old (i.e., red bar shifted to the right means increased count in young samples). In each panel, the histogram represents the null distribution of *t* values calculated using two-tailed Student's *t* test over slide cell counts randomly sampled from the entire population of the six samples, using 100 random iterations over 500 permutations where the true-label *t* statistics is depicted with a red bar, and the remaining distribution was calculated based on shuffled labels. The analysis was done on overall 8,766 young and 10,922 old group slides (left). From a total of 2,612 young and 1,828 old group high-density slides, the 50 slides with the highest density were selected per case for quantification. Similarly, from 1,154 young and 1,277 old group low-density slides, the 50 slides with the lowest density were chosen per sample for quantification.

(D) Cell counts in samples from old (red) and young (blue) groups, with significance calculated with *t* statistics as described in (C). The star marks bars with a *p* value < 0.05 and the mean *T* statistic, *p* value and SD of the permutation test are reported on top of the graphs.

body (small: 500–3,000 pixels,  $n = 92,947$ ; medium: 3,000–6,000 pixels,  $n = 202,239$ ; large: 6,000–9,000 pixels,  $n = 60,314$ ; very large: >9,000 pixels,  $n = 15,596$ ). In agreement with the previous

study (Kim et al., 2009), the intensity of cytoplasmic NeuN signal was strongest in the largest cells (Figure 6C). To account for the imbalance of tiles across samples, significance in the



**Figure 6. Decreased Counts of Specific Neuronal Populations in the Frontal Cortex upon Aging**

(A) An image of one NeuN-stained FCTX section, with re-defined tiles demonstrated by black rectangles (file size = 37.4 GB).

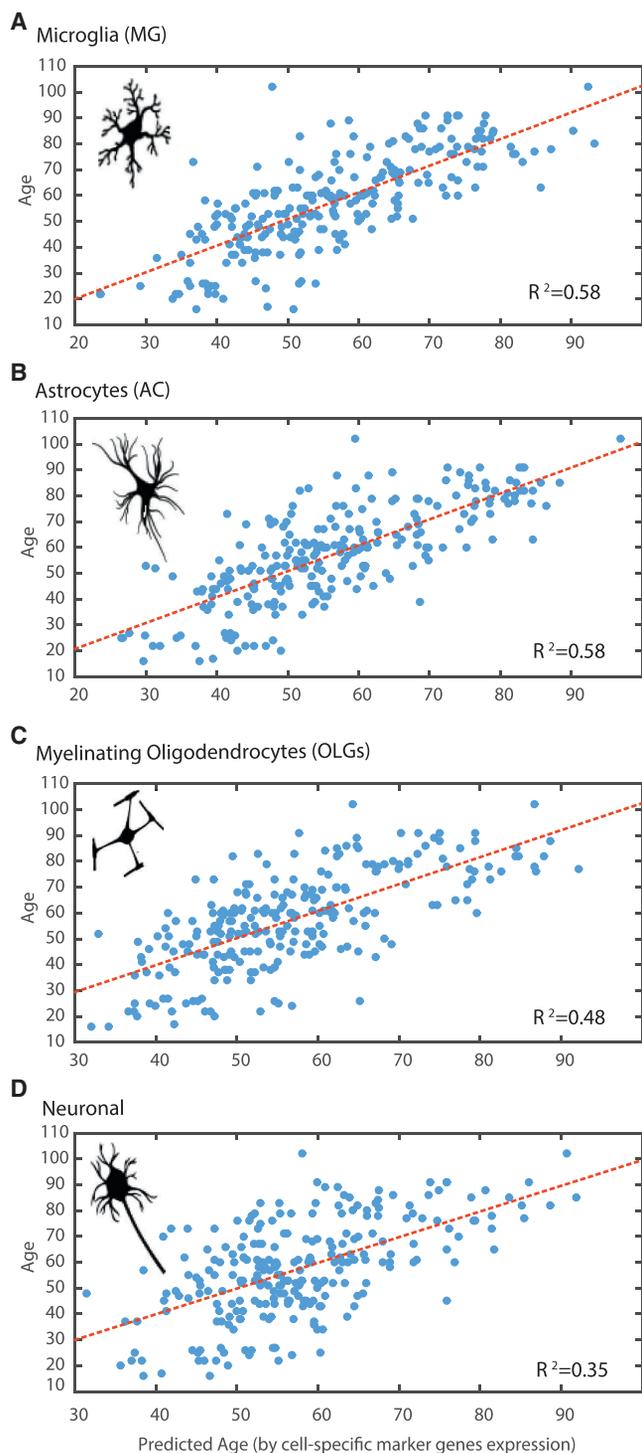
(B) (i) Enlargement of a single tile of 10,000 × 10,000 pixels (size = 225 MB). (ii) Enlargement of a 2,500 × 2,500 pixel section. (iii) Three cells as observed in the red channel (top, shown in light blue) and blue channel (middle), and intersection of the two channels (bottom) differentiates between neuronal cells (stained by NeuN in brown on the original slides) and other cells (stained by Hematoxylin in blue in the middle plot). (iv) The x axis represents the color frequency distribution of the red and blue channels across an intensity range of 256 gray levels, while the y axis represents the frequency of pixel intensity in the image tile depicted in (iii).

(C) Examples of detected neurons that contain small, medium, medium to large (E), or large cell body (F) with size given in pixels. Underneath each image is the histogram that asks if the number of cells of interest is different in young samples compared to old (i.e., red bar shifted to the right means increased count in young samples). The histogram shows the null distribution of t values, calculated using two-tailed Student's t test over slide counts using 100,000 random permutations from the entire population of the six samples (black bars), while the mean of the true-label t statistics is depicted with a red bar. The right graph shows the cell counts in samples from old (red) and young (blue) groups, with significance calculated with t statistics based on 10,000 random permutations. The star marks bars with a p value < 0.05, and the mean t statistics, p value, and SD of the permutation test are reported on top of the graphs.

age-dependent decrease in each neuronal population was tested with right-tailed two-sample t test. We observed no change in the number of neurons with small or medium-size cell bodies, which represent 80% of detected neurons. In contrast, the number of neurons with large or very large cell

bodies is significantly decreased ( $p = 0.029$  and  $p = 0.007$ , accordingly; right-tailed two-sample t test) Figure 6C).

Altogether, it is likely that changes in gene expression observed in our study reflect a combination of changes in expression profiles and changes in quantity of specific cell types.



**Figure 7. Glial-Specific Genes Are Most Capable of Predicting Biological Age**

(A–D) Analysis of the accuracy of cell-type-specific genes in predicting the biological age of UKBEC brain samples: (A) MG-specific genes ( $R^2 = 0.58$ ), (B) AC-specific genes ( $R^2 = 0.58$ ), (C) neuron-specific genes ( $R^2 = 0.35$ ), (D) OLG precursor-specific genes ( $R^2 = 0.48$ ).

In all plots, the y axis denotes the actual age and the x axis denotes the predicted age. MG, microglia; AC, astrocytes; OLG, oligodendrocytes. See Fig-

ure S4C for age association plots of endothelial, OLG precursor, and newly formed OLG cell-specific genes.

### MG- and Endothelial-Specific Genes Are the Best Predictors of Biological Age

While most neuron-specific genes are predominantly downregulated, the extent of this downregulation varies among brain regions. All other cell types have a more complex pattern of changes. These changes appear most pronounced in HIPP and SNIG, which show the strongest shifts in the regional expression pattern of AC- and OLG-specific genes upon aging.

Given our aforementioned findings, we next sought to gain insight into whether glial genes can predict age category with more fidelity than neurons and sought to understand the precise nature of gene expression changes driving this. Having established cross-regional and cell-type-specific gene expression relationships upon aging, we next asked which cell-type expression patterns within the brain are most associated with age. We applied a stepwise regression to construct an age-associative model based on the expression signals of cell-type-specific genes to compare them (Figure 7; Figure S4; Supplemental Experimental Procedures). We accounted for both the brain bank source and the cause of death. Application of the model uncovered defined groups of the highest age-relevant genes for each cell type. In a few cases, interaction between two genes was found to be age predictive (Table S3). Several multi-regional genes were also detected as age-predictive cell-type-specific genes, including *CP*, *SGPP1*, and *VWF*, which were detected as OLG or endothelial specific (Table S4). MG-, AC-, and endothelial-specific genes were most highly associated with biological age, while the smallest number of age-predictive genes was found among the neuron-specific genes (Figure 6). Altogether, our data implicate expression of glial-specific genes, rather than neuronal-specific genes, as the most reliable predictor of biological age in the human brain.

### Functional Enrichment Analysis of Multi-regional and Region-Specific Aging-Altered Genes

To gain further insight into the functional nature of aging-altered genes, we performed enrichment analysis of Gene Ontology (GO) terms of these genes (Edgar et al., 2013). The upregulated genes were enriched in the following functional terms: “MG cell development”, “interleukin-1 (IL-1) receptor activity”, and “immune response”. Supporting these observations, neuroinflammation is known to be involved in aging, with evidence implicating the interferon type I response in aging-associated cognitive decline (Baruch et al., 2014). Conversely, downregulated multi-regional genes were enriched in the processes of “protein transport and localization”, and aging-altered expression of these genes led to shifts in regional identity (Figure S2B). Moreover, 244 of a total of 253 genes annotated to the “protein transport” category were detected as altered upon aging in at least one brain region. These genes separated the CRBL from the other regions and maintained their regional sub-classification (Figure S5Bi). In addition, the CRBL samples of the old group remained clearly separated from the rest

ure S4C for age association plots of endothelial, OLG precursor, and newly formed OLG cell-specific genes.

(Figure S5Bii). Most aging-altered genes annotated to this category were downregulated in eight regions, apart from the CRBL and WHMT, which showed greater expression variability (Figure S5Biii). The WHMT aging-altered genes were functionally enriched in “regulation of cell adhesion”, “regulation of cell development”, “metabolic processes”, and “cognition” (Figure S6A). Conversely, among the top functional terms that were enriched in aging-altered genes in the FCTX were immune functions including “T cell differentiation”, “T cells”, and “leukocyte and lymphocyte activation” (Table S2). Among the CRBL-enriched functions were “cell adhesion”, “regulation of cell motion and migration”, and “neuron projection morphogenesis”. These results imply a region-specific functional heterogeneity of the brain aging process.

Analysis of the cell-type-specific aging-altered genes revealed enrichment of further functional pathways. “Synaptosome”, “regulation of programmed cell death”, and “metal ion transporter” were enriched in downregulated neuron-specific genes. “Regulation of adaptive immune response”, “natural killer cell-mediated cytotoxicity”, and “cell adhesion and motion” were enriched in MG-specific upregulated genes (Table S2). “Myelination”, “oxidoreductase”, and “RAS protein signal transduction” were enriched in upregulated OLG-specific genes. “Mitochondrial matrix”, “phosphate metabolic process”, and “Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway AD” were enriched in downregulated myelinating OLG-specific genes. Finally, “cell morphogenesis” and “cell-cell adhesion” were enriched in upregulated AC-specific genes. Some of these functions have also been identified in a study that examined the initial cell-type-specific transcriptional changes in a mouse model of amyotrophic lateral sclerosis (ALS), including synaptic functions in neurons and membrane signaling defects in OLGs (Sun et al., 2015).

## DISCUSSION

This study presents a comprehensive analysis of RNA expression in ten regions of the human brain and large-scale cell quantification in FCTX upon aging. Our findings show that cell-type-specific genes delineate samples based on both age group and brain region. Aging was the major determinant of glia-specific gene expression shifts in regional identity, while such changes were not evident in neuron-specific genes. Genes specific for neurons and OLGs generally decreased their expression upon aging, while MG-specific genes increased their expression profiles, consistent with the known MG activation in aging (Norden and Godbout, 2013). ACs showed a more complex pattern of reciprocal regional changes upon aging, with upregulation in the cortical regions and downregulation in the deeper brain structures. Among the genes specific for the non-neuronal cell types, those with the highest absolute expression in the young group decrease their expression upon aging in most brain regions, while those with the lowest expression in the young group increase their expression in a subset of regions. This leads to major shifts in region-specific gene expression, particularly of AC- and OLG-specific genes, which are most pronounced in the HIPP and SNIG, the regions that are archetypally affected in the most common age-related neurodegenerative diseases

(AD and PD, respectively). These findings reinforce a growing body of evidence implicating glia in aging (Norden and Godbout, 2013).

Age-related degeneration of OLG has been previously observed in the HIPP of the senescence-accelerated mice, as well as other animal models (Hayakawa et al., 2007; Hwang et al., 2006; Shimeda et al., 2005). OLG-specific genes were also found to have the strongest enrichment among genes with decreased age-related expression in human TCTX (Tollervey et al., 2011). We demonstrate that the age-related downregulation of OLG-specific gene expression is accompanied by a decrease in OLG cell numbers in the FCTX, consistent with previous observations of decreased OLGs in neocortical regions of old human brains (Fabricius et al., 2013; Pelvig et al., 2008). The OLG-specific aging-altered genes include *MBP*, a major constituent of the myelin sheath, and *LINGO1*, a regulator of myelination (Mi et al., 2005). It is of interest that the low-density tiles (corresponding to likely white matter) show the largest decrease in oligodendrocytes and a corresponding increase in other cells (Figure 5). Given the strongest upregulation of MG-specific genes in the brain, it appears possible that the increase of other cells is driven by the increase in MG, but this remains to be directly examined.

We found increased AC-specific gene expression in human aging HIPP, which agrees with data from aging mouse models, in which increased proliferation and activation of ACs are reported (Hayakawa et al., 2007) (Figure 4). An examination of three sub-regions of mouse HIPP using three AC-specific protein markers revealed complex, region-specific, and marker-dependent changes (Rodríguez et al., 2014). Regionally encoded AC expression is important for neuronal functions, as was demonstrated by the loss of ventral spinal cord AC-encoded *SEMA3A* gene expression, which leads to selective death of  $\alpha$ -motor neurons in mice (Molofsky et al., 2014). We find many regional differences in expression of AC-specific genes are largely erased in samples from the old group; for example, these genes cluster the HIPP and PUTM separately in the young group, but not in the old group (Figure 4B). This suggests that major changes in functional heterogeneity of AC take place in the aging brain, which might have deleterious consequences on the integrity of neuronal circuits.

A trend toward increased expression of MG-specific genes was observed in all regions upon aging, with corresponding upregulation of genes with immune or inflammatory functions. The upregulated genes include *C1Q*, which agrees with the increased C1Q protein levels that were observed in both mouse and human brains upon aging (Stephan et al., 2013). Another upregulated gene is *TREM2*, which is also upregulated in amyloid-plaque-associated MG (Frank et al., 2008) and contains variant alleles that increase AD risk (Guerreiro et al., 2013). Upregulation of inflammatory functions is in line with evidence implicating the interferon type I response in age-associated cognitive decline (Baruch et al., 2014).

In addition to glial changes, we also observed a decreased number of neurons with large cell bodies, which represent approximately 20% of neurons in the cortex. Although we did not attempt to directly identify the neuronal subtypes in the present study, neurons with the largest cell bodies are likely to be

associative pyramidal neurons (Zeba et al., 2008). Furthermore, these neurons were previously indicated to be most vulnerable to aging in a study of Rhesus monkeys (Gilman et al., 2016). While our analysis indicates that the decrease in these pyramidal neurons may be the primary source of the downregulation of neuron-specific genes, our findings regarding the cortical neuronal cells remain speculative due to the limited number of individuals used for the imaging analyses. Moreover, it remains possible that the change does not result from loss of these neurons, but rather from downregulation of Rbfox3 protein, or its loss from the cytoplasm of large neurons. Thus, our current analysis will need to be verified with the use of additional markers of specific neuronal cell types and increased sample size, which will potentially include additional brain regions; ideally, it will also be compared to the outcomes of cell-type-specific analyses of RNA sequencing datasets (Lake et al., 2016).

Age is the major risk factor for both AD and PD, the two most prevalent neurodegenerative diseases. It is becoming clear that the pre-clinical stage of AD begins decades before clinical manifestation (Dubois et al., 2014). This pre-clinical stage has been termed “the cellular phase,” because it involves changes in interactions among all cell types in the brain, with the most dramatic changes taking place in AC, MG, and vasculature (De Strooper and Karran, 2016). We find a corrosion of glial region-specific gene expression in aging, with the genes specific for AC, MG, and endothelial cells being the best predictors of age. HIPP and SNIG are affected in the early stages of AD and PD, respectively, and these are the two regions with major shifts in their regional expression profiles of AC- and OLG-specific genes upon aging. Thus, our data may provide insights into the role of glia in the region-specific vulnerability in these age-related neurodegenerative diseases.

By simultaneously assessing changes in cell-type-specific genes across multiple brain areas, our study takes a step toward providing a comprehensive framework of the molecular and cellular changes in human aging. While our primary aim was to deconvolute the cell-type-specific signatures present within large databases of age-related transcriptional changes, we also made a step toward interpreting these in light of changes in counts of OLG and neuronal cells. Integration of further genome-wide and single-cell data from human tissues samples and cell and animal models will be required to fully understand the cellular and molecular mechanisms underlying the observations in our study. Altogether, our study indicates that the cellular changes during aging involve a dramatic shift in the regional identity of glia, and it provides a resource for further studies of the relationship between aging and the cellular phase of dementia.

## Conclusions

Our study examines brain-wide gene expression patterns in the aging human brain across a wide physiological age range, coupled with complementary analysis of cell-type-specific marker genes and validation by direct cell quantification using immunohistochemical imaging followed by targeted computational analysis. In addition to the expected increase in expression of MG-specific genes and decrease in expression of neuron-specific genes, our analyses uncovered major changes in the re-

gion-specific expression of AC- and OLG-specific genes. The age-associated changes in the regional expression of glial-specific genes are most dramatic in HIPP and SNIG, the brain regions affected in AD and PD. The age-dependent decrease in expression of OLG- and neuron-specific genes aligns with the results of direct cortical cell counting, in which decreased numbers of OLGs and of neurons with large cell bodies are demonstrated. We believe that our data and computational approaches provide a powerful resource for further study of the cellular and molecular changes taking place during human brain aging and provide insights into the pre-clinical cellular phase of dementia.

## EXPERIMENTAL PROCEDURES

### Ethical Statement

All samples used for this study had fully informed consent for retrieval and were authorized for ethically approved scientific investigation (Research Ethics Committee number 10/H0716/3).

### Brain Samples

Post-mortem human brain material was produced under institutional guidelines governed by approved protocols. Tissue samples were produced from 99 individuals by the Sudden Death MRC brain bank, 35 individuals by the Sun Head Institute for the UKBEC, and 305 individuals by the American Brain Bank (NABEC).

### Quality Assessment and Array Pre-processing

For the UKBEC dataset, all quality measurements were extensively described in a previous publication (Trabzuni et al., 2011). The initial pre-processing of the microarray data, including application of RMA (robust multi-assay) average quantile normalization with guanine cytosine (GC) background corrections (GC-RMA) and expression data were log<sub>2</sub> transformed. The gene level signal estimates were calculated for a total of 26,493 transcripts using the median signal of each group of probe sets interrogating a transcript.

### Expression Data Analysis

A tailored analysis pipeline was developed for all computational analyses and data visualization of microarray and RNA-seq datasets that were analyzed in this study (in MATLAB, R2014-2016a). Those include construction of data structures and statistical significance inference using ANOVA, with false discovery rate (FDR) thresholding (of corrected  $p < 1e-3$ ), classification and clustering (e.g., using t-distributed stochastic neighbor embedding [t-SNE] and hierarchical clustering), and data visualization.

Cell-type-specific genes were defined by analysis of RNA-seq data from mouse brain ([http://web.stanford.edu/group/barres\\_lab/brain\\_rnaseq.html](http://web.stanford.edu/group/barres_lab/brain_rnaseq.html)) and were further used to find age-predictive cell-specific genes. The lists are under Tables S5 and S6, accordingly. Additional cell-specific lists were based on a previous microarray data on three of these cell types (Cahoy et al., 2008). Further details are in the Supplemental Information.

### High-Resolution Imaging and Analysis of Immunolabeled Brain Samples

Post-mortem human brain sections were placed into xylene and rehydrated. Antigen retrieval was performed with citric acid. For OLG staining, the samples were immunolabeled with OLIG2 antibody using the Leica Novolink Polymer detection kit. We used the Olig2 antibody from Millipore (catalog #AB9610) at 1/200 dilution. For staining of neurons, the samples were immunolabeled with NeuN antibody (Acris) and the Leica Bond Epitope Retrieval Solution#1 was used (AR9961 from Leica Biosystems) (AR9961 from Leica Biosystems). The images for both types of stains were acquired on the Zeiss AxioScan slide scanner. Details of the cell detection and quantification computational methods for neurons are given under the Supplemental Experimental Procedures. In addition, all raw jpeg images of the slides can be seen at <https://figshare.com/s/f2675361af1242f3565f>. For image analysis, we employed

some of the computational methods mentioned by Bjornsson et al. (2008), in addition to a targeted computational pipeline developed in-house in MATLAB (see details under Supplemental Experimental Procedures).

### ACCESSION NUMBERS

The accession numbers for the UKBEC exon and NABEC expression datasets reported in this paper are GEO: GSE46706 and GSE36192, respectively.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2016.12.011>.

### CONSORTIA

The members of the UK Brain Expression Consortium are John Hardy, Mina Ryten, Daniah Trabzuni, Sebastian Guelfi, Michael E. Weale, Adaikalavan Ramasamy, Paola Forabosco, Colin Smith, and Robert Walker. The members of the North American Brain Expression Consortium are Sampath Arepalli, Mark R. Cookson, Allissa Dillman, J. Raphael Gibbs, Dena G. Hernandez, Michael A. Nalls, Andrew B. Singleton, Bryan Traynor, Marcel van der Brug, Luigi Ferrucci, Robert Johnson, Ronal Zielke, Dan L. Longo, Juan Toncoso, and Alan Zonderman.

### AUTHOR CONTRIBUTIONS

L.S., R.P., and J.U. conceived and designed the project and wrote the manuscript with contributions from all co-authors; J.U. and R.P. contributed equally to the work. L.S. developed the computational pipelines, analyzed the data, and produced the figures. E.S. developed targeted imaging computational analyses. M.R. and D.T. provided RNA extraction and microarray sample preparations. J.H. and M.R. provided access to the UKBEC dataset, and M.R.C. provided access to the NABEC datasets. J.R. and C.S. provided immunohistochemistry preparation, and R.P. supervised the immunohistochemical slide analysis and interpretation.

### ACKNOWLEDGMENTS

This work was supported by the European Research Council (617837-Translate) to J.U.; the Marie Curie Intra European Fellowship (330430-PRANA), and the Alzheimer's Society for Junior Investigator award (award 172065, project 534121) to L.S.; the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001002), the UK Medical Research Council (FC001002); the Wellcome Trust (FC001002); the UK Medical Research Council (MRC) through the MRC Sudden Death Brain Bank to C.S.; a project grant (G0901254) to J.H. and a training fellowship (G0802462) to M.R.; and in part by the Intramural Research Program of the US National Institute on Aging, NIH, Department of Health and Human Services (project ZO1 AG000947). R.P. is a Wellcome Trust Intermediate Clinical Fellow (101149/Z/13/A) and an Anne Rowling Fellow in Regenerative Neurology. D.T. was supported by the King Faisal Specialist Hospital and Research Centre, Saudi Arabia. We are grateful to the Banner Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona, for the provision of human biospecimens. The Brain and Body Donation Program is supported by the US National Institute of Neurological Disorders and Stroke (U24 NS072026 to the National Brain and Tissue Resource for Parkinson's Disease and Related Disorders), the National Institute on Aging (P30 AG19610 to the Arizona Alzheimer's Disease Core Center), the Arizona Department of Health Services (contract 211002 to the Arizona Alzheimer's Research Center), the Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901, and 1001 to the Arizona Parkinson's Disease Consortium), and the Michael J. Fox Foundation for Parkinson's Research. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging.

Received: May 2, 2016

Revised: October 4, 2016

Accepted: December 2, 2016

Published: January 10, 2017

### REFERENCES

- Baruch, K., Deczkowska, A., David, E., Castellano, J.M., Miller, O., Kertser, A., Berkutzki, T., Barnett-Iltzhaki, Z., Bezalet, D., Wyss-Coray, T., et al. (2014). Aging. Aging-induced type I interferon response at the choroid plexus negatively affects brain function. *Science* 346, 89–93.
- Bjornsson, C.S., Lin, G., Al-Kofahi, Y., Narayanaswamy, A., Smith, K.L., Shain, W., and Roysam, B. (2008). Associative image analysis: a method for automated quantification of 3D multi-parameter images of brain tissue. *J. Neurosci. Methods* 170, 165–178.
- Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., et al. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* 28, 264–278.
- Cerbai, F., Lana, D., Nosi, D., Petkova-Kirova, P., Zecchi, S., Brothers, H.M., Wenk, G.L., and Giovannini, M.G. (2012). The neuron-astrocyte-microglia triad in normal brain ageing and in a model of neuroinflammation in the rat hippocampus. *PLoS ONE* 7, e45250.
- 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.
- De Strooper, B., and Karran, E. (2016). The cellular phase of Alzheimer's disease. *Cell* 164, 603–615.
- 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.
- Edgar, R., Mazor, Y., Rinon, A., Blumenthal, J., Golan, Y., Buzhor, E., Livnat, I., Ben-Ari, S., Lieder, I., Shitrit, A., et al. (2013). LifeMap Discovery™: the embryonic development, stem cells, and regenerative medicine research portal. *PLoS ONE* 8, e66629.
- Erraji-Benchekroun, L., Underwood, M.D., Arango, V., Galfalvy, H., Pavlidis, P., Smythiopoulos, P., Mann, J.J., and Sibille, E. (2005). Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. *Biol. Psychiatry* 57, 549–558.
- Fabricius, K., Jacobsen, J.S., and Pakkenberg, B. (2013). Effect of age on neocortical brain cells in 90+ year old human females—a cell counting study. *Neurobiol. Aging* 34, 91–99.
- Frank, S., Burbach, G.J., Bonin, M., Walter, M., Streit, W., Bechmann, I., and Deller, T. (2008). TREM2 is upregulated in amyloid plaque-associated microglia in aged APP23 transgenic mice. *Glia* 56, 1438–1447.
- Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., et al. (2010). Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* 6, e1000952.
- Gilman, J.P., Medalla, M., and Luebke, J.I. (2016). Area-specific features of pyramidal neurons—a comparative study in mouse and rhesus monkey. *Cereb. Cortex*, bhw062.
- Grabert, K., Michoel, T., Karavolos, M.H., Clohisey, S., Baillie, J.K., Stevens, M.P., Freeman, T.C., Summers, K.M., and McColl, B.W. (2016). Microglial brain region-dependent diversity and selective regional sensitivities to aging. *Nat. Neurosci.* 19, 504–516.
- Guerreiro, R., Wojtas, A., Bras, J., Carrasquillo, M., Rogaeve, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J.S., Younkin, S., et al.; Alzheimer Genetic Analysis Group (2013). TREM2 variants in Alzheimer's disease. *N. Engl. J. Med.* 368, 117–127.

- Harel, A., Inger, A., Stelzer, G., Strichman-Almashanu, L., Dalah, I., Safran, M., and Lancet, D. (2009). GIFTS: annotation landscape analysis with GeneCards. *BMC Bioinformatics* *10*, 348.
- Hayakawa, N., Kato, H., and Araki, T. (2007). Age-related changes of astrocytes, oligodendrocytes and microglia in the mouse hippocampal CA1 sector. *Mech. Ageing Dev.* *128*, 311–316.
- Head, D., Buckner, R.L., Shimony, J.S., Williams, L.E., Akbudak, E., Conturo, T.E., McAvoy, M., Morris, J.C., and Snyder, A.Z. (2004). Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: evidence from diffusion tensor imaging. *Cereb. Cortex* *14*, 410–423.
- Hwang, I.K., Yoo, K.Y., Kim, D.S., Kang, T.C., Lee, B.H., Kim, Y.S., and Won, M.H. (2006). Chronological distribution of Rip immunoreactivity in the gerbil hippocampus during normal aging. *Neurochem. Res.* *31*, 1119–1125.
- Kim, K.K., Adelstein, R.S., and Kawamoto, S. (2009). Identification of neuronal nuclei (NeuN) as Fox-3, a new member of the Fox-1 gene family of splicing factors. *J Biol Chem.* *284*, 31052–31061.
- Kumar, A., Gibbs, J.R., Beilina, A., Dillman, A., Kumaran, R., Trabzuni, D., Ryten, M., Walker, R., Smith, C., Traynor, B.J., et al. (2013). Age-associated changes in gene expression in human brain and isolated neurons. *Neurobiol. Aging* *34*, 1199–1209.
- Lake, B.B., Ai, R., Kaeser, G.E., Salathia, N.S., Yung, Y.C., Liu, R., Wildberg, A., Gao, D., Fung, H.L., Chen, S., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* *352*, 1586–1590.
- 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., Salihi, 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.
- Mi, S., Miller, R.H., Lee, X., Scott, M.L., Shulag-Morskaya, S., Shao, Z., Chang, J., Thill, G., Levesque, M., Zhang, M., et al. (2005). LINGO-1 negatively regulates myelination by oligodendrocytes. *Nat. Neurosci.* *8*, 745–751.
- Molofsky, A.V., Kelley, K.W., Tsai, H.H., Redmond, S.A., Chang, S.M., Madiredy, L., Chan, J.R., Baranzini, S.E., Ullian, E.M., and Rowitch, D.H. (2014). Astrocyte-encoded positional cues maintain sensorimotor circuit integrity. *Nature* *509*, 189–194.
- Mosher, K.I., and Wyss-Coray, T. (2014). Microglial dysfunction in brain aging and Alzheimer's disease. *Biochem. Pharmacol.* *88*, 594–604.
- Norden, D.M., and Godbout, J.P. (2013). Review: microglia of the aged brain: primed to be activated and resistant to regulation. *Neuropathol. Appl. Neurobiol.* *39*, 19–34.
- Orre, M., Kamphuis, W., Dooves, S., Kooijman, L., Chan, E.T., Kirk, C.J., DiMayuga Smith, V., Koot, S., Mamber, C., Jansen, A.H., et al. (2013). Reactive glia show increased immunoproteasome activity in Alzheimer's disease. *Brain* *136*, 1415–1431.
- Pelvig, D.P., Pakkenberg, H., Stark, A.K., and Pakkenberg, B. (2008). Neocortical glial cell numbers in human brains. *Neurobiol. Aging* *29*, 1754–1762.
- Perry, V.H., and Teeling, J. (2013). Microglia and macrophages of the central nervous system: the contribution of microglia priming and systemic inflammation to chronic neurodegeneration. *Semin. Immunopathol.* *35*, 601–612.
- Peters, A., and Sethares, C. (2004). Oligodendrocytes, their progenitors and other neuroglial cells in the aging primate cerebral cortex. *Cereb. Cortex* *14*, 995–1007.
- Rodríguez, J.J., Yeh, C.Y., Terzieva, S., Olabarria, M., Kulijewicz-Nawrot, M., and Verkhatsky, A. (2014). Complex and region-specific changes in astroglial markers in the aging brain. *Neurobiol. Aging* *35*, 15–23.
- Shimeda, Y., Hirotsani, Y., Akimoto, Y., Shindou, K., Ijiri, Y., Nishihori, T., and Tanaka, K. (2005). Protective effects of capsaicin against cisplatin-induced nephrotoxicity in rats. *Biol. Pharm. Bull.* *28*, 1635–1638.
- Sibille, E. (2013). Molecular aging of the brain, neuroplasticity, and vulnerability to depression and other brain-related disorders. *Dialogues Clin. Neurosci.* *15*, 53–65.
- Stephan, A.H., Madison, D.V., Mateos, J.M., Fraser, D.A., Lovelett, E.A., Cou-tellier, L., Kim, L., Tsai, H.H., Huang, E.J., Rowitch, D.H., et al. (2013). A dramatic increase of C1q protein in the CNS during normal aging. *J. Neurosci.* *33*, 13460–13474.
- Streit, W.J., and Xue, Q.S. (2010). The brain's aging immune system. *Aging Dis.* *1*, 254–261.
- Sun, S., Sun, Y., Ling, S.C., Ferraiuolo, L., Mcalonis-Downes, M., Zou, Y., Drenner, K., Want, Y., Ditsworth, D., Tokunaga, S., et al. (2015). Translational profiling identifies a cascade of damage initiated in motor neurons and spreading to glia in mutant SOD1-mediated ALS. *Proc. Natl. Acad. Sci. USA* *112*, E6993–E7002.
- Tollervey, J.R., Wang, Z., Hortobágyi, T., Witten, J.T., Zarnack, K., Kayikci, M., Clark, T.A., Schweitzer, A.C., Rot, G., Curk, T., et al. (2011). Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. *Genome Res.* *21*, 1572–1582.
- Trabzuni, D., Ryten, M., Walker, R., Smith, C., Imran, S., Ramasamy, A., Weale, M.E., and Hardy, J. (2011). Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. *J. Neurochem.* *119*, 275–282.
- Tremblay, M.E., Zettel, M.L., Ison, J.R., Allen, P.D., and Majewska, A.K. (2012). Effects of aging and sensory loss on glial cells in mouse visual and auditory cortices. *Glia* *60*, 541–558.
- Vernooij, M.W., de Groot, M., van der Lugt, A., Ikram, M.A., Krestin, G.P., Hofman, A., Niessen, W.J., and Breteler, M.M. (2008). White matter atrophy and lesion formation explain the loss of structural integrity of white matter in aging. *Neuroimage* *43*, 470–477.
- Zeba, M., Jovanov-Milosević, N., and Petanjek, Z. (2008). Quantitative analysis of basal dendritic tree of layer III pyramidal neurons in different areas of adult human frontal cortex. *Coll. Antropol.* *32* (Suppl 1), 161–169.
- Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keefe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* *34*, 11929–11947.

# Genomic Analyses Reveal the Influence of Geographic Origin, Migration, and Hybridization on Modern Dog Breed Development

Heidi G. Parker,<sup>1</sup> Dayna L. Dreger,<sup>1</sup> Maud Rimbault,<sup>1</sup> Brian W. Davis,<sup>1</sup> Alexandra B. Mullen,<sup>1</sup> Gretchen Carpintero-Ramirez,<sup>1</sup> and Elaine A. Ostrander<sup>1,2,\*</sup>

<sup>1</sup>Cancer Genetics and Comparative Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA

<sup>2</sup>Lead Contact

\*Correspondence: [eostrand@mail.nih.gov](mailto:eostrand@mail.nih.gov)

<http://dx.doi.org/10.1016/j.celrep.2017.03.079>

## SUMMARY

There are nearly 400 modern domestic dog breeds with a unique histories and genetic profiles. To track the genetic signatures of breed development, we have assembled the most diverse dataset of dog breeds, reflecting their extensive phenotypic variation and heritage. Combining genetic distance, migration, and genome-wide haplotype sharing analyses, we uncover geographic patterns of development and independent origins of common traits. Our analyses reveal the hybrid history of breeds and elucidate the effects of immigration, revealing for the first time a suggestion of New World dog within some modern breeds. Finally, we used cladistics and haplotype sharing to show that some common traits have arisen more than once in the history of the dog. These analyses characterize the complexities of breed development, resolving longstanding questions regarding individual breed origination, the effect of migration on geographically distinct breeds, and, by inference, transfer of trait and disease alleles among dog breeds.

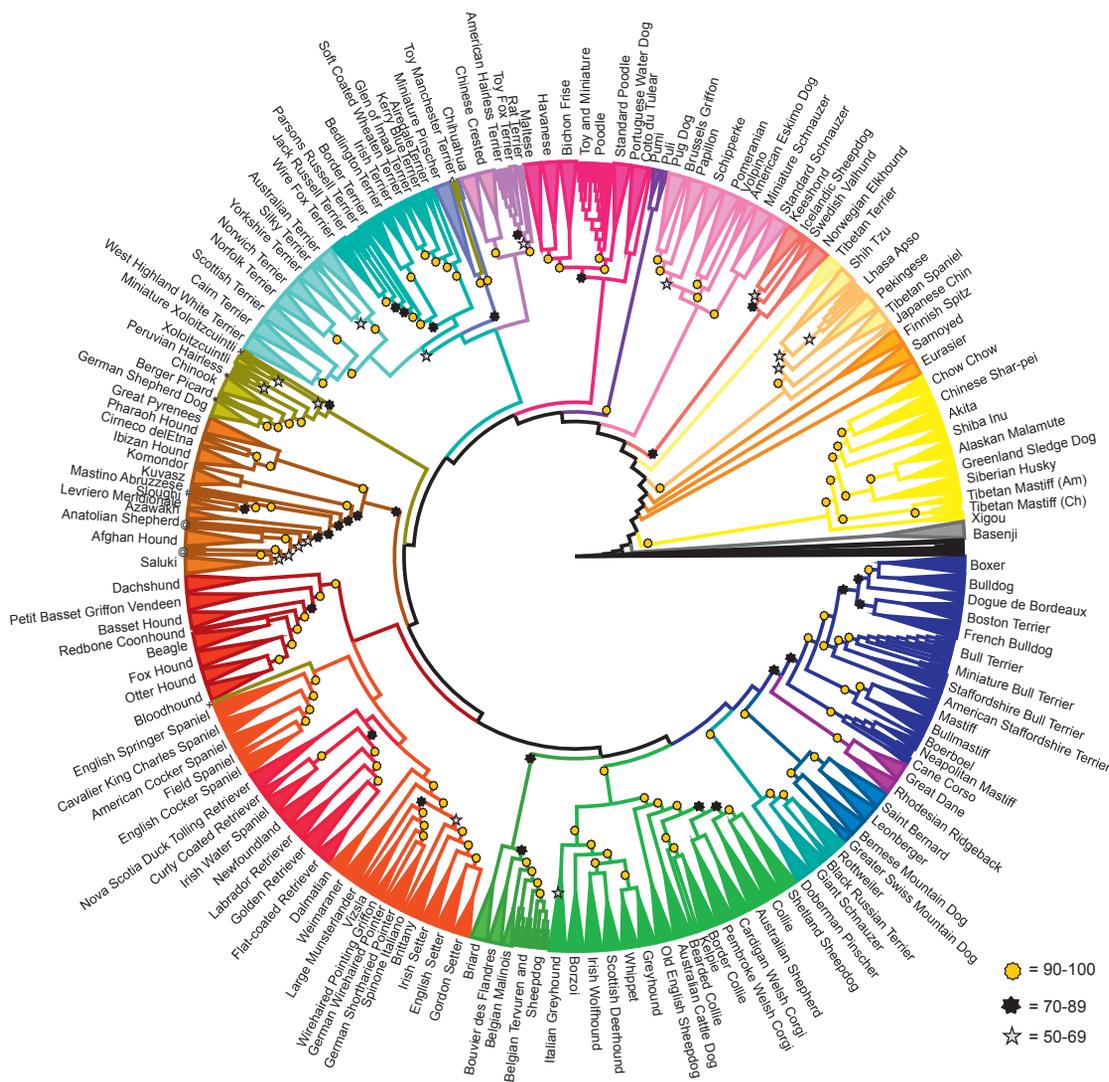
## INTRODUCTION

The dog, *Canis familiaris*, is the first domesticated animal earning a place within nearly every society across the globe for thousands of years (Druzhkova et al., 2013; Thalmann et al., 2013; Vilà et al., 1997, 1999). Over the millennia, dogs have assisted humans with hunting and livestock management, guarding house and field, and played crucial roles in major wars (Moody et al., 2006). Providing a range of services from companionship to production of fur and meat (Wilcox and Walkowicz, 1995), the diversity of talents and phenotypes combined with an unequalled emotional connection between dogs and humans has led to the creation of more than 350 distinct breeds, each of which is a closed breeding population that reflects a collage of defining traits (<http://www.akc.org>).

Previous studies have addressed the genomic makeup of a limited number of breeds, demonstrating that dogs from the same breed share common alleles and can be grouped using measures of population structure (Irion et al., 2003; Koskinen, 2003; Parker et al., 2004), and breeds that possess similar form and function often share similar allelic patterns (Parker et al., 2004, 2007; Vonholdt et al., 2010). However, none of these studies have effectively accounted for the variety of mechanisms through which modern breeds may have developed, such as geographic separation and immigration; the role of hybridization in the history of the breeds; and the timeline of the formation of breeds. In this study, we overcome these barriers by presenting an expansive dataset, including pure breeds sampled from multiple sections of the globe and genotyped on a dense scale. By applying both phylogenetic methods and a genome-wide analysis of recent haplotype sharing, we have unraveled common population confounders for many breeds, leading us to propose a two-step process of breed creation beginning with ancient separation by functional employment followed by recent selection for physical attributes. These data and analyses provide a basis for understanding which and why numerous, sometimes deleterious mutations are shared across seemingly unrelated breeds.

## RESULTS

We examined genomic data from the largest and most diverse group of breeds studied to date, amassing a dataset of 1,346 dogs representing 161 breeds. Included are populations with vastly different breed histories, originating from all continents except Antarctica, and sampled from North America, Europe, Africa, and Asia. We have specifically included breeds that represent the full range of phenotypic variation present among modern dogs, as well as three breeds sampled from both the United States and their country of origin. Samples from 938 dogs representing 127 breeds and nine wild canids were genotyped using the Illumina CanineHD bead array following standard protocols. Data were combined with publically available information from 405 dogs genotyped using the same chip (Hayward et al., 2016; Vaysse et al., 2011). For three dogs from one breed, genotypes were retrieved from publically available sequence



**Figure 1. Cladogram of 161 Domestic Dog Breeds**

Breeds that form unique clades supported by 100% of bootstraps are combined into triangles. For all other branches, a gold star indicates 90% or better, black star 70%–89%, and silver star 50%–69% bootstrap support. Breeds are listed on the perimeter of the circle. A small number of dogs do not cluster with the rest of their breed, indicated as follows: \*cane paratore, +Peruvian hairless dog, #sloughi, @country-of-origin salukis, and ^miniature xoloitzcuintle.

files, and all were merged into a single dataset (Table S1). After pruning for low quality or genotyping rate, 150,067 informative SNPs were retained.

Ascertainment bias has been shown to skew population genetic calculations that require estimation of allele frequencies and diversity measures (Lachance and Tishkoff, 2013). It has also been shown that ascertainment based on a single individual provides less bias than a mixed group (Patterson et al., 2012). The SNPs used in this study were identified primarily within the boxer or from boxer compared to another genome (Vaysse et al., 2011), which has exaggerated the boxer minor allele frequency (MAF; 0.351 in boxer compared to 0.260 overall) but has little effect the other breeds (MAF range, 0.247–0.284). To minimize the effect this might have, we have chosen to use distance measures based on allele sharing rather than frequency

and to enhance these analyses with unbiased haplotype sharing for a robust assessment of canine population structure.

A bootstrapped cladogram was obtained using an identity-by-state distance matrix and a neighbor-joining tree algorithm (Supplemental Experimental Procedures). After 100 bootstraps, 91% of breeds (146/161) formed single, breed-specific nodes with 100% bootstrap support (Figure 1). Of the 15 breeds that did not meet these criteria, seven (Belgian tervuren, Belgian sheepdog, cane corso, bull terrier, miniature bull terrier, rat terrier, and American hairless terrier) were part of two- or three-breed clades that were supported at 98% or greater, and two breeds (Lhasa apso and saluki) formed single-breed clades that were supported at 50% and 78%, respectively. Four breeds (redbone coonhound, sloughi, cane paratore, and Jack Russell terrier) were split within single multi-breed clades, and the last



**Figure 2. Representatives from Each of the 23 Clades of Breeds**

Breeds and clades are listed for each picture from left to right, top to bottom.

- (A) Akita/Asian spitz.
- (B) Shih tzu/Asian toy (by Mary Bloom).
- (C) Icelandic sheepdog/Nordic spitz (by Veronica Druk).
- (D) Miniature schnauzer/schnauzer.
- (E) Pomeranian/small spitz.
- (F) Brussels griffon/toy spitz (by Mary Bloom).
- (G) Puli/Hungarian.
- (H) Standard poodle/poodle.
- (I) Chihuahua/American toy.
- (J) Rat terrier/American terrier (by Stacy Zimmerman).
- (K) Miniature pinscher/pinscher.
- (L) Irish terrier/terrier.
- (M) German shepherd dog/New World (by Mary Bloom).
- (N) Saluki/Mediterranean (by Mary Bloom).
- (O) Basset hound/scent hound (by Mary Bloom).
- (P) American cocker spaniel/spaniel (by Mary Bloom).

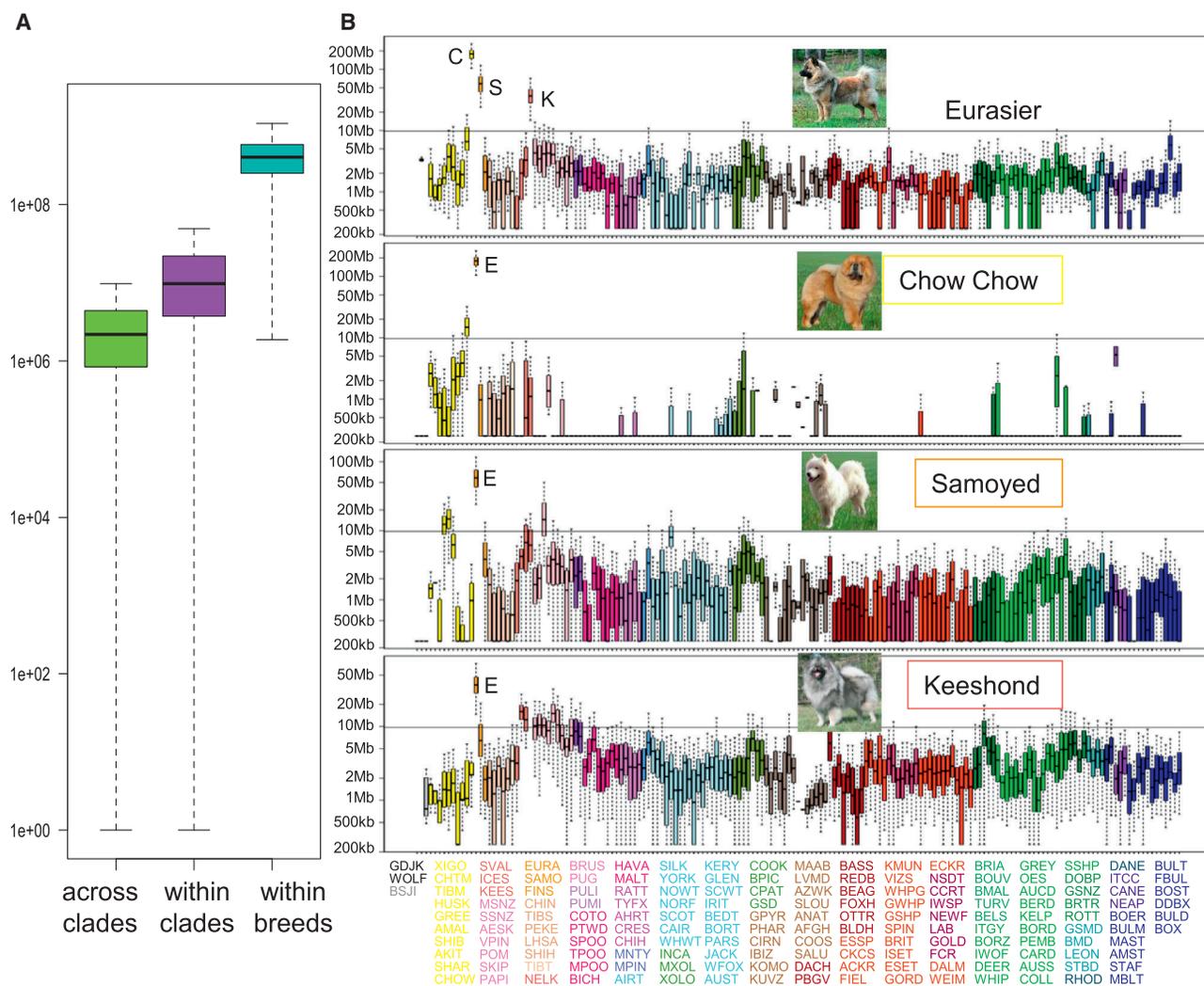
- (Q) Golden retriever/retriever (by Mary Bloom).
- (R) German shorthaired pointer/pointer setter (by Mary Bloom).
- (S) Briard/continental herder (by Mary Bloom).
- (T) Shetland sheepdog/UK rural.
- (U) Rottweiler/drover
- (V) Saint Bernard/alpine.
- (W) English mastiff/European mastiff (by Mary Bloom).

two breeds (xoloitzcuintli and Peruvian hairless dog) were split between divergent clades. Nine of the breeds that were not monophyletic were either newly recognized by the American Kennel Club (AKC) or not recognized at the time of sample collection and likely represent a breed under development. Two other non-monophyletic breeds are composed of dogs collected in two countries; the cane corsos collected in Italy form a fully supported, single clade, as do the salukis collected in the United States. However, the cane corsos collected in the United States form a paraphyletic clade near the Neapolitan mastiffs, and the salukis collected in the Middle East form multiple paraphyletic groups within a clade that includes the US salukis and Afghan hounds.

Not including those that are breed specific, this study defined 105 phylogenetic nodes supported by  $\geq 90\%$  of bootstrap replicates, 133 by  $\geq 70\%$ , and 150 supported by  $\geq 50\%$  of replicates. We identify 29 multi-breed clades that are supported at  $\geq 90\%$ . Each of these clades includes 2–16 breeds and together account for 78% of breeds in the dataset. 150 breeds, or 93% of the dataset, can be divided into 23 clades of 2–18 breeds each, supported at  $>50\%$ . These multi-breed clades reflect common behaviors, physical appearance, and/or related geographic origin (Figure 2).

Eleven breeds did not group with significance to any other breeds. Five breeds form independent clades and six others are paraphyletic to established clades with  $<50\%$  bootstrap support (Table S2). The lack of grouping may indicate that we have not sampled the closest relatives of these breeds or that these breeds comprise outcrossings that are not shared by similar breeds.

To assess hybridization across the clades, identical-by-descent (IBD) haplotype sharing was calculated between all pairs of dogs from the 161 breeds. Haplotypes were phased using the program Beagle (Browning and Browning, 2013) in 100-SNP windows, resulting in a minimum haplotype size of 232 kb, well above the shared background level established in previous studies (Lindblad-Toh et al., 2005; Sutter et al., 2004). The large haplotypes specifically target admixture resulting from breed formation rather than domestication, which previous studies have not addressed. The total length of the shared haplotypes was summed for each pair of dogs. Individuals from within the same breed clade share nearly four times more of their genome within large IBD haplotype blocks than dogs in different breed clades (median shared haplotype lengths of 9,742,000 bp and 2,184,000 bp, respectively;  $p$  [Kolmogorov-Smirnov (K-S) and Wilcox]  $< 2.2e^{-16}$ ; Figure 3A). Only 5% of the across-breed pairings have a median greater than 9,744,974 bp. These exceptions argue for recent admixture events between breeds, as evidenced by the example of the Eurasier breed, created in the 1970s by mixing chow chow with other spitz-type breeds (Fogle, 2000) (Figure 3B). The data reveal not only the components of the breed but also the explanation for its placement on the cladogram. The Eurasier (unclustered) shows significant haplotype sharing with the samoyed (unclustered), keeshond (Nordic spitz), and chow chow (Asian spitz) (Figure 3B). Because all three breeds are located in different clades, unrelated to each other, the Eurasier falls between the component breeds and forms its own single-breed clade. Haplotype-sharing bar graphs for each of 161 breeds, including 152 AKC breeds, are available in Data S1. This provides a long-term resource for identifying



**Figure 3. Gross Haplotype Sharing across Breeds**

(A) Boxplot of total haplotype sharing between all pairs of dogs from breeds within the same clade, across different clades, and within the same breed. The difference between the distributions is highly significant ( $p < 2e-16$ ).

(B) Example of haplotype sharing between three breeds (samoyed, chow chow, and keeshond) and a fourth (Eurasier) that was created as a composite of the other three. Combined haplotype length is displayed on the y axis, and 169 breeds and populations are listed on the x axis in the order they appear on the cladogram, starting with the jackal and continuing counterclockwise. Haplotype sharing of zero is set at 250,000 for graphing, a value just below what is detected in this analysis. Breeds are colored by clade. 95% significance level is indicated by the horizontal line. Breed abbreviations are listed under the graph in the order they appear and colored by clade. Definitions of the breed abbreviations can be found in Table S1.

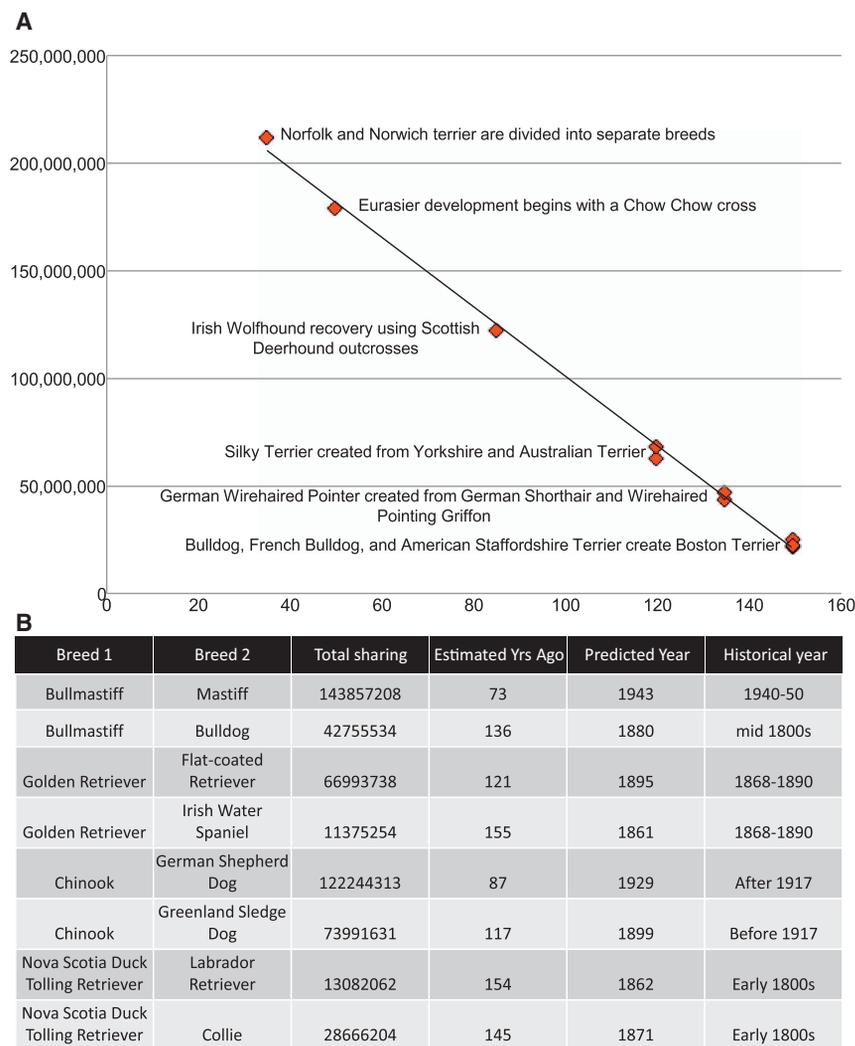
populations that likely share rare and common traits that will be invaluable for mapping the origins of deleterious and beneficial mutations.

Strong evidence of admixture across the clades was found in 117 breeds (Figure 4). A small number of these were identified in previous studies using migration analysis (Pickrell and Pritchard, 2012; Shannon et al., 2015) 30% of these breeds share with only one breed outside their clade. Therefore, more than half (54%) of the breeds that make up the 23 established clades share large haplotypes with one or zero breeds outside their clade, indicating breed creation by selection based on the initial founder population rather than recent admixture. Only 6 of the 161 breeds share extensive haplotypes with many (>8) different

groups, suggesting recent creation of these breeds from multiple others or that they provide a popular modern breed component. The overall low level of sharing across diverse breeds suggests that interclade crosses are done thoughtfully and for specific reasons, such as the introduction of a new trait or the immigration of a breed to a new geographic region.

As importation and establishment in a new country has been shown to have a measurable effect on breed structure (Quignon et al., 2007), we chose three breeds, the Tibetan mastiff, saluki, and cane corso, for inclusion in the study, with each collected in the country of origin as well as from established populations in the United States. In each case, there is division of the breed based on collection location. The split between the US and





**Figure 5. Total Haplotype Sharing Is Inversely Correlated with the Time of Hybridization between Breeds that Have Developed within the Last 200 Years**

(A) The time of hybridization in years before present is graphed on the x axis and the median total haplotype sharing on the y axis for six breeds of dog with reliable recent histories of admixture in breed formation or recovery. The trendline shows a linear correlation with  $r^2 = 1$ .

(B) The slope and intercept of the trendline from A was applied to the median haplotype sharing values from the data for four additional breeds with reliable breed creation dates to establish accuracy of estimated hybridization dates.

likelihood score over a zero migration tree (Figure 6A). Only 2 of the 18 clades, New World and Asian toy (Figures 6B and 6C), showed evidence of extensive hybridization between the breeds. Thus, the modern breeds were likely created through selection for unique traits within an ancient breed type with possible admixture from unrelated breeds to enhance the trait.

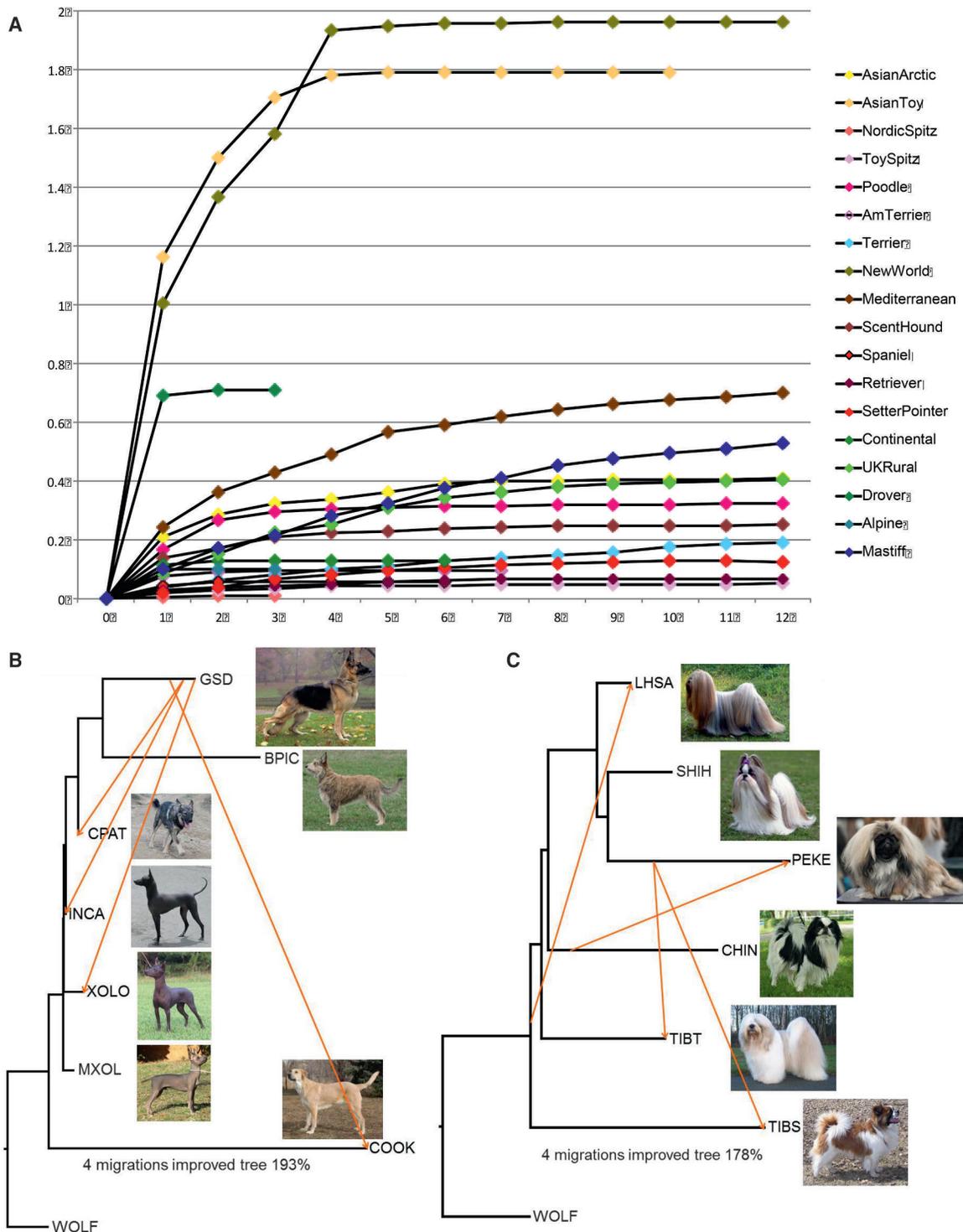
Our hybridization analysis reveals evidence for disease sharing across the clades. For instance, collie eye anomaly (CEA) is a disease that affects the development of the choroid in several herding breeds, including the collie, Border collie, Shetland sheepdog, and Australian shepherd, all members of the UK Rural clade (Lowe et al., 2003). The mutation and haplotype pattern displayed IBD across all affected breeds, and we speculated that all share a common obviously affected ancestor (Parker et al., 2007).

We were unable to explain, however, the presence of the disease in the Nova Scotia duck tolling retriever, a sporting dog developed in Canada from an unknown mixture of local breeds, which also shares the same haplotype. This perplexing observation can now be explained, as this analysis shows that collie and/or Shetland sheepdog were strong, undocumented, contributors to the formation of the Nova Scotia duck tolling retriever and, therefore, the likely source of the CEA mutation within that breed (Figure 7A).

Similarly, a mutation in the *MDR1* gene (multi-drug resistance 1), which causes life-threatening reactions to multiple drugs in many of the UK Rural breeds, has been reported in 10% of German shepherd dogs (Mealey and Meurs, 2008). These data display a link between the German shepherd dog and UK Rural breeds through the Australian shepherd, highlighting the unexpected role the Australian shepherd or its predecessor played in the development of the modern German shepherd dog (Figure 6B). Earlier this year, the *MDR1* mutation was identified in the chinook at a frequency of 15% (Donner et al., 2016). Our analysis reveals recent admixture between this breed and the German shepherd dog as

historically accurate time estimations (Figure 5B). Using the relationship equation  $y = -1,613,084.67x + 262,137,843.89$ , where  $y$  is the total shared haplotype length and  $x$  is the number of years, we can estimate the time at which undocumented crosses or divisions from older breeds took place. For example, based on a median haplotype sharing value of 66,993,738, the golden retriever was separated from the flat-coated retriever in 1895, and the written history of the golden retriever dates to crosses between multiple breeds taking place between 1868 and 1890 (Figure 5B), a near-perfect match.

To determine if the multi-breed clades are formed through recent admixture rather than through common ancestral sources, we examined migration in 18 groups of four or more breeds. These include 16 of the clades established on the tree, including nearby unclustered breeds, and two groups of small clades (American terrier/American toy and small spitz/toy spitz/schnauzer) that are monophyletic, but not well supported. Using the program Treemix (Pickrell and Pritchard, 2012), and allowing 0–12 predicted migration events, we determined the effect of admixture on clade formation by calculating the increase in maximum



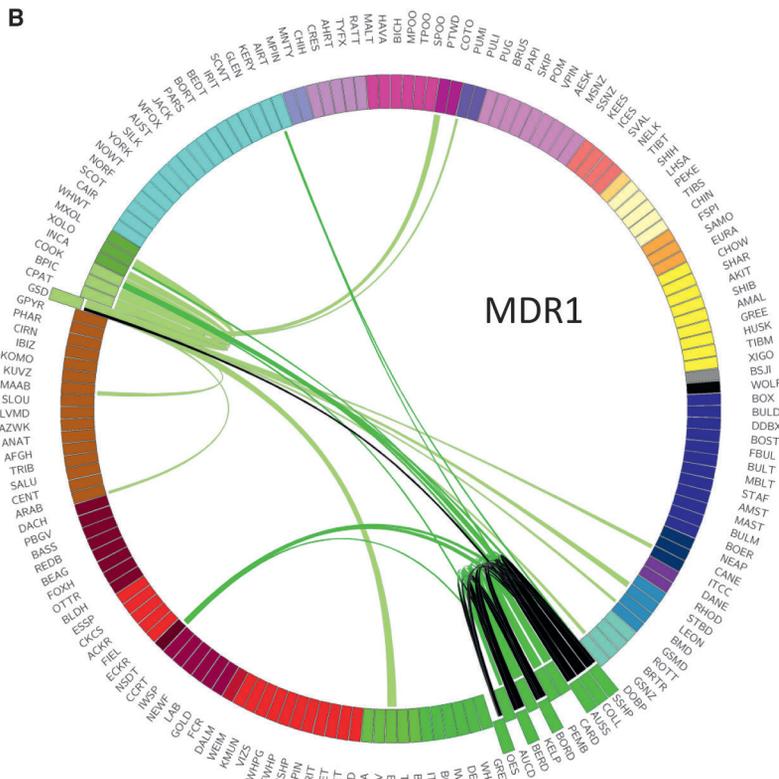
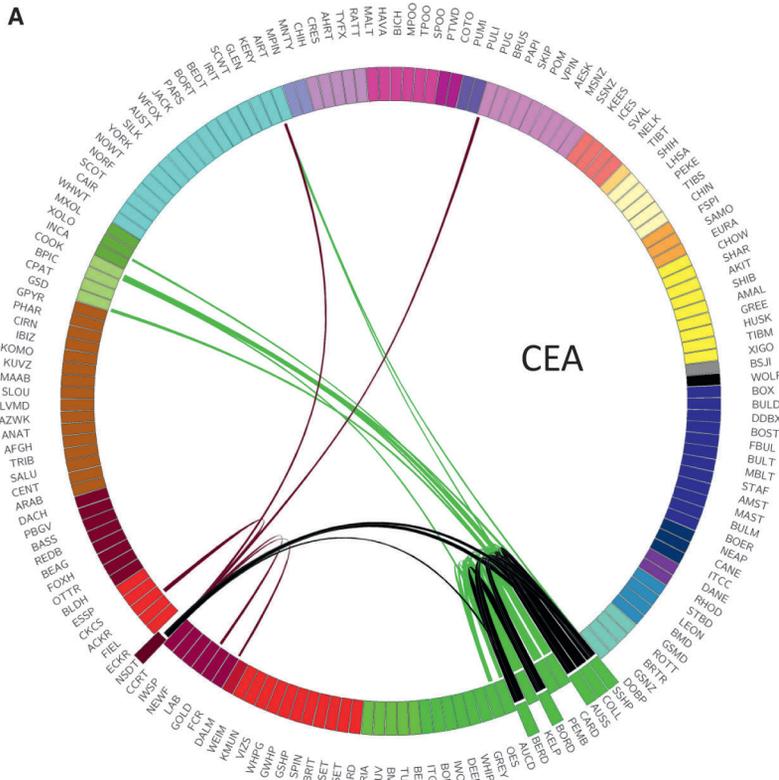
**Figure 6. Assessment of Migration between Breeds within Clades**

Admixture was measured in Treemix for 18 groups of breeds representing clades or combinations of small clades.

(A) Improvement to the maximum likelihood tree of each group as the result of admixture. The y axis shows fold improvement over the zero admixture tree.

(B) Cladogram of the New World breeds with European herders allowing four migration events. Arrows show estimated migration between breeds colored by weight (yellow to red = 0–0.5).

(C) Cladogram showing migration within the Asian toy clade, including a neighboring breed, the Tibetan terrier. Pictures by Terri Gueck (TIBT), Yuri Hooker (INCA), Mary Bloom (GSD and SHIH), Maurizio Marziali (CPAT), Mary Malkiel (COOK), and John and Debbie Caponetto (large and small XOLO/MXOL).



**Figure 7. Haplotypes Shared with Breeds that Carry Known Deleterious Mutations**

Breeds are connected if the median shared haplotype size exceeds the 95% threshold for interclade sharing. Sharing between breeds that are known to carry the mutation is colored black, and sharing with other breeds is colored according to the breed that carries the mutation.

(A) Collie eye anomaly is found in a number of herding breeds developed in the United Kingdom and some sporting breeds developed in the United States.

(B) Multi-drug resistance 1 mutation is carried by many UK herding breeds as well as the German shepherd.

well as previously unknown addition of Collie, both carriers of the *MDR1* mutation. Haplotype sharing with the same affected breeds is found in the xoloitzcuintli, which allows us to predict that this rare breed may also carry the deleterious allele but has yet to be tested.

## DISCUSSION

Phylogenetic analyses have often been applied to determine the relationships between dog breeds with the understanding that a tree structure cannot fully explain the development of breeds. Prior studies have shown that single mutations produce recognizable traits that are shared across breeds from diverse clades, suggesting that admixture across clades is a notable source of morphologic diversity (Cadieu et al., 2009; Parker et al., 2009; Sutter et al., 2007). Studies of linkage disequilibrium and haplotype sharing suggest further that within regions of ~10–15 kb, there exist a small number of haplotypes that are shared by the majority of breeds, while breed specificity is revealed only in large haplotypes (Lindblad-Toh et al., 2005; Sutter et al., 2004).

In this study, we observe that the majority of dog breeds either do not share large haplotypes outside their clade or share with only one remote breed. The small number of breeds that share excessively outside their assigned clade could be recently created from multiple diverse breeds or may have been popular contributors to other breeds. For example, the pug dog groups closely with the European toy breed, Brussels griffon (Figure 2F), in the toy spitz clade but also shares extensive haplotypes with the Asian toy breeds (Figure 2B) as well as many small dog breeds from multiple other clades. This likely indicates the pug's early exportation from Asia and subsequent contribution to many small breeds (Watson, 1906). Consider also the extensive cross-clade haplotype sharing in the chinook, a recently created breed with multiple ancestors from different breeds. Our data both recapitulates and enhances the written history of this breed (<http://www.chinook.org/history.html>) (Data S1). Extreme examples such as these underscore the complications implicit in relying on phylogeny alone to describe breed relationships. Overall, our data show that admixture has played an important role in the development of many breeds and, as new hybrids are added to phylogenetic analyses, the topology of the cladogram will likely rearrange to accommodate.

The ability to determine a time of hybridization for recent admixture events can refine sparse historical accounts of breed formation. For example, when dog fighting was a popular form of entertainment, many combinations of terriers and mastiff or bully-type breeds were crossed to create dogs that would excel in that sport. In this analysis, all of the bull and terrier crosses map to the terriers of Ireland and date to 1860–1870. This coincides perfectly with the historical descriptions that, though they do not clearly identify all breeds involved, report the popularity of dog contests in Ireland and the lack of stud book veracity, hence undocumented crosses, during this era of breed creation (Lee, 1894).

The dates estimated from these data are approximations, as selection for or against traits that accompanied each cross, as well as the size of the population at the time of the cross, would have affected retention of the haplotypes within the genome.

Based on these estimates, the excess haplotype sharing that we have identified represents the creation of breeds since the Victorian era breed explosion. Most breeds within each clade share haplotypes at this level (<200 ybp); however, the lack of sharing across the clades, outside of very specific crosses, suggests the clades were developed much earlier than the breed registries. Dividing the data by clade, the median haplotype sharing is lowest in the Asian spitz (median = 0) and the Mediterranean clades (median = 516,900) (median range across all clades = 0–3,459,000), indicating that these clades are most divergent and possibly older than the rest. This fits well with previous studies that suggest the earliest dogs came from Central and East Asia (Pang et al., 2009; Shannon et al., 2015). Interestingly, the mean haplotype sharing is slightly higher in the Asian spitz clade than it is in the Mediterranean clade (mean = 1,596,000 and 1,317,000, respectively) (Figure S1), implying that the Asian spitz breeds have been used in recent crosses while the Mediterranean breeds are currently more segregated. These data describe a staggered pattern of dog breed creation starting with separation by type based on required function and the form necessary to carry out that function. This would have taken place as the need arose during early human progression from hunter-gather to pastoral, agricultural, and finally urban lifestyles. During the last 200 years, these breed types were refined into very specific breeds by dividing the original functional dog into morphotypes based on small changes in appearance and with occasional outcrosses to enhance appearance or alter behavior (e.g., reduce aggression, increase biddability).

Though most breeds within a clade appear to be the result of descent from a common ancestor, the New World dogs and the Asian toys showed nearly 200% improvement in the maximum likelihood score by allowing for admixture between the breeds within the clade. Based on this analysis, the Asian toy dogs were likely not considered separate breeds when first exported from their country of origin resulting in multiple admixture events (Figure 6C). Unexpectedly, the New World clade admixture events center exclusively on the German shepherd dog, which informs both the development of this breed as well as immigration of dog breeds to the New World (Figure 6B). The inclusion of German shepherd dog with cane paratore, an Italian working farm dog, likely indicates a recent common ancestor among these breeds, as the German shepherd dog was derived from a herding dog of unknown ancestry in the late 1800s (<http://gsdca.org>). However, the hybridization of the German shepherd dog with the Peruvian hairless dog and the xoloitzcuintli, also a hairless breed, is unexpected and could be the result of recent admixture to enhance the larger varieties of these breeds or could indicate admixture of generic herding dogs from Southern Europe into South America during the Columbian Exchange.

Dogs have been in the Americas for more than 10,000 years, likely traveling from East Asia with the first humans (Wang et al., 2016). However, studies of mitochondrial DNA suggest that the original New World dogs were almost entirely replaced through European contact (Castroviejo-Fisher et al., 2011; Wayne and Ostrander, 1999; Witt et al., 2014) and additional Asian migrations (Brown et al., 2015). As colonists came to the Americas from the 16<sup>th</sup> to the 19<sup>th</sup> centuries, they brought Old

World livestock, and therefore the dogs required to manage and tend the livestock, to the New World (Crosby, 1972). Many of the newly introduced animals outcompeted the native animals, which may explain the surprising and very strong herding dog signature in the native hairless breeds of South and Central America that were not developed to herd. In this analysis, we observe that the ancient hairless breeds show extensive hybridization with herding dogs from Europe and, to a lesser extent, with each other. We also identify two additional clades of New World breeds, the American terriers and the American toys (Figures 2I and 2J), two monophyletic clades of small-sized breeds from North/Central America, which include a set of related terriers, and the Chihuahua and Chinese crested. Written records state that the terriers trace their ancestry to the feists, a North American landrace dog bred for hunting (<http://www.americantreeingfeist.com,akc.org>). The Chihuahua and Chinese crested are both believed to have originated in Central America (American Kennel Club, 1998; Parker et al., 2017), despite the nomenclature of the latter, which implies Asian ancestry. In contrast, most new breeds developed in the Americas were created from crosses of European breeds and cluster accordingly (i.e., Boston terrier [European mastiff], Nova Scotia duck tolling retriever [retriever], and Australian shepherd [UK Rural]). The separation of the older American breeds on the cladogram, despite recent European admixture, suggests that both clades may retain the aboriginal New World dog genomic signatures intermixed with the European breed haplotypes, similar to the admixture among European, African, and Native American genomes that can be found in modern South American human populations (Mathias et al., 2016; Ruiz-Linares et al., 2014). This is the first indication that the New World dog signature may not be entirely extinct in modern dog breeds, as has been previously suggested (Leonard et al., 2002).

In addition to the effects on the native population, our analysis of geographically distinct subsets of the same breeds shows that some degree of admixture also occurs in the imported breeds when first introduced into a new country. These data suggest two outcomes of breed immigration that mirror human immigration into a new region: the immigrant population is less diverse than the founding population, and there is often admixture with the native population in early generations (Baharian et al., 2016; Zhai et al., 2016).

We observe further evidence of the role geography plays in the distribution of breeds within the clades. For instance, both the UK Rural and the Mediterranean clades include both sighthound and working dog breeds, two highly divergent groups in terms of physical and behavioral phenotype. Sighthounds are lithe and leggy hunters, built to run fast, and have a strong prey drive. Working dogs include both the tall and heavy flock guards that are bred to live among herds without human interaction, preventing predator attacks, and mid-sized herders (Figure 2T), which are agile and bred to work closely with humans to control the movement of the flock without harming them. Despite the opposing phenotypes under selection, both breed types form single clades stemming from distinct geographical regions. Haplotype analysis shows no recent admixture between the geographically distinct clades, suggesting that these groups arose independently (Figures S2A and S2B). Archeological

depictions show sighthound-type hunting dogs that date back 4,000 ybp (Alderton, 2002; Fogle, 2000), and one of the earliest known writings regarding segregation of dogs based by type clearly delineates hunting dogs from working dogs (Columella, 1954). The new cladogram presented herein suggests that the switch from hunting to agricultural pursuits may have initiated early breed formation and that this occurred in multiple regions. These data show that geographical region can define a foundational canid population within which selection for universally relevant behaviors occurred independently, separating the regional groups also by function long ago.

The lack of admixture across clades that appear to share a common trait suggests that these traits may have arisen independently, multiple times. For example, these data show no recent haplotype sharing between the giant flock guards of the Mediterranean and the European mastiffs (Figure S2D). These breed types required large size for guarding; however, each used that size in a different way, a fact that was recognized at least 2,000 years ago (Columella, 1954). The flock guards use their size to defeat animal predators, while the mastiffs use their size to keep human predators at bay, often through fierce countenance rather than action. The phylogenetic placement of these breeds and lack of recent admixture suggests that giant size developed independently in the different clades and that it may have been one of the earliest traits by which breeds were segregated thousands of years ago.

The cladogram of 161 breeds presented here represents the most diverse dataset of domestic dog breeds analyzed to date, displaying 23 well-supported clades of breeds representing breed types that existed before the advent of breed clubs and registries. While the addition of more rare or niche breeds will produce a denser tree, the results here address many unanswered questions regarding the origins of breeds. We show that many traits such as herding, coursing, and intimidating size, which are associated with specific canine occupations, have likely been developed more than once in different geographical locales during the history of modern dog. These data also show that extensive haplotype sharing across clades is a likely indicator of recent admixture that took place in the time since the advent of breed registries, thus leading to the creation of most of the modern breeds. However, the primary breed types were developed well before this time, indicating selection and segregation of dog populations in the absence of formal breed recognition. Breed prototypes have been forming through selective pressures since ancient times depending on the job they were most required to perform. A second round of hybridization and selection has been applied within the last 200 years to create the many unique combinations of traits that modern breeds display. By combining genetic distance relationships with patterns of haplotype sharing, we can now elucidate the complex makeup of modern dogs breeds and guide the search for genetic variants important to canine breed development, morphology, behavior, and disease.

## EXPERIMENTAL PROCEDURES

Further details and an outline of resources used in this work can be found in the Supplemental Experimental Procedures.

## ACCESSION NUMBERS

The accession numbers for the raw data files for the SNP genotype arrays reported in this paper are GEO: GSE90441, GSE83160, GSE70454, and GSE96736.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, two figures, two tables, and two data files and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.03.079>.

## AUTHOR CONTRIBUTIONS

H.G.P. conceived of project, performed analyses, created figures, and prepared the manuscript. D.L.D. created figures and assisted in manuscript preparation. M.R. ran SNP chips and worked on early analysis. B.W.D. and A.B. performed experiments. G.C.-R. performed sample collection and DNA isolation. E.A.O. organized and directed the study and contributed to manuscript preparation

## ACKNOWLEDGMENTS

We gratefully acknowledge support from the Intramural Program of the National Human Genome Research Institute. We thank Sir Terence Clark for collecting DNA samples from multiple breeds of sighthounds from their countries of origin in Africa and Asia; Mauricio Lima, Flavio Bruno, and Robert Gennari for collecting samples from native Italian breeds; and Lei Song for collecting samples from native Tibetan mastiffs.

Received: January 3, 2017

Revised: February 10, 2017

Accepted: March 28, 2017

Published: April 25, 2017

## REFERENCES

- Alderton, D. (2002). *Dogs* (Dorling Kindersley, Ltd.).
- American Kennel Club (1998). *The Complete Dog Book*, 19th Edition Revised (Howell Book House).
- Baharian, S., Barakatt, M., Gignoux, C.R., Shringarpure, S., Errington, J., Blot, W.J., Bustamante, C.D., Kenny, E.E., Williams, S.M., Aldrich, M.C., and Gravel, S. (2016). The great migration and African-American genomic diversity. *PLoS Genet.* *12*, e1006059.
- Brown, S.K., Darwent, C.M., Wictum, E.J., and Sacks, B.N. (2015). Using multiple markers to elucidate the ancient, historical and modern relationships among North American Arctic dog breeds. *Heredity (Edinb.)* *115*, 488–495.
- Browning, B.L., and Browning, S.R. (2013). Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* *194*, 459–471.
- Cadiou, E., Neff, M.W., Quignon, P., Walsh, K., Chase, K., Parker, H.G., Vonholdt, B.M., Rhue, A., Boyko, A., Byers, A., et al. (2009). Coat variation in the domestic dog is governed by variants in three genes. *Science* *326*, 150–153.
- Castroviejo-Fisher, S., Skoglund, P., Valadez, R., Vilà, C., and Leonard, J.A. (2011). Vanishing native American dog lineages. *BMC Evol. Biol.* *11*, 73.
- Columella, L.J.M. (1954). *On Agriculture (De Re Rustica)*, Vol. Books 5–9 (E.S. Forster and E.H. Heffner, Trans.) (Harvard University Press).
- Crosby, A.W., Jr. (1972). *The Columbian Exchange* (Greenwood Publishing Company).
- Donner, J., Kaukonen, M., Anderson, H., Möller, F., Kyöstilä, K., Sankari, S., Hytönen, M., Giger, U., and Lohi, H. (2016). Genetic panel screening of nearly 100 mutations reveals new insights into the breed distribution of risk variants for canine hereditary disorders. *PLoS ONE* *11*, e0161005.
- Druzhkova, A.S., Thalmann, O., Trifonov, V.A., Leonard, J.A., Vorobieva, N.V., Ovodov, N.D., Graphodatsky, A.S., and Wayne, R.K. (2013). Ancient DNA analysis affirms the canid from Altai as a primitive dog. *PLoS ONE* *8*, e57754.
- Fogle, B. (2000). *The New Encyclopedia of the Dog*, Second Edition (Dorling Kindersley Publishing, Inc.).
- Hayward, J.J., Castelano, M.G., Oliveira, K.C., Corey, E., Balkman, C., Baxter, T.L., Casal, M.L., Center, S.A., Fang, M., Garrison, S.J., et al. (2016). Complex disease and phenotype mapping in the domestic dog. *Nat. Commun.* *7*, 10460.
- Irion, D.N., Schaffer, A.L., Famula, T.R., Eggleston, M.L., Hughes, S.S., and Pedersen, N.C. (2003). Analysis of genetic variation in 28 dog breed populations with 100 microsatellite markers. *J. Hered.* *94*, 81–87.
- Koskinen, M.T. (2003). Individual assignment using microsatellite DNA reveals unambiguous breed identification in the domestic dog. *Anim. Genet.* *34*, 297–301.
- Lachance, J., and Tishkoff, S.A. (2013). SNP ascertainment bias in population genetic analyses: why it is important, and how to correct it. *BioEssays* *35*, 780–786.
- Lee, R.B. (1894). *A History and Description of the Modern Dogs of Great Britain and Ireland* (Horace Cox).
- Leonard, J.A., Wayne, R.K., Wheeler, J., Valadez, R., Guillén, S., and Vilà, C. (2002). Ancient DNA evidence for Old World origin of New World dogs. *Science* *298*, 1613–1616.
- Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., 3rd, Zody, M.C., et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* *438*, 803–819.
- Lowe, J.K., Kukekova, A.V., Kirkness, E.F., Langlois, M.C., Aguirre, G.D., Acland, G.M., and Ostrander, E.A. (2003). Linkage mapping of the primary disease locus for collie eye anomaly. *Genomics* *82*, 86–95.
- Mathias, R.A., Taub, M.A., Gignoux, C.R., Fu, W., Musharoff, S., O'Connor, T.D., Vergara, C., Torgerson, D.G., Pino-Yanes, M., Shringarpure, S.S., et al.; CAAPA (2016). A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. *Nat. Commun.* *7*, 12522.
- Mealey, K.L., and Meurs, K.M. (2008). Breed distribution of the ABCB1-1Delta (multidrug sensitivity) polymorphism among dogs undergoing ABCB1 genotyping. *J. Am. Vet. Med. Assoc.* *233*, 921–924.
- Moody, J.A., Clark, L.A., and Murphy, K.E. (2006). Canine history and breed clubs. In *The Dog and Its Genome*, E.A. Ostrander, U. Giger, and K. Lindblad-Toh, eds. (Cold Spring Harbor Laboratory Press), pp. 1–18.
- Pang, J.F., Kluetsch, C., Zou, X.J., Zhang, A.B., Luo, L.Y., Angleby, H., Ardalan, A., Ekström, C., Skölleremo, A., Lundeberg, J., et al. (2009). mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol. Biol. Evol.* *26*, 2849–2864.
- Parker, H.G., Kim, L.V., Sutter, N.B., Carlson, S., Lorentzen, T.D., Malek, T.B., Johnson, G.S., DeFrance, H.B., Ostrander, E.A., and Kruglyak, L. (2004). Genetic structure of the purebred domestic dog. *Science* *304*, 1160–1164.
- Parker, H.G., Kukekova, A.V., Akey, D.T., Goldstein, O., Kirkness, E.F., Baysac, K.C., Mosher, D.S., Aguirre, G.D., Acland, G.M., and Ostrander, E.A. (2007). Breed relationships facilitate fine-mapping studies: a 7.8-kb deletion cosegregates with Collie eye anomaly across multiple dog breeds. *Genome Res.* *17*, 1562–1571.
- Parker, H.G., Vonholdt, B.M., Quignon, P., Margulies, E.H., Shao, S., Mosher, D.S., Spady, T.C., Elkahoulou, A., Cargill, M., Jones, P.G., et al. (2009). An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science* *325*, 995–998.
- Parker, H.G., Harris, A., Dreger, D.L., Davis, B.W., and Ostrander, E.A. (2017). The bald and the beautiful: hairlessness in domestic dog breeds. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *372*, 372.
- Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. *Genetics* *192*, 1065–1093.

- Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* 8, e1002967.
- Quignon, P., Herbin, L., Cadieu, E., Kirkness, E.F., Hédan, B., Mosher, D.S., Galibert, F., André, C., Ostrander, E.A., and Hitte, C. (2007). Canine population structure: assessment and impact of intra-breed stratification on SNP-based association studies. *PLoS ONE* 2, e1324.
- Ruiz-Linares, A., Adhikari, K., Acuña-Alonzo, V., Quinto-Sanchez, M., Jaramillo, C., Arias, W., Fuentes, M., Pizarro, M., Everardo, P., de Avila, F., et al. (2014). Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. *PLoS Genet.* 10, e1004572.
- Shannon, L.M., Boyko, R.H., Castelhamo, M., Corey, E., Hayward, J.J., McLean, C., White, M.E., Abi Said, M., Anita, B.A., Bondjengo, N.I., et al. (2015). Genetic structure in village dogs reveals a Central Asian domestication origin. *Proc. Natl. Acad. Sci. USA* 112, 13639–13644.
- Sutter, N.B., Eberle, M.A., Parker, H.G., Pullar, B.J., Kirkness, E.F., Kruglyak, L., and Ostrander, E.A. (2004). Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. *Genome Res.* 14, 2388–2396.
- Sutter, N.B., Bustamante, C.D., Chase, K., Gray, M.M., Zhao, K., Zhu, L., Padhukasahasram, B., Karlins, E., Davis, S., Jones, P.G., et al. (2007). A single IGF1 allele is a major determinant of small size in dogs. *Science* 316, 112–115.
- Thalmann, O., Shapiro, B., Cui, P., Schuenemann, V.J., Sawyer, S.K., Greenfield, D.L., Germonpré, M.B., Sablin, M.V., López-Giráldez, F., Domingo-Roura, X., et al. (2013). Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* 342, 871–874.
- Vaysse, A., Ratnakumar, A., Derrien, T., Axelsson, E., Rosengren Pielberg, G., Sigurdsson, S., Fall, T., Seppälä, E.H., Hansen, M.S., Lawley, C.T., et al.; LUPA Consortium (2011). Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genet.* 7, e1002316.
- Vilà, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J., and Wayne, R.K. (1997). Multiple and ancient origins of the domestic dog. *Science* 276, 1687–1689.
- Vilà, C., Maldonado, J.E., and Wayne, R.K. (1999). Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. *J. Hered.* 90, 71–77.
- Vonholdt, B.M., Pollinger, J.P., Lohmueller, K.E., Han, E., Parker, H.G., Quignon, P., Degenhardt, J.D., Boyko, A.R., Earl, D.A., Auton, A., et al. (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 464, 898–902.
- Wang, G.D., Zhai, W., Yang, H.C., Wang, L., Zhong, L., Liu, Y.H., Fan, R.X., Yin, T.T., Zhu, C.L., Poyarkov, A.D., et al. (2016). Out of southern East Asia: the natural history of domestic dogs across the world. *Cell Res.* 26, 21–33.
- Watson, J. (1906). *The Dog Book, Volume II* (Doubleday, Page & Company).
- Wayne, R.K., and Ostrander, E.A. (1999). Origin, genetic diversity, and genome structure of the domestic dog. *BioEssays* 21, 247–257.
- Wilcox, B., and Walkowicz, C. (1995). *Atlas of Dog Breeds of the World, Fifth Edition* (T.F.H. Publications).
- Witt, K.E., Judd, K., Kitchen, A., Grier, C., Kohler, T.A., Ortman, S.G., Kemp, B.M., and Malhi, R.S. (2014). DNA analysis of ancient dogs of the Americas: identifying possible founding haplotypes and reconstructing population histories. *J. Hum. Evol.* 79, 105–118.
- Zhai, G., Zhou, J., Woods, M.O., Green, J.S., Parfrey, P., Rahman, P., and Green, R.C. (2016). Genetic structure of the Newfoundland and Labrador population: founder effects modulate variability. *Eur. J. Hum. Genet.* 24, 1063–1070.

# Entosis Is Induced by Glucose Starvation

Jens C. Hamann,<sup>1,2</sup> Alexandra Surcel,<sup>3</sup> Ruoyao Chen,<sup>1,4</sup> Carolyn Teragawa,<sup>5</sup> John G. Albeck,<sup>5</sup> Douglas N. Robinson,<sup>3</sup> and Michael Overholtzer<sup>1,6,\*</sup>

<sup>1</sup>Cell Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

<sup>2</sup>Louis V. Gerstner, Jr. Graduate School of Biomedical Sciences, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

<sup>3</sup>Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

<sup>4</sup>BCMB Allied Program, Weill Cornell Medical College, New York, NY 10065, USA

<sup>5</sup>Department of Molecular and Cellular Biology, University of California, Davis, Davis, CA 95616, USA

<sup>6</sup>Lead Contact

\*Correspondence: overhom1@mskcc.org

<http://dx.doi.org/10.1016/j.celrep.2017.06.037>

## SUMMARY

Entosis is a mechanism of cell death that involves neighbor cell ingestion. This process occurs in cancers and promotes a form of cell competition, where winner cells engulf and kill losers. Entosis is driven by a mechanical differential that allows softer cells to eliminate stiffer cells. While this process can be induced by matrix detachment, whether other stressors can activate entosis is unknown. Here, we find that entosis is induced in adherent cells by glucose withdrawal. Glucose withdrawal leads to a bimodal distribution of cells based on their deformability, where stiffer cells appear in a manner requiring the energy-sensing AMP-activated protein kinase (AMPK). We show that loser cells with high levels of AMPK activity are eliminated by winners through entosis, which supports winner cell proliferation under nutrient-deprived conditions. Our findings demonstrate that entosis serves as a cellular response to metabolic stress that enables nutrient recovery through neighbor cell ingestion.

## INTRODUCTION

Programmed cell death is essential for promoting proper tissue development and homeostasis and for inhibiting the development of diseases such as cancer. While programmed cell death was once considered to occur only by apoptosis, many alternative forms of cell death have recently been identified that may also regulate cell turnover in a context-dependent manner. For example, regulated forms of necrosis (necroptosis [Degterev et al., 2005] and ferroptosis [Dixon et al., 2012]) and autophagic cell death are now known to contribute to cell death induced by viral infection (Linkermann and Green, 2014) and nutrient deprivation (Gao et al., 2015; Liu et al., 2013) and to programmed cell death during development (Nelson et al., 2014).

In addition to these regulated forms of cell death, other alternative forms have been reported that may represent yet-additional programmed mechanisms that eliminate cells in certain

contexts (Galluzzi et al., 2012). Among these, entosis is a mechanism that targets cells for death following their engulfment by neighboring cells (Overholtzer et al., 2007). Entotic cells are killed non-cell autonomously by engulfing cells through autophagy protein-dependent lysosomal digestion (Florey et al., 2011). Entosis occurs in human cancers, and we have shown that it inhibits transformed growth by inducing cell death. However, this process also promotes the development of aneuploidy in host cells (Krajcovic et al., 2011) and facilitates nutrient recovery by engulfing cells that could function to promote tumor progression (Krajcovic et al., 2013). Recently, we demonstrated that entosis acts as a form of cell competition, where the engulfment of loser cells by neighboring winners can promote clonal selection within heterogeneous tumor cell populations (Sun et al., 2014b). Competition is driven by a mechanical differential between softer (reduced elasticity) cells and stiffer cells, where stiffer cells are eliminated by softer winners (Sun et al., 2014b).

While recent studies have elucidated consequences of entosis on cell populations, the signals that could promote this process remain poorly characterized, with a lack of suitable matrix adhesion as the only clear known inducer of entosis (Overholtzer et al., 2007). As entosis, like autophagy, can allow for nutrient recovery that supports cell survival and proliferation under conditions of starvation (Krajcovic et al., 2013), we considered if entosis might also be induced by nutrient deprivation. Here, we identify glucose starvation, acting through an AMP-activated protein kinase (AMPK)-dependent mechanical differential, as a potent activator of entosis in matrix-adherent cancer cells.

## RESULTS

### Entosis Is Induced in Cancer Cells by Glucose Withdrawal

We previously reported that cell engulfment by entosis allows cells to recover nutrients that can support cell survival and proliferation under conditions of amino acid deprivation (Krajcovic et al., 2013). However, amino acid withdrawal did not induce high levels of entosis, prompting us to examine whether other forms of nutrient starvation could induce high rates of neighbor cell ingestion. To identify potential entotic triggers, we cultured MCF-7 human breast tumor cells, which undergo high levels of entosis in matrix-detached conditions, in various nutrient-depleted conditions.

After 72 hr in medium containing low serum, no glucose, and no amino acids, neighboring cells engulfed each other at high rates, with more than 30% of adherent cells containing an average of two engulfed neighbors (Figure 1A). Often, more complicated cell structures were observed, with three or more cells involved in sequential engulfments (Figures 1Ai and 1Aii), similar to the entotic structures reported in matrix-detached cultures. Interestingly, the withdrawal of glucose from growth medium, unlike starvation for other nutrients, was sufficient to induce a high level of cell engulfment (Figure 1A), and re-addition of D-glucose to glucose-free medium completely rescued this effect (Figure 1C). Glucose starvation therefore appears to be a primary trigger of cell engulfment.

Entotic cell engulfment is known to involve cell contractility of internalizing cells, regulated by RhoA and Rho-kinase (ROCK) signaling, and cell-cell adhesions, mediated by E-cadherin. To determine if cell engulfment induced by glucose starvation occurs by entosis, we examined cells for the presence of these characteristics. First, activated myosin II, indicated by phosphorylation of myosin light chain 2 on the ROCK-dependent site serine 19 (P-MLC<sup>S19</sup>), was localized at the cortex of the internalizing cells, as reported (Sun et al., 2014a) (Figure 1B). Further, treatment with Y-27632, an inhibitor of ROCK and a potent entosis inhibitor, also completely blocked cell engulfment induced by glucose withdrawal (Figure 1C). Second, the cell-cell adhesion protein  $\beta$ -catenin localized at the interface between internalizing and engulfing cell pairs, consistent with a cell-cell-adhesion-based mechanism of uptake (Figure 1B). Accordingly, disruption of E-cadherin in MCF-7 cells, by CRISPR/Cas9-mediated gene editing, significantly inhibited cell engulfment due to glucose withdrawal, which was rescued by exogenous expression of E-cadherin-GFP (Figures 1D and S1A). We also examined whether MDA-MB-231 breast cancer cells, which are E-cadherin deficient and do not normally undergo entosis (Sun et al., 2014a), could undergo cell engulfment in response to glucose deprivation. MDA-MB-231 cells showed no evidence of engulfment in either glucose-starved or full media conditions (Figure 1E). However, upon expression of exogenous E-cadherin, MDA-MB-231 cells exhibited a significant rate of cell engulfment with glucose starvation (Figure 1E). To examine whether this effect was unique to cancer cells or common among non-transformed cells as well, we cultured a non-transformed mammary epithelial cell line (MCF-10A) in glucose-free conditions for multiple days and found that levels of entosis within the culture also increased, although to a lower extent than in MCF-7 cells (Figures S1B and S1C). We therefore conclude that glucose starvation is an inducer of entosis in breast cancer cells and non-transformed mammary epithelial cells.

### Entotic Cell Death Is Increased in Glucose-Starved Conditions

Entosis leads to the death of internalized cells through a non-apoptotic mechanism that involves lipidation of the autophagy protein LC3 onto entotic vacuoles (Florey et al., 2011). Typically, 50%–70% of internalized cells undergo death by this mechanism within a 24 hr period, while others (10%–20%) manage to escape from their hosts (Overholtzer et al., 2007; Florey et al.,

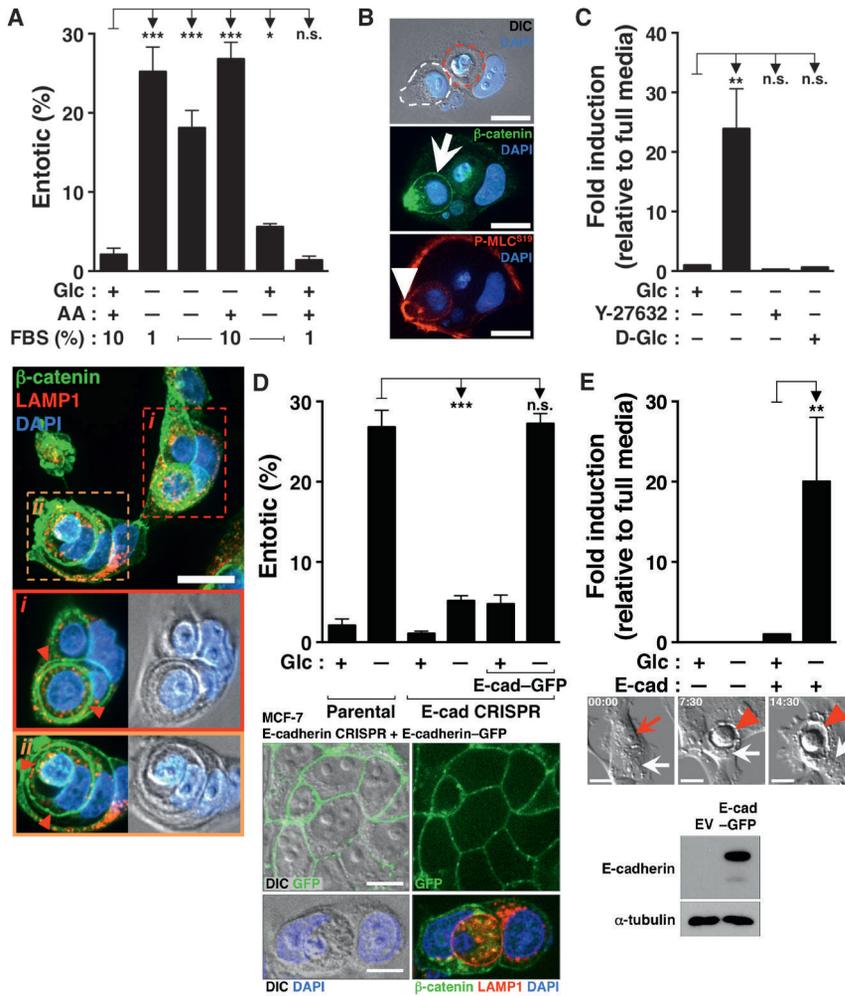
2011). Interestingly, under conditions of glucose withdrawal, we noted that entotic cells died more rapidly after their engulfment and rarely escaped (Figure 2A). More than 60% of engulfed cells died within 5 hr after engulfment, compared to 10% of engulfed cells cultured in full media. Ten hours after engulfment was complete, more than 90% of internalized cells cultured in glucose-free media had undergone cell death, while only ~50% of engulfed cells grown in full media had died (Figure 2A). Like entotic cell deaths occurring in nutrient-rich conditions, those occurring in glucose-starved cultures involved lipidation of LC3 onto entotic vacuoles, and the frequency of cell death was reduced by knockdown of the autophagy protein Atg5 (Figures 2A, 2B, and S2A).

### The Energy-Sensing Kinase AMPK Regulates Loser Cells during Glucose-Starvation-Induced Entosis

To identify the signaling mechanisms that might control glucose starvation-induced entosis, we considered AMPK, a well-known starvation-induced kinase that allows cells to respond to starvation stress by inducing autophagy (Yuan et al., 2013). Treatment with compound C, an inhibitor of AMPK (Zhou et al., 2001), as well as expression of dominant-negative isoforms of AMPK (AMPK-DN) (Mu et al., 2001; Young et al., 2016) inhibited entosis induction in the absence of glucose (Figures 3A, S3A, and S3B). Conversely, entosis was induced in nutrient-rich media by the induction of AMPK activity using two AMPK activators, the AMP analog 5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR) (Sullivan et al., 1994) and the allosteric activator A-769662 (Göransson et al., 2007) (Figures 3A, S3A, and S3C), indicating that AMPK plays an important role in promoting this process.

To determine if AMPK activation occurs in winner or loser cells, we utilized a fluorescence resonance energy transfer (FRET)-based sensor (Tsou et al., 2011) to monitor temporal AMPK dynamics. Shortly before engulfment, we observed an increase in FRET-based fluorescence (indicating increased AMPK activity) within internalizing MCF-7 cells undergoing glucose-withdrawal-induced entosis (Figure 3B). This finding is consistent with previous reports of this sensor that showed similar fold increases in FRET signal upon glucose starvation (Banko et al., 2011; Tsou et al., 2011). We then co-cultured MCF-7 cells expressing AMPK-DN with control MCF-7 cells expressing GFP in the absence of glucose. While control mCherry-expressing MCF-7 cells were winners at ~50% frequency, MCF-7 cells expressing AMPK-DN were winner cells in ~75% of heterotypic entotic structures, consistent with the idea that AMPK acts within loser cells (Figures 3C and S3D). Interestingly, inhibition of AMPK, either by expression of AMPK-DN or treatment with compound C, or activation of AMPK by treatment with AICAR did not alter the rates of cell death once loser cells became internalized, suggesting the rapid rate of loser cell death observed in glucose-free conditions is independent of AMPK activity (Figures S2B and S2C).

We then tested whether glucose deprivation primarily promotes entosis by inducing loser cell status. We took advantage of our previous observation that expression of a constitutively active form of Rac1 (Rac1<sup>V12</sup>, hereafter Rac1<sup>CA</sup>-GFP) can inhibit cell uptake by blocking loser cell internalization while promoting



**Figure 1. Glucose Starvation Induces Entosis in Breast Cancer Cell Lines**

(A) Nutrient starvation causes engulfment. Graph shows the percentage of engulfment in MCF-7 cells grown in the indicated conditions for 72 hr, as determined by immunofluorescence. Glc, glucose; AA, amino acids; FBS, fetal bovine serum. Error bars depict mean  $\pm$  SEM; data are from at least three independent experiments. Images show engulfed cell structures from a 72-hr glucose-starved culture. Immunostaining for  $\beta$ -catenin (green) and Lamp1 (red) and DAPI-stained nuclei (blue) are shown. Arrowheads show completed engulfments. Scale bar, 20  $\mu$ m.

(B) Localization of cellular machinery required for entosis. Immunofluorescence image shows localization of phosphorylated myosin light chain (P-MLC<sup>S19</sup>, red, arrowhead) in MCF-7 cells grown in glucose-free conditions for 72 hr. Cell-cell junctions are shown by  $\beta$ -catenin staining (green, arrow), and DAPI-stained nuclei are shown in blue. The white dashed line shows the outline of partial engulfment, while the red dashed line shows the engulfed corpse. Scale bars, 15  $\mu$ m.

(C) Induction of engulfment is blocked by ROCK inhibition and readdition of glucose. Graph shows the quantification of engulfment over 72 hr, as determined by time-lapse microscopy. Error bars depict mean  $\pm$  SEM; data are from at least three independent experiments.

(D) Loss of E-cadherin blocks entosis. Graph shows percentage of engulfment in parental MCF-7 cells, as well as E-cad CRISPR and E-cad CRISPR + E-cad-GFP cells, grown in the indicated media conditions for 72 hr, as determined by immunofluorescence. Error bars depict mean  $\pm$  SEM; data are from at least three independent experiments. Representative images show expected plasma membrane localization of exogenous E-cadherin-GFP in MCF-7 E-cad CRISPR + E-cad-GFP cells (top row) and induction of entosis in these cells

(bottom row). Immunostaining for  $\beta$ -catenin (green) and Lamp1 (red) and DAPI-stained nuclei (blue) is shown. Scale bars, 10  $\mu$ m. See also Figure S1A.

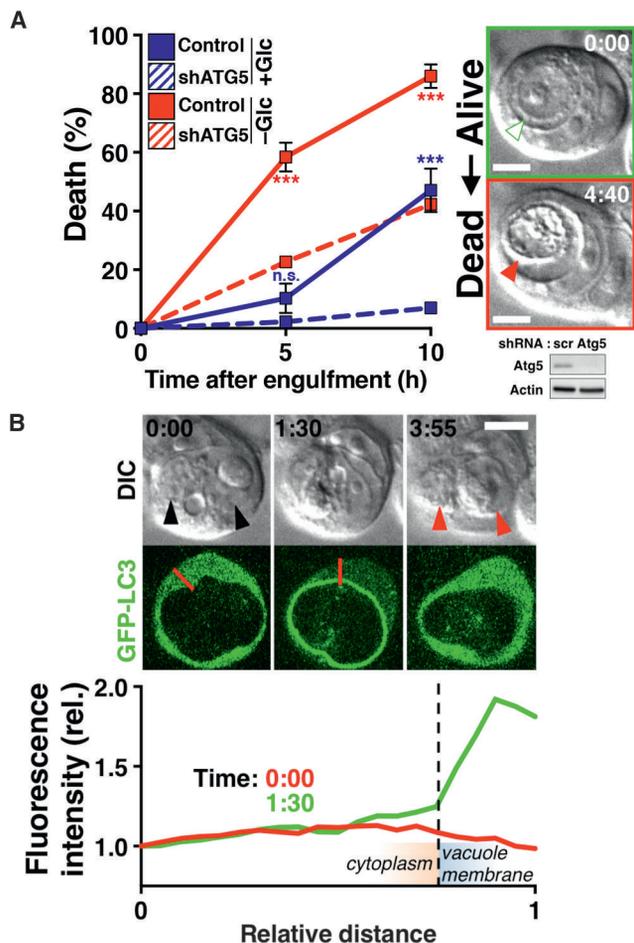
(E) Engulfment is induced during glucose starvation in MDA-MB-231 cells expressing E-cadherin. Shown is the quantification of engulfment over 72 hr for MDA-MB-231 cells expressing either empty vector (EV) or E-cadherin-GFP, cultured in full or glucose-free media, as determined by time-lapse microscopy. Error bars depict mean  $\pm$  SEM; data are from at least three independent experiments. A representative image sequence of MDA-MB-231 cells expressing E-cadherin-GFP undergoing engulfment is shown in glucose-free conditions. The red arrow shows a cell to be internalized before engulfment. White arrows show the host cell. The red arrowheads show engulfment. Time shown is in hours:minutes. Scale bars, 10  $\mu$ m. Western blot shows expression of E-cadherin in parental and E-cadherin-GFP-expressing MDA-MB-231 cells.

winner cell behavior (Sun et al., 2014b). Rac1<sup>CA</sup>-GFP expression completely blocked the induction of entosis in glucose-depleted conditions (Figure 3D), consistent with the model that glucose deprivation activates entosis by stimulating loser cell behavior through increased AMPK activity, and activation of Rac1 can override this signal.

### Glucose Starvation Results in Changes in Cell Deformability

A differential in cell deformability (inverse elasticity) appears to be a requisite for entosis (Sun et al., 2014b). We therefore sought to determine if mechanical deformability is altered in glucose-starved cell populations using micropipette aspiration (MPA). Control cells display a monomodal distribution of deformability (Figure 4A). However, glucose-starved cells were distributed

into a bimodal population based on their mechanical profiles, with cells exhibiting relatively more (low elastic modulus) or less (high elastic modulus) deformability (Figure 4A). The measured elastic modulus of the less deformable cell population that appeared upon glucose starvation was significantly higher than that of MCF-7 cells in nutrient-rich medium, suggesting this population could represent loser cells within the population, similar to what we observed previously (Sun et al., 2014b) (Figure 4C). To examine if AMPK plays a role in controlling the appearance of this population, we performed MPA with control and glucose-starved MCF-7 cells expressing AMPK-DN. After glucose starvation, the number of cells clustering into the high elasticity population was significantly reduced, suggesting that AMPK activity impacts loser cell behavior, in part, by altering cell deformability (Figure 4B).



**Figure 2. Glucose Starvation Increases Entotic Cell Death**

(A) Entotic death rate of internalized cells increases during glucose starvation. Graph shows death rate of internalized cells within 10 hr after completed engulfment, as determined by the appearance of a vacuole within the host cell by DIC microscopy. Entotic death rates for cells cultured in full media (+Glc, blue lines) or glucose-free media (-Glc, red lines) in control MCF-7 cells (solid lines) and MCF-7 shATG5 cells (dashed lines) are shown. Error bars depict mean  $\pm$  SEM; data are from at least three independent experiments, with  $n \geq 40$  entotic cell structures per experiment. *p* values according to two-way ANOVA multiple comparisons test. *p* values compare control and shATG5 cells in +Glc or -Glc media, respectively, and are color-coded accordingly. Western blot shows expression of Atg5 relative to actin in MCF-7 cells expressing a scrambled or Atg5-targeted shRNA. Images are representative of live and dead internalized cells as scored by DIC. The green and white arrowhead shows a live internalized cell ("alive"); the red arrowhead shows a morphologically dead cell inside the host cell vacuole ("dead"). Time shown is in hours:minutes. Scale bars, 10  $\mu$ m. See also Figure S2A.

(B) Internalized cell death involves transient lipidation of LC3 (light chain 3) onto entotic vacuoles. Confocal time-lapse images of MCF-7 cells expressing GFP-LC3 cultured in glucose-free media. Black arrowheads show two live engulfed cells within a GFP-LC3-expressing host (one inner cell contains a corpse). Red arrowheads show death of both internalized cells  $\sim$ 4 hr after LC3 recruitment. Time shown is in hours:minutes. Scale bar, 10  $\mu$ m. Graph shows relative GFP fluorescence intensity both before (0:00) and during (1:30) GFP-LC3 recruitment. The line used for the fluorescence intensity profile is shown in red.

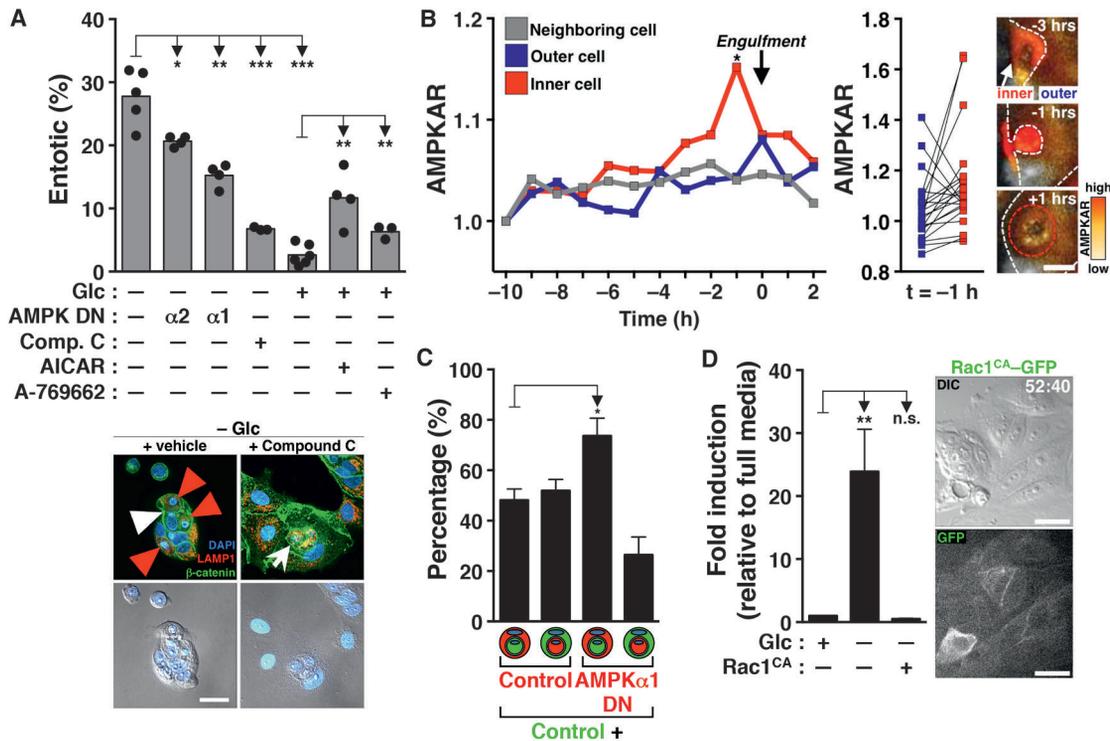
### Entosis Supports Proliferation in Nutrient-Limiting Conditions

We next examined the consequences of glucose-withdrawal-induced entosis on cell populations. We asked whether the ingestion and degradation of loser cells could provide winner cells with nutrients that support cell survival or proliferation during starvation, similar to what we have shown for amino acid starvation (Krajcovic et al., 2013). Indeed, even under stringent conditions of dual glucose and amino acid deprivation, cells that had ingested their neighbors proliferated 10-fold more frequently than control single-cell neighbors (Figures 5A and S4A). We further examined the effects of entosis in glucose-starved conditions over a longer time course. Over the first 72 hr of glucose starvation, the population size was significantly reduced by frequent cell deaths occurring with entotic, necrotic, or apoptotic morphologies, as quantified by time-lapse microscopy (Figures 5B and 5C). Following 72 hr, MCF-7 cells expressing E-cadherin exhibited a population doubling after 6 days of continued growth in the absence of glucose (Figure 5D). Conversely, cells lacking E-cadherin, which are deficient for entosis induction, had an impaired ability to grow under starvation conditions, despite reduced levels of cell death overall, and a slightly increased ability to proliferate under nutrient-replete conditions (Figures 5B and 5D; data not shown). Altogether, these data are consistent with a model that entosis induced by nutrient starvation supports the proliferation of winner cells under conditions of continued nutrient withdrawal.

Typically, entotic cell structures induce the generation of aneuploid cell lineages due to the failure of engulfing cells to divide properly (Krajcovic et al., 2011). Because most cell divisions occurring under starvation conditions involved entotic cell structures, we examined if nutrient withdrawal led to the appearance of multinucleated cells. Indeed, nutrient starvation induced a 5-fold increase in the percentage of cells exhibiting multinucleation (Figures 5E and S4B), suggesting that starvation can disrupt cell ploidy by inducing entosis.

To examine the properties of cancer cell populations selected by glucose starvation, we took advantage of our observation that MCF-7 cells could be grown in the absence of glucose for extended periods (MCF-7<sup>-Glc</sup>). MCF-7<sup>-Glc</sup> cells were selected in continuous culture for 36–78 days in the absence of glucose. We co-cultured MCF-7<sup>-Glc</sup> cells with passage-matched parental starvation-naive MCF-7 cells (MCF-7<sup>parental</sup>) and quantified winner and loser status of each cell population in heterotypic entotic cell structures. In glucose-free conditions, MCF-7<sup>-Glc</sup> displayed a marked increase in winner cell activity (Figure 5F). Thus, cells continually grown in glucose-depleted conditions maintain altered characteristics that confer winner status when co-cultured with parental cells under starvation conditions. Consistent with AMPK controlling loser cell activity, selected cells also exhibited reduced AMPK activation, as well as lowered levels of P-MLC<sup>S19</sup>, indicative of winner cells (Figure 5G).

Altogether, our data demonstrate that glucose starvation induces a high level of entosis in matrix-adherent breast cancer cell populations, in a manner controlled by AMPK activity in loser cells. Glucose starvation results in the emergence of a bimodal mechanical distribution of cells, where AMPK is required for the appearance of the less deformable subpopulation that we



**Figure 3. AMPK Regulates Loser Cells during Glucose-Starvation-Induced Entosis**

(A) Modulation of AMPK pathway activity alters the formation of entotic structures. Graph shows number of entotic structures in MCF-7 cells grown in respective media with or without an AMPK inhibitor or two distinct activators, or MCF-7 cells expressing dominant-negative (DN) isoforms of AMPK (AMPK $\alpha 1$  and  $\alpha 2$ ), for 72 hr, as determined by immunofluorescence. Shown are individual data points from independent experiments. Immunostaining for  $\beta$ -catenin (green) and Lamp1 (red) and DAPI-stained nuclei (blue) is shown. Arrow shows internalized cells, red arrowheads mark entotic corpses, and white arrowhead marks partial engulfment. Scale bar, 15  $\mu$ m. See also Figures S3A–S3C.

(B) AMPK activity increases within internalizing cells during glucose starvation. Graph on the left shows normalized AMPKAR FRET ratio of internalizing cells (“inner cell”) in glucose-free media throughout 12 hr (10 hr prior to completion of entosis and 2 hr after engulfment is complete, as determined by imaging), as well as AMPKAR values of entotic host cells (“outer cell”) and neighboring single cells. Data are from two independent experiments (n = 10 and n = 12, respectively). p values according to two-way ANOVA multiple comparisons test. Graph on right shows FRET values of matched outer and inner cell pairs 1 hr before completion of engulfment. Images show representative entotic engulfment of cells expressing AMPKAR. The white dashed line marks the boundary between cells (inner cell is labeled by white arrow); the red dashed line marks an inner cell within an outer cell after entosis. Scale bar, 10  $\mu$ m.

(C) Overexpression of DN AMPK $\alpha 1$  blocks loser cell behavior in MCF-7 cells cultured in glucose-free media. Graph shows quantification of heterotypic entotic structures of mCherry-expressing AMPK $\alpha 1$ -DN cells cultured 1:1 with parental cells (expressing GFP) and cultured in glucose-free media for 72 hr. Error bars show mean  $\pm$  SEM; data are from at least three independent experiments. See also Figures S3B and S3D.

(D) Overexpression of constitutively active Rac1<sup>V12</sup> (Rac1<sup>CA</sup>-GFP) blocks glucose starvation-induced entosis. Graph shows quantification of entosis events over 72 hr in parental MCF-7 cells or cells expressing constitutively active Rac1 grown in full or glucose-free media, as determined by time-lapse microscopy. Error bars show mean  $\pm$  SEM; data are from at least three independent experiments. Time-lapse images show cells expressing Rac1<sup>CA</sup>-GFP. Time shown is in hours:minutes. Scale bars, 15  $\mu$ m.

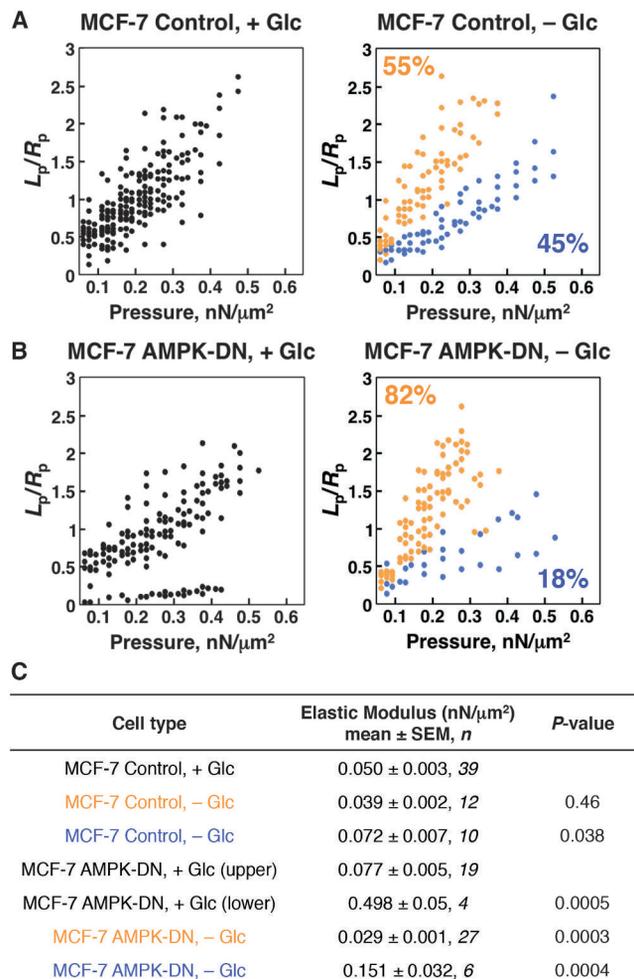
have shown previously has loser cell activity. Long-term glucose starvation also selects for winner cell behavior that is associated with increased proliferation, changes in cell ploidy, and a long-term ability to grow in the absence of glucose.

## DISCUSSION

Here, we find that nutrient deprivation in the form of glucose withdrawal is an inducer of the cell engulfment and death mechanism entosis. Thus, in addition to matrix detachment, metabolic stress resulting from insufficient glucose availability, which is known to occur during tumorigenesis (Denko, 2008), may induce entosis in human cancers. As entosis promotes the scavenging of nutrients by winner cells from losers, its induction in this

context allows cell populations to respond to starvation stress, promoting competition between cells.

We find that glucose starvation induces entosis by activating the energy-sensing kinase AMPK within loser cells. Consistent with this, the activation of Rac1, which blocks loser cell behavior, inhibits starvation-induced entosis. Mechanical measurements of cells undergoing glucose starvation revealed the appearance of a bimodal population of cells consisting of one group that is much more deformable and a second group that is less deformable in a manner dependent on AMPK activity. These data support a model where glucose starvation activates entosis by upregulating loser cell behavior, which is known to be controlled by RhoA and ROCK. Long-term glucose withdrawal can lead to AMPK-dependent cell death by apoptosis (Okoshi et al., 2008;



**Figure 4. Glucose Starvation Alters Cell Deformability in an AMPK-Dependent Manner**

(A) Glucose starvation induces two mechanically distinct cell populations. The scatterplot on the left shows individual measurements of control MCF-7 cells cultured in full media conditions for 48 hr (+Glc). The scatterplot on the right shows control MCF-7 cells cultured in glucose-free media for 48 hr (-Glc). Note the appearance of two distinct populations upon glucose starvation. The y axis represents the deformation of individual cells ( $L_p$ ) normalized by the radius of the micropipette ( $R_p$ ). Cell measurements were color-coded based on each cell's degree of deformability at high pressures. The distribution of slopes for individual cells in both control and -Glc conditions were tested using the dip test for multimodality (R statistical package). The control dataset was consistent with a monomodal distribution ( $p = 0.854$ ), while the -Glc dataset was consistent with a bimodal population distribution ( $p = 0.015$ ).

(B) Expression of AMPK-DN blocks the appearance of less-deformable cell population in glucose starvation. On the left are individual measurements of MCF-7 cells expressing a DN isoform of AMPK (AMPK-DN) grown in full media for 48 hr. On the right are individual measurements of AMPK-DN cells grown in glucose-free conditions for 48 hr. The number of cells present in the blue, less deformable category in glucose starvation is significantly reduced by expression of AMPK-DN (compare to the -Glc graph in A).

(C) Summary of measured mechanical parameters. Font color corresponds to populations labeled in the same color in (A) and (B), respectively. p values listed for MCF-7 control, -Glc set were compared to MCF-7 control, +Glc cells; p values for MCF-7 AMPK-DN, -Glc were compared to MCF-7 AMPK-DN, +Glc (top) cells. p values were obtained by ANOVA with Fisher's least significant difference test.

El-Masry et al., 2012). Our data demonstrate that AMPK can also promote cell death through entosis, an effect potentially linked to the known AMPK-dependent control over myosin contractility (Thaiparambil et al., 2012; Bultot et al., 2009). Entosis may have the unique property of distributing nutrients to winner cells within a starved population, thereby supporting population regrowth following acute induction of cell death that initially reduces cell number. Intriguingly, we find that entosis inhibition by depletion of E-cadherin or treatment with Y-27632 increases rates of necrosis while having no effect on the overall death percentage (Figures 5B and 5C). Future studies to explore if this observed relationship results from co-regulation of these mechanisms will be informative.

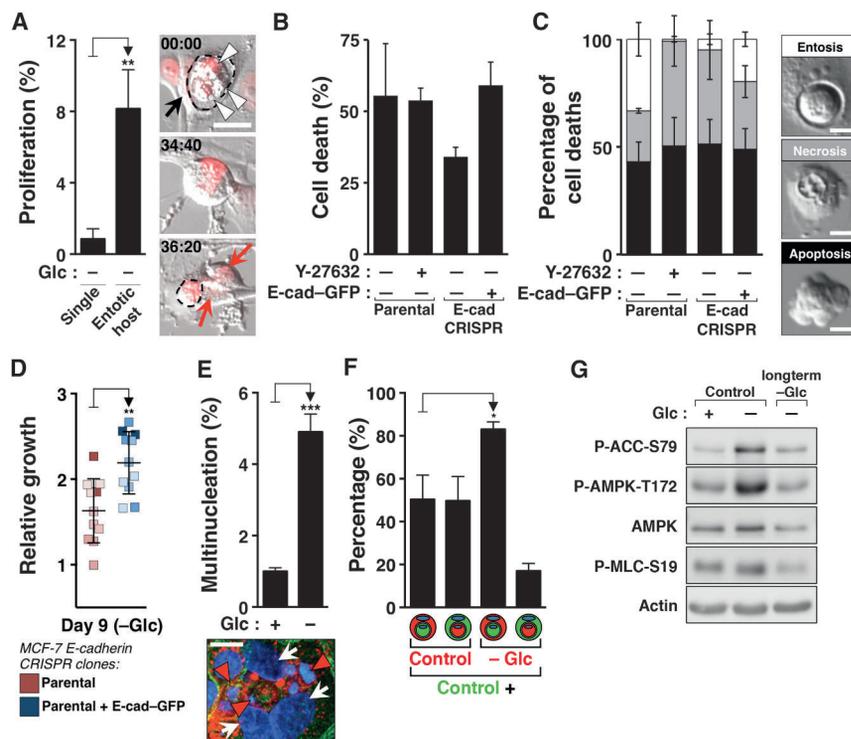
In addition to AMPK-dependent regulation of loser cell mechanics, about half of the cells in the population show increased deformability compared to controls (see Figure 4A). This highly deformable population becomes the dominant population upon AMPK inhibition (Figure 4B). Thus, glucose starvation triggers two events: high levels of AMPK activation in one population, leading to low deformability, and increased deformability in the other population, promoting winner status. In cells cultured in nutrient-replete medium, we have found that activators of AMPK (AICAR and A-769662) are sufficient to induce entosis, but less effectively than glucose starvation, suggesting that winner cell mechanics induced in starved cultures could contribute significantly to entosis induction. How glucose starvation promotes winner cell mechanics, as well as entotic cell death, and whether these activities could be coupled are important questions for further study.

Overall, we show that entosis is induced by glucose starvation and promotes competition between cancer cells. In addition to starvation responses, such as autophagy, which promotes nutrient recycling to support cell survival (Mizushima et al., 2008), and macropinocytosis, which allows cancer cells to scavenge extracellular protein to support proliferation (Commisso et al., 2013; Palm et al., 2015), entosis may be an important mechanism utilized by cancer cell populations to support metabolism under conditions of limiting nutrient availability. In the long-term, some cancer cells may also activate gluconeogenesis to adapt to the continual absence of glucose, as reported previously (Méndez-Lucas et al., 2014). We previously found that entosis disrupts cell ploidy, and we show here that starving cell populations exhibit multinucleation, suggesting that an additional consequence of engaging this mechanism may be to promote tumorigenesis through the promotion of gross aneuploidy. Dying cells have been shown to provide nutrients to support the survival and proliferation of neighboring cells in single-cell yeast and bacterial populations undergoing starvation (Gourlay et al., 2006; Fabrizio et al., 2004; Büttner et al., 2006). Our data suggest that some cancer cell populations may also respond to starvation by redistributing nutrients in a manner that maintains the proliferation of selected cells.

## EXPERIMENTAL PROCEDURES

### Cell Culture and Reagents

MCF-7 cells (Lombardi Cancer Center, Georgetown University, Washington, DC) were cultured in DMEM (11965-092; Life Technologies) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (F2442; Sigma-Aldrich)



**Figure 5. Consequences of Starvation-Induced Entosis**

(A) Entotic host cells have a proliferative advantage in starvation conditions compared to non-engulfing single cells. Graph shows quantification of proliferation events of single and entotic host cells in glucose/amino acid-free media, imaged by time-lapse microscopy over 72 hr after an initial starvation period. Error bars show mean  $\pm$  SEM; data are from at least three independent experiments. Images show entotic host cell (black arrow) with three engulfed corpses (arrowheads) in glucose/amino acid-free media undergoing cell division; red arrows mark two daughter cells. Time shown is in hours:minutes. Scale bar, 15  $\mu$ m. See also Figure S4A.

(B) Inhibition of entosis does not increase overall rates of cell death. Graph shows quantification of all types of cell death in MCF-7 cells over 72 hr of glucose starvation, as determined by time-lapse microscopy. Error bars show mean  $\pm$  SEM; data are from at least three independent experiments.

(C) Inhibition of entosis results in increased rates of necrotic cell death. Graph shows percentage of different types of cell death observed over 72 hr of glucose starvation by time-lapse microscopy. Error bars show mean  $\pm$  SEM; data are from at least three independent experiments. Representative DIC images show the three types of death morphologies used to score cell death. Scale bars, 10  $\mu$ m.

(D) Entosis-competent cells proliferate more than entosis-deficient cells in glucose-free conditions. Scatterplot shows the number of MCF-7 E-cad CRISPR (red) and E-cad CRISPR + E-cad-GFP cells (blue) after 9 days of glucose starvation relative to the cell number after 72 hr. Shown are individual values from three independent experiments (data points from experimental replicates are shaded accordingly). Error bars show mean  $\pm$  SD.

(E) Glucose starvation induces multinucleation in MCF-7 cells. Graph shows quantification of multinucleated cells in cultures of MCF-7 cells grown in the presence or absence of glucose for 72 hr, as determined by immunofluorescence. Error bars show mean  $\pm$  SEM; data are from at least three independent experiments. Image shows appearance of multinucleated cell after 72 hr of glucose starvation, with three nuclei (arrows) and three engulfed cells (arrowheads). Immunostaining for  $\beta$ -catenin (green) and Lamp1 (red) and DAPI-stained nuclei (blue) is shown. Scale bar, 10  $\mu$ m. See also Figure S4B.

(F) Cells cultured under prolonged glucose starvation preferentially internalize parental cells. Graph shows the percentage of red-labeled cells (either parental MCF-7 or MCF-7 cells grown for 36–78 days in glucose-free conditions) that are hosts or internalized cells when mixed with passage-matched, green-labeled parental MCF-7 cells and cultured in glucose-free conditions for 72 hr. Error bars show mean  $\pm$  SEM; data are from at least three independent experiments.

(G) AMPK and myosin activity is reduced in long-term glucose-starved cells. Western blot shows activation of AMPK and downstream signaling, as well as activation of myosin light chain, in MCF-7 cells cultured in the indicated conditions.

and penicillin/streptomycin (30-002-CI; Mediatech). MDA-MB-231 cells and its derivative were grown in RPMI (11875-093; Life Technologies) supplemented with 10% heat-inactivated FBS with penicillin/streptomycin, as described previously (Sun et al., 2014a). MCF-10A cells were cultured in DMEM/F-12 (11320-033; Life Technologies) supplemented with 5% horse serum (HS) (S12150; Atlanta Biologicals), 20 ng/mL epidermal growth factor (EGF) (AF-100-15; Peprotech), 10  $\mu$ g/mL insulin (I-1882; Sigma-Aldrich), 0.5  $\mu$ g/mL hydrocortisone (H-0888; Sigma-Aldrich), 100 ng/mL cholera toxin (C-8052; Sigma-Aldrich), and penicillin/streptomycin (30-002-CI; Mediatech), as previously described (Debnath et al., 2003). Glucose-free, amino acid-free, and glucose/amino acid-free medium was prepared by dialyzing heat-inactivated FBS for 4 hr at 4°C in PBS (P3813; Sigma-Aldrich) in MWCO 3500 dialysis tubing (21-152-9; Fisherbrand), followed by overnight at 4°C in fresh PBS and subsequent addition to base media prepared without respective component (glucose, amino acids, or glucose and amino acids, respectively) to a 10% final concentration. Cells expressing the H2B-mCherry nuclear marker were prepared by transducing cells with retroviruses made with the pBabe-H2B-mCherry construct, as described previously (Florey et al., 2011). The mCherry-AMPK-DN construct was generated by inserting the AMPK $\alpha$ 2 K45R gene (from Plasmid #15992; Addgene) into the pQCXIP-mCherry retroviral vector. pEGFP-Rac1<sup>CA</sup> was a gift from Dr. Alan Hall (Memorial Sloan Kettering Cancer Center, New York, NY). For disruption of E-cadherin in

MCF-7 cells by CRISPR-Cas9, guide RNAs (gRNAs) were designed using the online CRISPR design tool from Feng Zhang's laboratory (<http://crispr.mit.edu>), and the gRNA with one of the highest theoretical mutagenic efficiencies was used (5'-CGCCGAGAGCTACACGTTCCACGG-3'). The vector encoding for Cas9 (pCDNA3.3-TOPO-hCas9, plasmid #41815; Addgene), as well as vector encoding the gRNA (pCR-Blunt II-TOPO, plasmid #41824; Addgene), were introduced into control MCF-7 cells by nucleofection (Cell Line Nucleofector Kit V, VCA-1003; Lonza). Single-cell clones were selected and examined for disruption of E-cadherin by sequencing and western blotting. Cells were treated with Y-27632 (#1254; Tocris Bioscience) at 10  $\mu$ M, AICAR (#9944; Cell Signaling Technology) at 2 mM, A-769662 at 500 nM (#3336; Tocris Bioscience), and compound C (P5499; Sigma-Aldrich) at 10  $\mu$ M. Inhibitors (or vehicle) were added to cultures ~30 min before the start of biological assays unless indicated otherwise.

#### Western Blotting

Cells were lysed in ice-cold RIPA buffer and western blotting was performed as described previously (Florey et al., 2011). The following antibodies were used: anti-E-cadherin (1:500; 3195, Cell Signal), anti-tubulin (1:2,000; 3873, Cell Signal), anti-Atg5 (1:500; 2630, Cell Signal), anti-phospho-ACC-S79 (1:500; 3661, Cell Signal), anti-ACC (1:500; 3662, Cell Signal), anti-phospho-AMPK-T172 (1:500; 2531, Cell Signaling), anti-mCherry (1:500; ab125096, Abcam),

anti- $\beta$ -actin (1:2000; A1978, Sigma-Aldrich), anti-rabbit immunoglobulin G (IgG) horseradish peroxidase (HRP)-linked antibody (1:5,000; 7074, Cell Signal), and anti-mouse IgG HRP-linked antibody (1:5,000; 7076, Cell Signal).

### Immunofluorescence

The following antibodies were used for immunofluorescence (IF): anti- $\beta$ -catenin (1:100; C2206; Sigma-Aldrich), anti-Lamp1 (1:100; 555798; BD Biosciences), Alexa Fluor 568 goat anti-mouse secondary (1:500; A-11031; Life Technologies), and Alexa Fluor 488 goat anti-rabbit secondary (1:500; A-11034; Life Technologies). IF was performed on cells cultured on glass-bottom dishes (P35G-1.5-20-C; MatTek), as described previously (Overholtzer et al., 2007). Briefly, cells were fixed in 1:1 methanol/acetone for 5 min at  $-20^{\circ}\text{C}$ , followed by three 5-min PBS washes and blocking in 5% BSA, 100 mM glycine in PBS for 1 hr, followed by incubation with primary antibodies at  $4^{\circ}\text{C}$  overnight. Samples were then incubated with secondary antibodies and counterstained with DAPI (1:1,000; D1306; Life Technologies). Confocal microscopy was performed with the Ultraview Vox spinning-disk confocal system (PerkinElmer) equipped with a Yokogawa CSU-X1 spinning-disk head and an electron-multiplying charge-coupled device camera (Hamamatsu C9100-13) coupled to a Nikon Ti-E microscope; image analysis was done using Volocity software (PerkinElmer).

### Time-Lapse Microscopy

Cells were cultured on glass-bottom dishes (P06G-1.5-20-F; MatTek), and time-lapse microscopy was performed in  $37^{\circ}\text{C}$  and 5%  $\text{CO}_2$  live-cell incubation chambers, as described previously (Florey et al., 2011). Fluorescence and differential interference contrast (DIC) images were acquired every 20 min for 72 hr using a Nikon Ti-E inverted microscope attached to a CoolSNAP charge-coupled device camera (Photometrics) and NIS Elements software (Nikon). For tetramethylrhodamine, ethyl ester, perchlorate (TMRE) imaging, TMRE (T669; Thermo Fisher Scientific) was added to cultures to be imaged at a final concentration of 100 nM.

### Entosis Quantification

For quantification of entosis in MCF-7 cells by immunofluorescence, 250,000 cells were plated on 35-mm glass-bottom dishes and allowed to adhere overnight, washed briefly three times with PBS, cultured in the indicated conditions for 72 hr, and fixed and stained as described above. The percentage of entotic cells was determined by counting at least 300 cells in each sample and quantifying the number of single cells and cell-in-cell structures; both dead (LAMP1-positive compartments) and live cells were counted as entotic. If one host cell contained two cells in separate compartments, it was scored as two engulfment events. In sequential cell-in-cell structures (see Figures 1Ai and 1Aii), only the outermost cell was counted as an entotic host. Fold induction of entosis in MCF-7 and MDA-MB-231 cells by time-lapse microscopy was determined as follows: engulfment events were scored by the appearance of a vacuole within the host cell throughout 72 hr and normalized to cell number at time 0 hr; normalized engulfment numbers from all conditions were then normalized to full media conditions ("fold induction"). For fate of internalized cells, time 0 of engulfment was determined by the appearance of a host cell vacuole; cell death was scored by changes in cell morphology in the DIC channel. For quantification of entosis in MCF-10A cells by immunofluorescence, 500,000 cells were plated on 60-mm tissue culture dishes (353002; Corning) and allowed to adhere overnight, washed briefly three times with PBS, and cultured in the indicated media for 96 hr (media was replaced after 48 hr). After 96 hr, cells were trypsinized (25053C1; Corning) to achieve a single-cell suspension, and 250,000 cells were re-plated on 35-mm glass-bottom dishes in fresh media (either full or glucose-free) and allowed to adhere for 12 hr (for a total of 108 hr of starvation), at which point cells were fixed and stained as described above.

### AMPK FRET Measurements

A modified version of the AMPK activity reporter (AMPKAR) (Tsou et al., 2011) with an extended "EV" linker (Komatsu et al., 2011) was stably integrated into MCF-7 cells using PiggyBac transposase. Homogeneous populations of reporter-expressing cells were isolated by limited dilution cloning; three

independent clones were analyzed. Time-lapse imaging was performed as previously described (Sparta et al., 2015) using a Nikon TiE with a  $20\times 0.75$  NA Plan-Apochromat objective and CFP and YFP filter cubes (Chroma 49001 and 49003, respectively). Images were recorded with an Andor Zyla sCMOS camera, using Nikon Elements software. For imaging, cells were plated on #1.5 glass-bottom 96-well plates (P96-1.5H-N; In Vitro Scientific) and maintained at  $37^{\circ}\text{C}$  and 5%  $\text{CO}_2$ . Raw image files were imported to ImageJ, where tracking of cells and measurement of fluorescence values was performed manually using a modified version of the "Manual Tracking" plugin. AMPKAR signals were calculated as the average background-subtracted CFP/YFP ratio within three 3-by-3-pixel regions of the cytoplasm.

### MPA Assay

MPA was performed as described previously (Sun et al., 2014b; Zhou et al., 2010). In short, after culturing cells in either full or glucose-free media for 48 hr, cells were trypsinized, pelleted by spinning down for 5 min at  $1,500\times g$ , and re-suspended in the appropriate media. Prior to being measured, the cells were incubated for 15 min at  $37^{\circ}\text{C}$  and 5%  $\text{CO}_2$ . Cells were aspirated with 6- to 8- $\mu\text{m}$  pipettes at varying and increasing pressures. For data analysis, the length of the deformation of the cell cortex pulled into the micropipette ( $L_p$ ) was divided by the radius of the pipette ( $R_p$ ), and  $L_p/R_p$  values were plotted as a function of applied pressure. These data were converted into apparent elastic moduli (Hochmuth, 2000). The distribution of slopes for individual cells in both the control and the  $-\text{Glc}$  conditions were tested using the dip test for multimodality (R statistical package), and the resulting subpopulations were analyzed using ANOVA with Fisher's least significant difference test.

### Proliferation Advantage Assay

250,000 cells per 35-mm well were plated on glass and allowed to adhere overnight. Cells were washed in PBS three times and grown in glucose/amino acid-free media for 72 hr to induce entosis. After, cells were washed in PBS and either full or glucose/amino acid-free media containing 10  $\mu\text{M}$  Y-27632 was added to cells to inhibit further cell engulfment; cells were then imaged for 72 hr at 20-min intervals. Cell fates of entotic hosts or single cells in each field of view were determined throughout this time.

### Population Growth Assay

100,000 cells were seeded in triplicate in 12-well culture dishes (#3512; Corning) and allowed to adhere overnight. Cells were washed in PBS three times and grown in glucose-free media for 9 days, with three PBS washes and media replacement every 72 hr. For crystal violet staining, cells were washed once in PBS, fixed in 4% paraformaldehyde (PFA) in PBS for 15 min, washed once with  $\text{H}_2\text{O}$ , and subsequently stained with 0.1% crystal violet solution (in 10% ethanol) for 20 min. Crystal violet solution was aspirated, and cells were washed three times with  $\text{H}_2\text{O}$  and allowed to air-dry overnight. The following day, crystal violet was extracted by incubating the cells with 1 mL 10% acetic acid for 20 min with gentle shaking, followed by absorbance measurements at 570 nm. Values were normalized to respective absorbance at day 3 of glucose starvation.

### Quantification of Winner and Loser Cell Identity

Passage-matched cells (either control MCF-7 or  $-\text{Glc}$  MCF-7 (starved for glucose for 36–78 days with change of media every 3 days) were labeled with 10  $\mu\text{M}$  CellTracker dyes (green or red, C7025 and C34552, respectively; Life Technologies) for 20 min at  $37^{\circ}\text{C}$  and then plated at a 1:1 ratio at a total cell density of 250,000 cells in 35-mm glass-bottom dishes overnight in media containing 10  $\mu\text{M}$  Y-27632 to block entosis. The next day, cells were washed three times with PBS and glucose-free media was added for 72 hr, at which point cells were analyzed by confocal microscopy. Heterotypic cell-in-cell structures were counted and the number of structures of green-inside-red and red-inside-green determined. For competition assays with AMPK $\alpha$ 1 DN cells, stable mCherry-expressing MCF-7 cells were transfected (using Amara Nucleofector, VCA-1003; Lonza) according to the manufacturer's protocol with either a vector expressing AMPK  $\alpha$ 1 DN (K47R) (plasmid #79011; Addgene) or empty vector and allowed to recover overnight. Cells were then mixed at a 1:1 ratio with stable GFP-expressing MCF-7 cells and plated and analyzed for cell-in-cell structures according to the same protocol described

above. For competition assays with AMPK $\alpha$ 2 DN cells, MCF-7 cells stably expressing either mCherry (control) or mCherry-AMPK $\alpha$ 2 K45R DN were mixed at a 1:1 ratio with stable GFP-expressing MCF-7 cells and plated and analyzed as described above.

### Statistics

The indicated p values were obtained using Student's t test unless otherwise noted (\*\*\*, p < 0.001; \*\*, p < 0.01; \*, p < 0.05; n.s., not significant).

### SUPPLEMENTAL INFORMATION

Supplemental Information contains four figures and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.06.037>.

### AUTHOR CONTRIBUTIONS

J.C.H. and M.O. designed and carried out experiments and wrote the paper. D.N.R. and A.S. designed, performed, and analyzed the MPA experiments and participated in writing the paper. J.G.A. and C.T. designed, performed, and analyzed the AMPKAR FRET experiments. R.C. generated CRISPR knockout cell lines. All authors participated in editing the manuscript.

### ACKNOWLEDGMENTS

This work was supported by NIH grants CA154649 (to M.O.) and GM66817 (to D.N.R.) and the Benjamin Friedman Research Fund (to M.O.). We thank members of the Overholtzer lab for critical reading of the manuscript and Jennifer Nguyen and Pablo Iglesias (Johns Hopkins University School of Medicine) for help with statistical analysis of MPA data.

Received: September 9, 2016

Revised: May 5, 2017

Accepted: June 13, 2017

Published: July 5, 2017

### REFERENCES

Banko, M.R., Allen, J.J., Schaffer, B.E., Wilker, E.W., Tsou, P., White, J.L., Villén, J., Wang, B., Kim, S.R., Sakamoto, K., et al. (2011). Chemical genetic screen for AMPK $\alpha$ 2 substrates uncovers a network of proteins involved in mitosis. *Mol. Cell* **44**, 878–892.

Bultot, L., Horman, S., Neumann, D., Walsh, M.P., Hue, L., and Rider, M.H. (2009). Myosin light chains are not a physiological substrate of AMPK in the control of cell structure changes. *FEBS Lett.* **583**, 25–28.

Büttner, S., Eisenberg, T., Herker, E., Carmona-Gutierrez, D., Kroemer, G., and Madeo, F. (2006). Why yeast cells can undergo apoptosis: death in times of peace, love, and war. *J. Cell Biol.* **175**, 521–525.

Commisso, C., Davidson, S.M., Soydaner-Azeloglu, R.G., Parker, S.J., Kamphorst, J.J., Hackett, S., Grabocka, E., Nofal, M., Drebin, J.A., Thompson, C.B., et al. (2013). Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. *Nature* **497**, 633–637.

Debnath, J., Muthuswamy, S.K., and Brugge, J.S. (2003). Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **30**, 256–268.

Degterev, A., Huang, Z., Boyce, M., Li, Y., Jagtap, P., Mizushima, N., Cuny, G.D., Mitchison, T.J., Moskowitz, M.A., and Yuan, J. (2005). Chemical inhibitor of nonapoptotic cell death with therapeutic potential for ischemic brain injury. *Nat. Chem. Biol.* **1**, 112–119.

Denko, N.C. (2008). Hypoxia, HIF1 and glucose metabolism in the solid tumour. *Nat. Rev. Cancer* **8**, 705–713.

Dixon, S.J., Lemberg, K.M., Lamprecht, M.R., Skouta, R., Zaitsev, E.M., Gleason, C.E., Patel, D.N., Bauer, A.J., Cantley, A.M., Yang, W.S., et al. (2012). Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* **149**, 1060–1072.

El-Masry, O.S., Brown, B.L., and Dobson, P.R. (2012). Effects of activation of AMPK on human breast cancer cell lines with different genetic backgrounds. *Oncol. Lett.* **3**, 224–228.

Fabrizio, P., Battistella, L., Vardavas, R., Gattazzo, C., Liou, L.L., Diaspro, A., Dossen, J.W., Gralla, E.B., and Longo, V.D. (2004). Superoxide is a mediator of an altruistic aging program in *Saccharomyces cerevisiae*. *J. Cell Biol.* **166**, 1055–1067.

Florey, O., Kim, S.E., Sandoval, C.P., Haynes, C.M., and Overholtzer, M. (2011). Autophagy machinery mediates macroendocytic processing and entotic cell death by targeting single membranes. *Nat. Cell Biol.* **13**, 1335–1343.

Galluzzi, L., Vitale, I., Abrams, J.M., Alnemri, E.S., Baehrecke, E.H., Blagosklonny, M.V., Dawson, T.M., Dawson, V.L., El-Deiry, W.S., Fulda, S., et al. (2012). Molecular definitions of cell death subroutines: recommendations of the Nomenclature Committee on Cell Death 2012. *Cell Death Differ.* **19**, 107–120.

Gao, M., Monian, P., Quadri, N., Ramasamy, R., and Jiang, X. (2015). Glutaminolysis and transferrin regulate ferroptosis. *Mol. Cell* **59**, 298–308.

Göransson, O., McBride, A., Hawley, S.A., Ross, F.A., Shpiro, N., Foretz, M., Viollet, B., Hardie, D.G., and Sakamoto, K. (2007). Mechanism of action of A-769662, a valuable tool for activation of AMP-activated protein kinase. *J. Biol. Chem.* **282**, 32549–32560.

Gourlay, C.W., Du, W., and Ayscough, K.R. (2006). Apoptosis in yeast—mechanisms and benefits to a unicellular organism. *Mol. Microbiol.* **62**, 1515–1521.

Hochmuth, R.M. (2000). Micropipette aspiration of living cells. *J. Biomech.* **33**, 15–22.

Komatsu, N., Aoki, K., Yamada, M., Yukinaga, H., Fujita, Y., Kamioka, Y., and Matsuda, M. (2011). Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Mol. Biol. Cell* **22**, 4647–4656.

Krajcovic, M., Johnson, N.B., Sun, Q., Normand, G., Hoover, N., Yao, E., Richardson, A.L., King, R.W., Cibas, E.S., Schnitt, S.J., et al. (2011). A non-genetic route to aneuploidy in human cancers. *Nat. Cell Biol.* **13**, 324–330.

Krajcovic, M., Krishna, S., Akkari, L., Joyce, J.A., and Overholtzer, M. (2013). mTOR regulates phagosome and entotic vacuole fission. *Mol. Biol. Cell* **24**, 3736–3745.

Linkermann, A., and Green, D.R. (2014). Necroptosis. *N. Engl. J. Med.* **370**, 455–465.

Liu, Y., Shoji-Kawata, S., Sumpster, R.M., Jr., Wei, Y., Ginet, V., Zhang, L., Posner, B., Tran, K.A., Green, D.R., Xavier, R.J., et al. (2013). Autosis is a Na<sup>+</sup>,K<sup>+</sup>-ATPase-regulated form of cell death triggered by autophagy-inducing peptides, starvation, and hypoxia-ischemia. *Proc. Natl. Acad. Sci. USA* **110**, 20364–20371.

Méndez-Lucas, A., Hyroššová, P., Novellasdemunt, L., Viñals, F., and Perales, J.C. (2014). Mitochondrial phosphoenolpyruvate carboxykinase (PEPCK-M) is a pro-survival, endoplasmic reticulum (ER) stress response gene involved in tumor cell adaptation to nutrient availability. *J. Biol. Chem.* **289**, 22090–22102.

Mizushima, N., Levine, B., Cuervo, A.M., and Klionsky, D.J. (2008). Autophagy fights disease through cellular self-digestion. *Nature* **451**, 1069–1075.

Mu, J., Brozinick, J.T., Jr., Valladares, O., Bucan, M., and Birnbaum, M.J. (2001). A role for AMP-activated protein kinase in contraction- and hypoxia-regulated glucose transport in skeletal muscle. *Mol. Cell* **7**, 1085–1094.

Nelson, C., Ambros, V., and Baehrecke, E.H. (2014). miR-14 regulates autophagy during developmental cell death by targeting ip3-kinase 2. *Mol. Cell* **56**, 376–388.

Okoshi, R., Ozaki, T., Yamamoto, H., Ando, K., Koida, N., Ono, S., Koda, T., Kamijo, T., Nakagawara, A., and Kizaki, H. (2008). Activation of AMP-activated protein kinase induces p53-dependent apoptotic cell death in response to energetic stress. *J. Biol. Chem.* **283**, 3979–3987.

Overholtzer, M., Mailleux, A.A., Mouneimne, G., Normand, G., Schnitt, S.J., King, R.W., Cibas, E.S., and Brugge, J.S. (2007). A nonapoptotic cell death process, entosis, that occurs by cell-in-cell invasion. *Cell* **131**, 966–979.

Palm, W., Park, Y., Wright, K., Pavlova, N.N., Tuveson, D.A., and Thompson, C.B. (2015). The utilization of extracellular proteins as nutrients is suppressed by mTORC1. *Cell* **162**, 259–270.

- Sparta, B., Pargett, M., Minguet, M., Distor, K., Bell, G., and Albeck, J.G. (2015). Receptor level mechanisms are required for epidermal growth factor (EGF)-stimulated extracellular signal-regulated kinase (ERK) activity pulses. *J. Biol. Chem.* *290*, 24784–24792.
- Sullivan, J.E., Carey, F., Carling, D., and Beri, R.K. (1994). Characterisation of 5'-AMP-activated protein kinase in human liver using specific peptide substrates and the effects of 5'-AMP analogues on enzyme activity. *Biochem. Biophys. Res. Commun.* *200*, 1551–1556.
- Sun, Q., Cibas, E.S., Huang, H., Hodgson, L., and Overholtzer, M. (2014a). Induction of entosis by epithelial cadherin expression. *Cell Res.* *24*, 1288–1298.
- Sun, Q., Luo, T., Ren, Y., Florey, O., Shirasawa, S., Sasazuki, T., Robinson, D.N., and Overholtzer, M. (2014b). Competition between human cells by entosis. *Cell Res.* *24*, 1299–1310.
- Thaiparambil, J.T., Eggers, C.M., and Marcus, A.I. (2012). AMPK regulates mitotic spindle orientation through phosphorylation of myosin regulatory light chain. *Mol. Cell. Biol.* *32*, 3203–3217.
- Tsou, P., Zheng, B., Hsu, C.H., Sasaki, A.T., and Cantley, L.C. (2011). A fluorescent reporter of AMPK activity and cellular energy stress. *Cell Metab.* *13*, 476–486.
- Young, N.P., Kamireddy, A., Van Nostrand, J.L., Eichner, L.J., Shokhirev, M.N., Dayn, Y., and Shaw, R.J. (2016). AMPK governs lineage specification through Tfeb-dependent regulation of lysosomes. *Genes Dev.* *30*, 535–552.
- Yuan, H.X., Xiong, Y., and Guan, K.L. (2013). Nutrient sensing, metabolism, and cell growth control. *Mol. Cell* *49*, 379–387.
- Zhou, G., Myers, R., Li, Y., Chen, Y., Shen, X., Fenyk-Melody, J., Wu, M., Ventre, J., Doebber, T., Fujii, N., et al. (2001). Role of AMP-activated protein kinase in mechanism of metformin action. *J. Clin. Invest.* *108*, 1167–1174.
- Zhou, Q., Kee, Y.S., Poirier, C.C., Jelinek, C., Osborne, J., Divi, S., Surcel, A., Will, M.E., Eggert, U.S., Müller-Taubenberger, A., et al. (2010). 14-3-3 coordinates microtubules, Rac, and myosin II to control cell mechanics and cytokinesis. *Curr. Biol.* *20*, 1881–1889.

# Host Genotype and Gut Microbiome Modulate Insulin Secretion and Diet-Induced Metabolic Phenotypes

Julia H. Kreznar,<sup>1,7</sup> Mark P. Keller,<sup>2,7</sup> Lindsay L. Traeger,<sup>1</sup> Mary E. Rabaglia,<sup>2</sup> Kathryn L. Schueler,<sup>2</sup> Donald S. Stapleton,<sup>2</sup> Wen Zhao,<sup>5</sup> Eugenio I. Vivas,<sup>1</sup> Brian S. Yandell,<sup>3,4</sup> Aimee Teo Broman,<sup>5</sup> Bruno Hagenbuch,<sup>6</sup> Alan D. Attie,<sup>2,\*</sup> and Federico E. Rey<sup>1,8,\*</sup>

<sup>1</sup>Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, USA

<sup>2</sup>Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, USA

<sup>3</sup>Department of Statistics, University of Wisconsin-Madison, Madison, WI 53706, USA

<sup>4</sup>Department of Horticulture, University of Wisconsin-Madison, Madison, WI 53706, USA

<sup>5</sup>Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI 53706, USA

<sup>6</sup>Department of Pharmacology, Toxicology and Therapeutics, University of Kansas, Kansas City, KS 66160, USA

<sup>7</sup>Co-first author

<sup>8</sup>Lead Contact

\*Correspondence: [adattie@wisc.edu](mailto:adattie@wisc.edu) (A.D.A.), [ferey@wisc.edu](mailto:ferey@wisc.edu) (F.E.R.)

<http://dx.doi.org/10.1016/j.celrep.2017.01.062>

## SUMMARY

Genetic variation drives phenotypic diversity and influences the predisposition to metabolic disease. Here, we characterize the metabolic phenotypes of eight genetically distinct inbred mouse strains in response to a high-fat/high-sucrose diet. We found significant variation in diabetes-related phenotypes and gut microbiota composition among the different mouse strains in response to the dietary challenge and identified taxa associated with these traits. Follow-up microbiota transplant experiments showed that altering the composition of the gut microbiota modifies strain-specific susceptibility to diet-induced metabolic disease. Animals harboring microbial communities with enhanced capacity for processing dietary sugars and for generating hydrophobic bile acids showed increased susceptibility to metabolic disease. Notably, differences in glucose-stimulated insulin secretion between different mouse strains were partially recapitulated via gut microbiota transfer. Our results suggest that the gut microbiome contributes to the genetic and phenotypic diversity observed among mouse strains and provide a link between the gut microbiome and insulin secretion.

## INTRODUCTION

The intestinal microbiota exerts a profound influence on development, physiology, and health (Clemente et al., 2012; Sommer and Bäckhed, 2013; Tremaroli and Bäckhed, 2012). Although there is substantial interpersonal variation in the composition of the gut microbiota among unrelated healthy subjects, sequencing studies have revealed distal gut community patterns associated with different pathological states, including obesity

and diabetes (Ridaura et al., 2013; Qin et al., 2012; Karlsson et al., 2013). Remarkably, alterations in the intestinal microbiota composition have been shown to modulate insulin sensitivity (Vrieze et al., 2010), a key feature in metabolic disease and type 2 diabetes (T2D), and thus play a role in diabetes susceptibility.

Dietary components that are not efficiently absorbed in the proximal intestine reach the distal gut, where they are metabolized by gut microbes. Intestinal microbes impact our health in part by generating numerous metabolites from our diet. Short-chain fatty acids (SCFAs), mainly acetate, propionate, and butyrate, are produced through bacterial fermentation of dietary carbohydrates. SCFAs serve as energy and signaling molecules in the intestine and peripheral organs (den Besten et al., 2013). Specifically, SCFAs are important regulators of both energy and glucose homeostasis (den Besten et al., 2013; Koh et al., 2016). For example, butyrate improves insulin sensitivity (Gao et al., 2009; Hartstra et al., 2015) and T2D patients have reduced levels of butyrate-producing bacteria (Qin et al., 2012). Additionally, acetate modulates insulin secretion from  $\beta$  cells (Priyadarshini et al., 2015; Perry et al., 2016). While primarily associated with metabolic benefits, increased concentrations of butyrate and acetate have been found in the cecum of obese mice, suggesting an increased ability of the microbiome to harvest energy from the diet (Turnbaugh et al., 2006).

Gut microbes also impact host physiology by modifying bile acids (BAs) synthesized by the host (Houten et al., 2006; Kuipers et al., 2014; Ryan et al., 2014; Sayin et al., 2013). In addition to their role in emulsifying lipids, BAs function as hormones through their ability to activate nuclear hormone receptors (Parks et al., 1999) and G-coupled protein receptors (Kawamata et al., 2003). They modulate glucose homeostasis, lipid metabolism, energy expenditure, and intestinal motility (Kuipers et al., 2014). Primary BAs are synthesized from cholesterol in the liver (Russell, 2009), stored in the gallbladder, and secreted into the duodenum upon ingestion of a meal. The gut microbiota catalyzes the production of secondary BAs via deconjugation, dehydrogenation, epimerization, and dehydroxylation of primary BAs

(Ridlon et al., 2006). BAs with different modifications vary in their ability to activate receptors and affect host physiology (Makishima et al., 1999; Kuipers et al., 2014). Subjects with T2D have altered circulating BA profiles. Treatment of T2D subjects with compounds that increase fecal excretion of BAs and modify BA composition improves their glycemic status (Handelsman, 2011).

Mouse genetics can be employed to explore the relationships between diet, host genetics, and metabolic responses (O'Connor et al., 2014; Parks et al., 2013; Ussar et al., 2015). The Collaborative Cross (CC) is a systems genetics mouse resource that consists of a panel of recombinant inbred lines and an outbred stock derived from eight genetically diverse founder strains. These include five classical inbred strains (A/J, C57BL/6J, 129S1/SvImJ, non-obese diabetic [NOD]/ShiLtJ, and NZO/HILtJ) and three wild-derived strains (CAST/EiJ, PWK/PhJ, and WSB/EiJ; Churchill et al., 2004; Roberts et al., 2007; Aylor et al., 2011).

We examined the metabolic phenotypes and gut microbiota composition of the eight CC founder strains in response to chronic consumption of two defined diets: a high-fat/high-sucrose diet (HF/HS) and a control diet. We found remarkable variation in diabetes-related phenotypes and gut microbiota composition as a function of host genotype and diet, and we identified bacterial taxa that correlate with metabolic traits, including body weight, glucose, and insulin levels. Germ-free (GF) mice were colonized with microbiota derived from two founder strains that exhibited divergent metabolotypes, C57BL/6J and CAST/EiJ. The transplanted animals were maintained on the HF/HS diet and then subjected to metabolomic and metagenomic analyses. We identified functional differences attributable to the two transplanted microbial communities, including insulin secretion responses and susceptibility to diet-induced metabolic disease.

## RESULTS

### Host Metabolic Responses to Diet Are Influenced by Genetic Background

We assessed the variability of diet-induced metabolic responses of the eight genetically diverse CC founder strains: A/J; C57BL/6J (B6); 129S1/SvImJ (129); NOD/ShiLtJ (NOD); NZO/HILtJ (NZO); CAST/EiJ (CAST); PWK/PhJ (PWK); and WSB/EiJ (WSB). All mice were obtained from The Jackson Laboratory, maintained in the same vivarium, and fed the same diet so that the only known difference among the strains is genetics. We placed 4-week-old male mice from each strain on either a control or a HF/HS diet for 22 weeks (Table S1).

The CC founder strains displayed a wide range of body weight and metabolic responses to the dietary challenge (Figures 1 and S1). Two-way ANOVA analysis of the clinical traits revealed a significant strain effect for fasting insulin ( $F = 14.94$ ;  $p < 0.0001$ ). We also observed significant strain-diet interactions for body weight ( $F = 3.19$ ;  $p < 0.01$ ) and fasting glucose ( $F = 2.81$ ;  $p < 0.01$ ). Significant strain and diet effects were also seen for hepatic triglyceride content ( $F = 10.96$ ,  $p < 0.0001$ ;  $F = 11.92$ ,  $p < 0.001$ , respectively). Liver triglyceride content showed high inter-strain variation, with 129 having the most significant response to diet ( $p < 0.05$ ; Figure 1D). NZO mice were the only strain to become

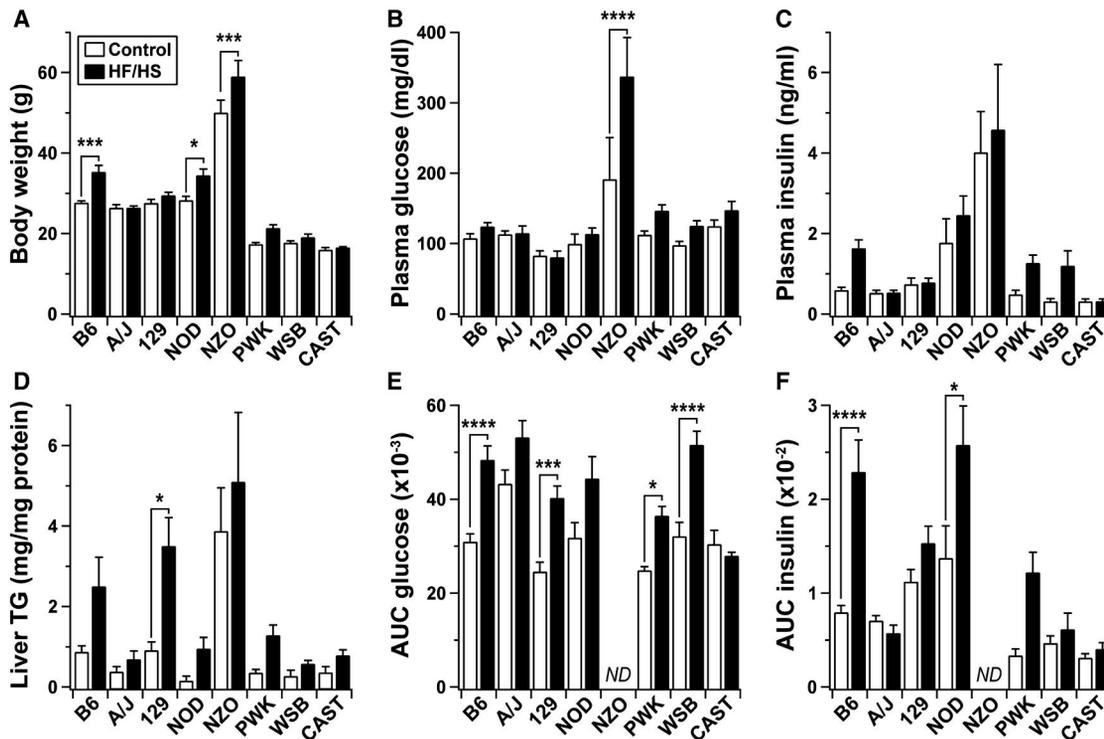
overtly diabetic (glucose levels  $> 300$  mg/dL) as a consequence of HF/HS feeding. With the exception of NZO mice, which did not survive past 18 weeks on the HF/HS diet, B6 mice were the most responsive to diet. HF/HS-fed B6 mice became obese ( $p < 0.01$ ) and developed insulin resistance and glucose intolerance after  $\sim 8$  weeks (Figures 1A and S1A–S1C). In addition to differences in diet responsiveness, the strains varied in both absolute levels of insulin and change in insulin levels over time, suggesting a significant divergence in insulin sensitivity among the strains (Figure S1B).

To assess whole-body glucose homeostasis and more directly evaluate the underlying role of the pancreatic islets in the control of plasma insulin, we measured plasma glucose and insulin during an oral glucose tolerance test (oGTT). Both plasma glucose and insulin during the oGTT varied dramatically between the strains. We computed the area under the curve (AUC) for each trait to determine the overall excursion in glucose and insulin that occurred during the oGTT (Figures 1E, 1F, and S2). We observed a wide inter-strain range of responses in plasma insulin during the oGTT ( $F = 12.84$ ;  $p < 0.0001$ ; Figures 1F and 2B). Changes in plasma insulin may reflect altered insulin secretion from  $\beta$  cells, peripheral insulin resistance, reduced insulin clearance, or any combination thereof. 129 and WSB showed diet-induced glucose intolerance but minimal changes in their insulin response during the oGTT (Figures 1E, 1F, and S2A), suggesting that their glucose intolerance may be driven by altered insulin secretion and/or enhanced insulin clearance. Remarkably, insulin secretion and glucose tolerance were completely unaffected by the HF/HS diet in CAST. Furthermore, the kinetics of the glucose and insulin responses were more rapid in CAST than in all other strains (Figure S2), suggesting that CAST mice may employ different pathways underlying glucose-stimulated insulin secretion and whole-body glucose disposal.

### Diet and Host Genotype Influence Microbiota Composition

Gut microbes influence the development of metabolic disease. We characterized the cecal microbiomes of the eight CC founder strains by 16S rRNA sequencing. We compared the cecal microbiomes employing UniFrac, a phylogenetic distance metric used to measure differences in bacterial community structure (Lozupone and Knight, 2005). Principal-coordinates analysis (PCoA) of 16S rRNA unweighted UniFrac distances revealed a strong influence of strain (PERMANOVA;  $p < 0.001$ ) and diet (PERMANOVA;  $p < 0.001$ ) on microbial community composition (Figure S3A). Consistent with previous studies, the effect of diet on gut microbial composition varied among the strains (O'Connor et al., 2014; Parks et al., 2013; Carmody et al., 2015), where B6, CAST, and NOD mice showed the greatest microbiome response to diet (Figure S3A).

We detected eight bacterial phyla among the mice (Figure S3B). Bacteroidetes and Firmicutes dominated the gut of all strains on either diet, accounting for  $>90\%$  of the sequenced reads. As reported by other studies, we observed a decrease in the Bacteroidetes:Firmicutes ratio and an increase in Proteobacteria in the HF/HS-fed mice (Ley et al., 2005; Hildebrandt et al., 2009). In fact, Proteobacteria showed the greatest fold change in abundance in response to diet: HF/HS feeding caused an



**Figure 1. Segregation of Metabolic Syndrome among CC Founder Mice**

Male mice were maintained on the high-fat/high-sucrose (HF/HS) or a control diet for 22 weeks beginning at 4 weeks of age.

(A–D) Body weight (A), fasting plasma glucose (B) and insulin (C), and hepatic triglyceride (D) content determined for all mice at 26 weeks of age.

(E and F) Areas under the curve (AUC) for (E) glucose and (F) insulin during oral glucose tolerance test (oGTT) conducted at 22 weeks of age. Insulin and glucose values were determined from plasma following a 4-hr fast. No data (ND) were collected for NZO mice during oGTT.

In all panels, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and \*\*\*\* $p < 0.0001$  by two-way ANOVA (diet and strain) with Bonferroni's multiple comparisons test to assess within-strain differences. Data are mean  $\pm$  SEM;  $n \geq 9$  mice/genotype/diet.

average 5.4-fold change ( $p < 0.0001$ ), although the relative increase varied among strains.

### Microbial Taxa Correlate with Metabolic Phenotypes

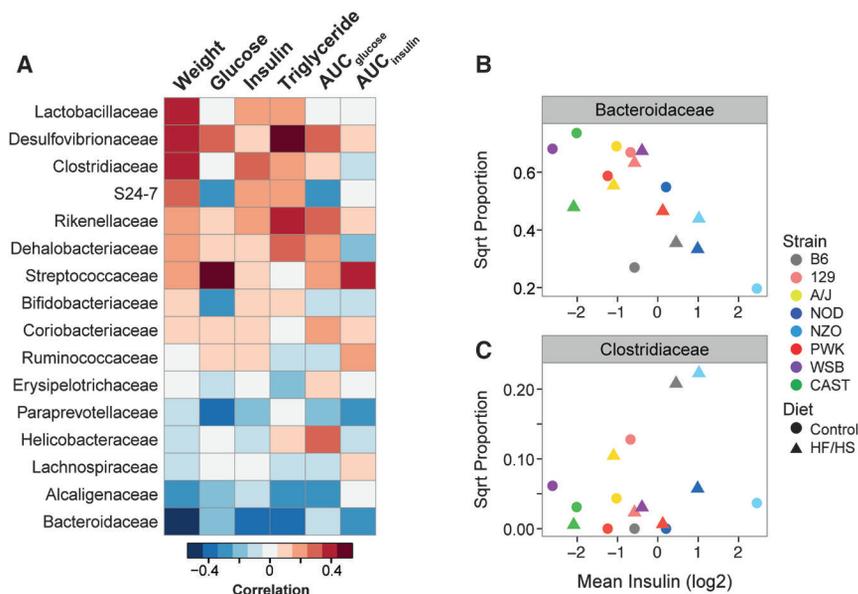
To determine whether strain-dependent variability in microbiota composition was associated with the dramatic differences in the diabetes-related clinical traits, we computed Pearson's correlations between abundance of family-level taxa and the metabolic traits among the eight CC founder mice (Figure 2A). We focused our analysis on families that were present in at least seven of the founder strains. Bacteroidaceae was among the most negatively correlated with several metabolic phenotypes, including body weight, fasting plasma insulin, and AUC<sub>insulin</sub> during the oGTT. The Bacteroidaceae family belongs to the Bacteroidetes phylum and is typically found at higher levels in fecal samples of lean versus obese individuals (Ley et al., 2005; Turnbaugh et al., 2009). Conversely, Clostridiaceae and Rikenellaceae showed the strongest positive correlations with plasma insulin levels. Our analysis also identified strong positive correlations between fasting plasma glucose and the Streptococcaceae and Desulfovibrionaceae families. Members of these families have previously been shown to be enriched in the fecal microbiome of patients with T2D (Qin et al., 2012; Karlsson et al., 2013).

Some of the correlations mentioned above varied significantly as a function of host diet and strain (Table S2). For example, the negative correlation observed between fasting insulin levels and Bacteroidaceae had a significant strain effect ( $p < 0.0001$ ). We also observed a slight diet effect ( $p < 0.001$ ), which is likely driven by the low abundance and high fasting insulin levels in the chow-fed NZO mice (Figure 2B). We also observed a significant diet effect for the relationship between Clostridiaceae and fasting insulin levels ( $p < 0.05$ ), but there was also a strain difference that seems to be driven by NZO on chow diet ( $p < 0.001$ ; Figure 2C).

These results suggest that diet and genetic background are major determinants of gut microbial composition and metabolic disease. However, the relative contributions of host genetic variance versus microbial-derived genetic variation across different mouse strains in the development of diet-induced metabolic phenotypes remain largely unknown.

### The Gut Microbiome Is a Source of Genetic Variation that Influences Host-Associated Differences in Diet-Induced Metabotypes

To directly test the influence of gut microbes on the metabolic phenotypes observed among the founder strains, we performed cecal transplants into germ-free B6 (B6-GF) hosts, leveraging two CC founder strains that showed disparate responsiveness



**Figure 2. Gut Microbial Taxa Correlate with Metabolic Phenotypes**

(A) Heatmap illustrates Pearson's pairwise correlation between microbial families and diabetes-related clinical traits measured in the eight CC founder mice ( $n \geq 9$  mice/genotype/diet). Microbial families are ordered by their correlation to body weight. Red, positive correlation; blue, negative. Area under the curve (AUC) values for insulin and glucose were computed from oGTT conducted at 22 weeks; other metrics were collected at 26 weeks. Correlation coefficients and p values are found in Table S2.

(B and C) Contributions of strain and diet on the correlations observed between fasting insulin and (B) the Bacteroidaceae family and (C) the Clostridiaceae family.

to the HF/HS diet. The B6 strain became obese, insulin resistant, and glucose intolerant, whereas the CAST strain remained lean and insulin-sensitive despite HF/HS feeding (Figure 1).

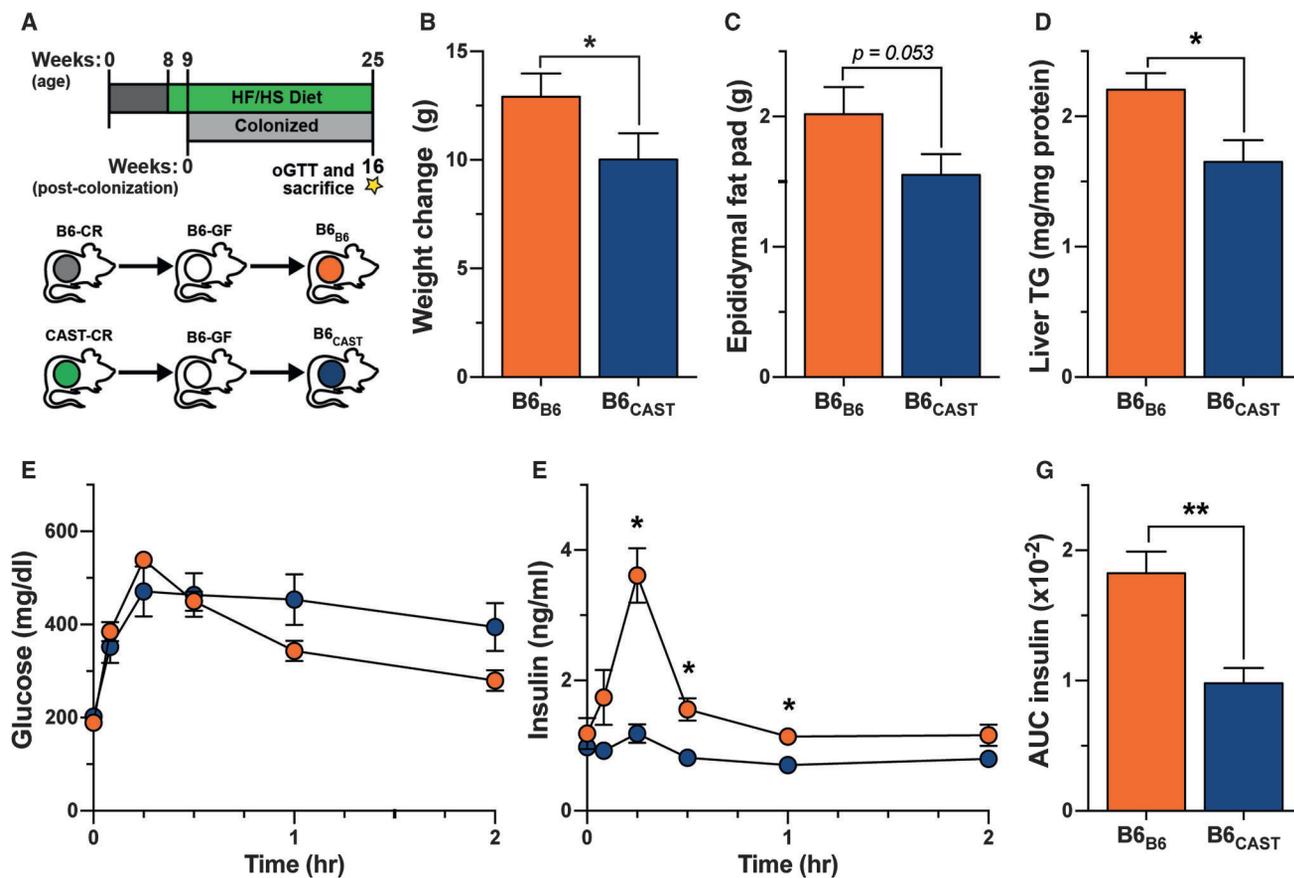
As mentioned above, B6 and CAST mice had significantly different intestinal microbiota (PERMANOVA;  $F = 4.86$ ;  $p < 0.001$ ; Figure S3A). B6 mice harbored a significantly greater abundance of microbial families with strong positive correlations with metabolic traits, such as weight and insulin (i.e., Clostridiaceae;  $p < 0.05$ ), whereas CAST mice had a greater representation of families with significant negative correlations (i.e., Bacteroidaceae;  $p < 0.01$ ; Figures 2A and S3C).

We transplanted cecal microbiota from either conventionally raised B6 (B6-CR) or CAST (CAST-CR) donor mice into 9-week-old B6-GF recipient mice to yield B6<sub>B6</sub> or B6<sub>CAST</sub> mice, respectively. Transplanted animals were housed by treatment group in separate vinyl gnotobiotic isolators and maintained on a HF/HS diet for 16 weeks following colonization (Figure 3A). A dietary treatment of 16 weeks allows robust development of metabolic phenotypes associated with consumption of HF/HS diet.

Recipient mice recapitulated microbial and metabolic phenotypes observed in the respective donor strains (Figures 3 and 4). B6<sub>B6</sub> mice gained ~25% more weight, had larger epididymal fat pad mass, and showed greater hepatic triglyceride accumulation than B6<sub>CAST</sub> mice (Figure 3). Additionally, oGTT revealed that, whereas the plasma glucose levels resulting from an orally administered bolus of glucose did not significantly differ between the two groups of transplanted mice (Figure 3E), the insulin responses were dramatically different (Figure 3F). The glucose challenge evoked a much larger insulin response in B6<sub>B6</sub> mice than in B6<sub>CAST</sub> mice. The low insulin response in B6<sub>CAST</sub> mice resembled the insulin response of the CAST-CR donors (Figures 1F and S1F). These results suggest that the effectiveness of insulin to maintain euglycemia was greater in the mice receiving the CAST microbiota than in mice receiving the B6 microbiota (Figures 3E and 3F).

16S rRNA gene profiling of the donor cecal inoculum and transplant recipient fecal samples show that recipient mice were successfully colonized with the donor's microbiota. B6<sub>B6</sub> and B6<sub>CAST</sub> mice assumed a phylogenetically similar composition to that of their respective donors as confirmed by PCoA of unweighted UniFrac distances (Figure 4A). As seen in the founders, Bacteroidetes and Firmicutes comprised ~90% of the microbiome, although the abundance of Firmicutes was higher in B6<sub>B6</sub> ( $p < 0.05$ ; Figure 4B). We identified taxonomic differences in the microbiota composition between the two recipient groups using linear discriminant analysis (LDA) effect size (LefSe) with LDA score  $> 2$  (Segata et al., 2011). We found 20 microbial families that were differentially enriched in the fecal microbiota of B6<sub>B6</sub> versus B6<sub>CAST</sub> mice. There were 12 microbial families that were enriched in B6<sub>B6</sub>, of which seven belonged to the Firmicutes phyla (Figure 4C). Some of the families differentially represented in the transplanted animals overlap with taxa that are significantly correlated with metabolic phenotypes in the founder strains (Figure 2). Notably, B6<sub>B6</sub> mice exhibited higher levels of Clostridiaceae ( $p < 0.01$ ), which is positively associated with insulin secretion in the founder strains (Figure 2), whereas B6<sub>CAST</sub> mice had higher levels of Bacteroidaceae ( $p < 0.01$ ), which is negatively associated with body weight and insulin secretion (Figure 2). These results are concordant with the metabolic phenotypes observed in the transplanted mice and suggest that the distinct microbial gut communities influence metabolic changes evoked by HF/HS feeding, including insulin secretion.

We characterized the functional potential of transplanted communities by sequencing and analyzing their metagenomes. Metagenomic analysis of the same samples further validated that the B6- and CAST-derived microbiota were distinct from one another, with donors clustering with their respective transplant recipients (Figure 4D). We identified several thousand genes differentially represented between the B6 and CAST microbiota (Table S4). This metagenomic analysis also revealed microbial functions that were enriched in each transplanted microbial community (Table S5). The most enriched microbial pathways in B6<sub>B6</sub> mice included genes involved in membrane



**Figure 3. Divergent Effects of B6 and CAST Microbiomes on Diet-Induced Metabolic Phenotypes**

(A) Transplant experimental design.

(B–D) Total weight change (B), epididymal fat pad mass (C), and quantification of hepatic triglyceride (TG) contents (D).

(E–G) Glucose and insulin values during oGTT (E and F) and AUC insulin (G) in B6<sub>B6</sub> and B6<sub>CAST</sub> mice. All measurements shown were collected 16 weeks post-colonization.

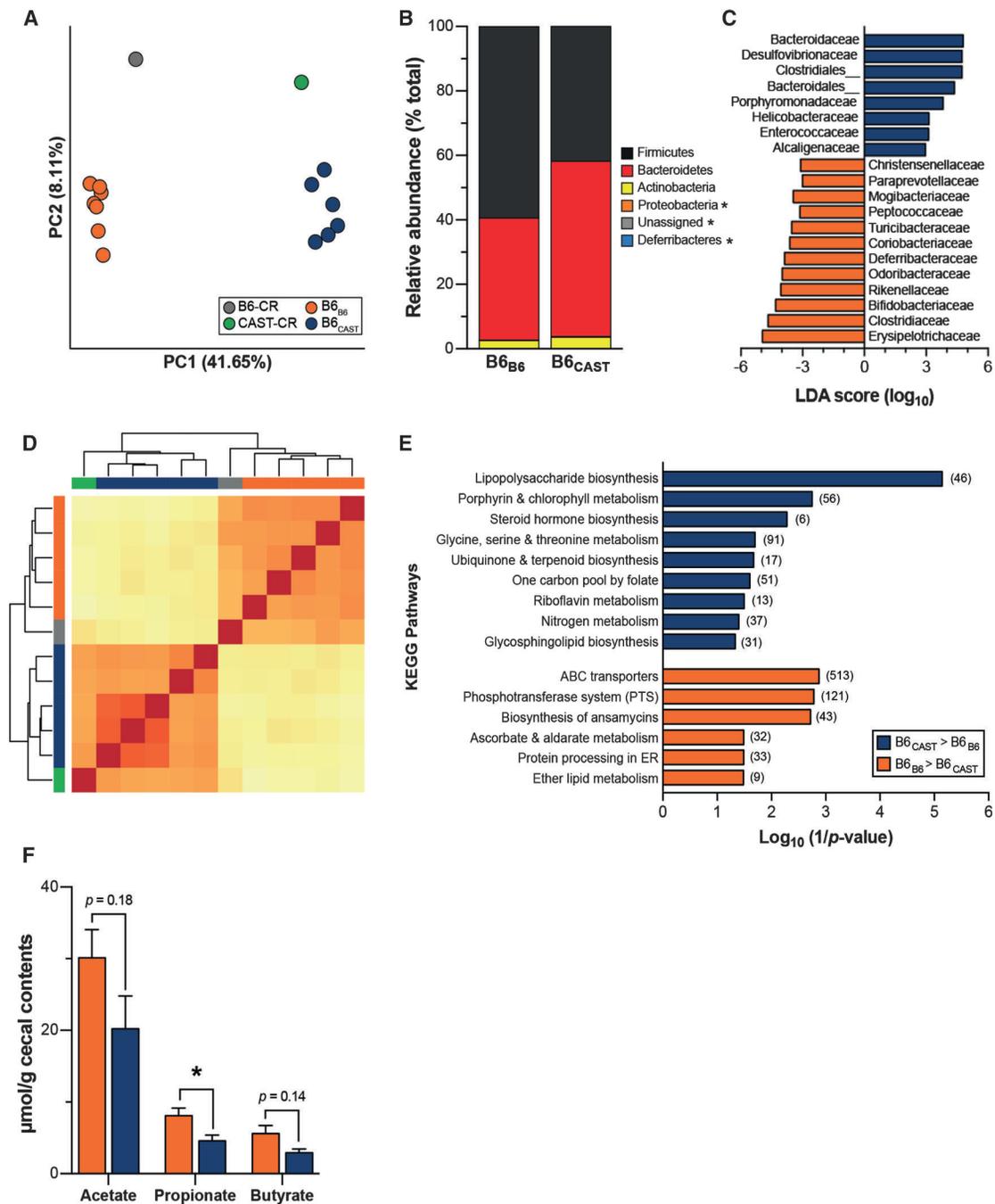
\*p < 0.05; \*\*p < 0.01 by Student's t test. Data are mean ± SEM; n = 7 for B6<sub>B6</sub> and n = 6 for B6<sub>CAST</sub> mice.

transport and carbohydrate and lipid metabolism (Figure 4E). For example, the ABC transporters and phosphotransferase system (PTS) pathways were enriched in mice colonized with the B6 microbiota (p < 0.01). PTS are a class of transport systems involved in the uptake and phosphorylation of a variety of carbohydrates that can be subsequently fermented to SCFAs (Deutscher et al., 2006). It has been previously reported that diet-induced obese mice have a concomitant enrichment of microbial pathways involved in PTS and elevated concentrations of SCFAs (Turnbaugh et al., 2008), reflecting an increased capacity for energy harvest. Consistent with these results, targeted gas chromatography-mass spectrometry (GC-MS) analysis of SCFAs in cecal contents disclosed that B6<sub>B6</sub> mice had an increased concentration of the major fermentation end products compared with B6<sub>CAST</sub> (Figure 4F). Conversely, B6<sub>CAST</sub> microbiota were enriched in genes related to the vitamin B12 (cobalamin) biosynthetic pathway (Figure S4A), synthesis of other B vitamins and enzyme co-factors, as well as lipopolysaccharide (LPS) biosynthesis (Figures 4E and S4B). A difference in LPS biosynthetic potential may reflect the composition of the B6<sub>CAST</sub> microbiota,

which has a significantly higher relative abundance of gram-negative Bacteroidetes than the B6<sub>B6</sub> microbiota (Figure S3B). Our findings mirror those described previously in T2D patients relative to diabetes-free control patients (Qin et al., 2012; Karlsson et al., 2013)—both the microbiota of T2D patients and our metabolically diseased mice with B6 microbiota show enrichment in Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways involved in membrane transport, whereas diabetes-free patients and mice with the CAST microbiota exhibit enrichment in vitamin and co-factor biosynthesis.

### B6 and CAST Microbiota Produce Divergent Bile Acid Profiles

Gut microbes impact host physiology in part by modulating the composition of the BA pool. We determined fecal BA profiles of the transplanted mice and HF/HS-fed B6-CR and CAST-CR mice by ultra performance liquid chromatography-mass spectrometry (UPLC/MS)-based quantification of primary and the most abundant secondary BAs. The BA composition of B6<sub>B6</sub> mice closely resembled that of B6-CR donor mice, whereas



**Figure 4. Gut Microbiota Composition and Function of Transplant Recipients**

(A) Principal-coordinate analysis (PCoA) of unweighted UniFrac distances for the fecal microbiota of transplant donors and recipients at sacrifice. Each circle represents an individual mouse. Percent variation explained by each PC is shown in parentheses.

(B) Relative abundance of major microbial phyla ordered by increasing mean abundance; \* denotes mean phyla abundance <1%.

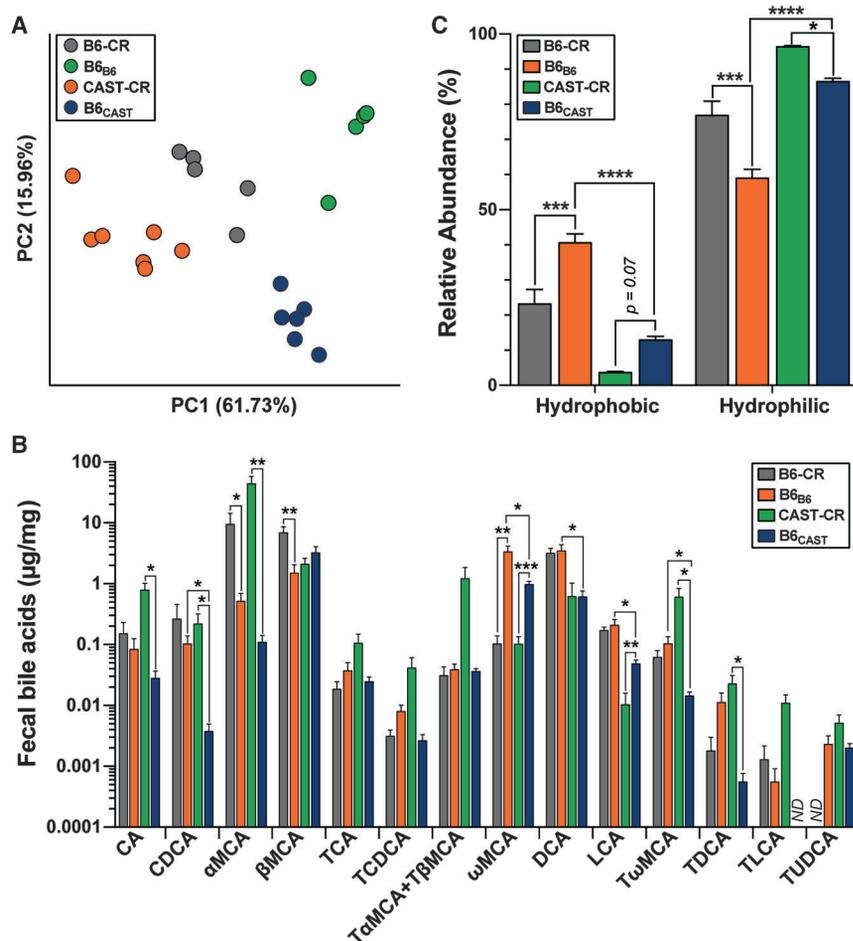
(C) Microbial families differentially enriched in either B6<sub>CAST</sub> (blue) or B6<sub>B6</sub> (orange) as determined by linear discriminant analysis (LDA) with effect size (LEfSe).

(D) Clustering of mice based on relative abundance of KEGG metabolic pathways using Euclidian distance measurement with complete linkage hierarchical clustering; B6-CR (gray), CAST-CR (green), B6<sub>B6</sub> (orange), and B6<sub>CAST</sub> (blue).

(E) KEGG categories enriched in either CAST (blue) or B6 (orange) transplanted microbiomes.

(F) Targeted GC-MS analysis of cecal short-chain fatty acids; \* $p < 0.05$  by Student's t test.

Data are mean  $\pm$  SEM;  $n = 6-7$  mice/recipient group and  $n = 2-3$  mice/donor group. For metagenomics analysis,  $n = 5$  mice/recipient group.



**Figure 5. B6 and CAST Microbiota Produce Different Bile Acid Profiles**

(A) Principal-component analysis of the square root proportion of 14 major bile acid species (ng/mg). Each dot represents the bile acid profile of an individual mouse. Percent variation explained by each PC is shown in parentheses.

(B and C) Abundance of fecal bile acids (B) and relative abundance of hydrophobic and hydrophilic BA species (C) determined by UPLC-MS/MS from fecal samples collected at 12 weeks post-colonization. \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$  by one-way ANOVA with Bonferroni's multiple comparisons test.

Data are mean  $\pm$  SEM;  $n = 6-7$  for transplant recipients and  $n = 5$  for CR mice.

B6<sub>CAST</sub> exhibited a BA profile that was intermediate between CAST-CR and B6-CR mice (Figure 5A). Microbiota composition was also a significant predictor of BA composition. Bray-Curtis dissimilarity-based principal-component analysis (PCA) revealed clustering of the BA profiles by microbiota composition.

Although the B6<sub>CAST</sub> microbiota composition resembled that of CAST-CR (Figure 4A), there were significant differences in BA profiles between these groups, suggesting that variation in circulating BAs is under the control of both host genetics and gut microbiota. For example, the primary BAs cholic acid (CA), chenodeoxycholic acid (CDCA), and  $\alpha$ -muricholic acid ( $\alpha$ -MCA) were significantly higher in CAST-CR mice compared to B6<sub>CAST</sub> mice ( $p < 0.01$ ,  $p < 0.05$ , and  $p < 0.01$ , respectively; Figure 5B). Moreover, taurine-conjugated muricholic acids (MCAs) were significantly higher in CAST-CR mice compared with B6<sub>CAST</sub> mice. In contrast, these differences in taurine conjugation were not present between B6-CR and B6<sub>B6</sub> mice. Taurine conjugation of MCAs is a host process (Ridlon et al., 2006), further highlighting the interaction of host genetics and microbiome in modulating host BA profiles.

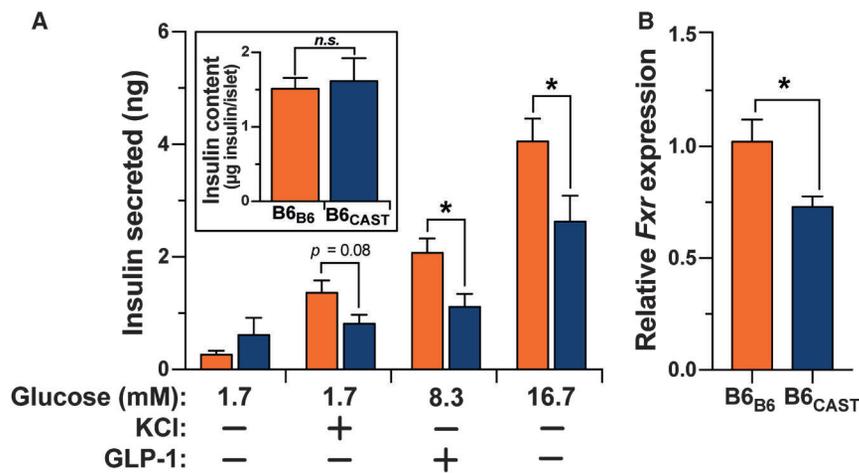
B6-CR and B6<sub>B6</sub> mice had a significantly greater representation of hydrophobic BA species (e.g., deoxycholic acid and lithocholic acid; Figures 5B and 5C), which are elevated in humans

and mice with insulin resistance (Ryan et al., 2014; Prawitt et al., 2011). Microbial metabolism of bile acids generally leads to a more hydrophobic bile acid pool, which facilitates fecal elimination of bile acids. Bile salt hydrolases (BSH) are involved in the hydrolysis of conjugated BAs, a necessary step for the production of secondary BAs. Consistent with the results presented above, there were a higher number of distinct BSH genes in the B6 microbiota relative to CAST microbiota (13 annotated BSH genes highly abundant in the B6 microbiota relative to CAST versus two annotated BSH genes highly abundant in the CAST microbiota relative to B6; Table S4). Furthermore, the two

groups of recipient mice had vastly different fecal BA profiles. Chenodeoxycholic acid (CDCA) ( $p < 0.05$ ), deoxycholic acid (DCA) ( $p < 0.01$ ), lithocholic acid (LCA) ( $p < 0.01$ ),  $\omega$ -muricholic acid ( $\omega$ MCA) ( $p < 0.05$ ), and tauro- $\omega$ -muricholic acid (T $\omega$ MCA) ( $p < 0.05$ ) were all significantly higher in B6<sub>B6</sub> than in B6<sub>CAST</sub> (Figure 5B). DCA was the most abundant BA species in B6<sub>B6</sub> mice and was also  $\sim 5$ -fold more abundant in B6-CR versus CAST-CR mice. DCA contributes to microbial dysbiosis, a hallmark of metabolic disease, and is positively associated with higher levels of Firmicutes (Islam et al., 2011). Tauroursodeoxycholic acid (TUDCA) was  $>2$ -fold higher in CAST-CR mice compared to the transplanted animals but was not detected in B6-CR mice. Interestingly, administration of TUDCA has been shown to decrease hepatic steatosis and improve insulin resistance in genetically obese mice (Kars et al., 2010; Ozcan et al., 2006), suggesting a potential protective role. These results reveal differences in BA profiles linked to both host genotype and gut microbial composition. They also suggest that the differential responses to prolonged HF/HS diet consumption between B6 and CAST mice could be mediated at least in part by differences in microbial BA metabolism.

#### Gut Microbiota Influences Insulin Secretion

The most dramatic phenotype difference we observed between B6<sub>B6</sub> and B6<sub>CAST</sub> mice was in insulin secretion, where B6<sub>CAST</sub>



**Figure 6. CAST and B6 Microbiomes Differentially Regulate Insulin Secretion and *Fxr* Expression in Pancreatic Islets**

(A) Total islet insulin content and glucose-stimulated insulin secretion in response to low glucose (3.3 mM), low glucose plus KCl (40 mM), high glucose (16.7 mM), and high glucose plus GLP-1 (100 nM) from islets isolated from B6<sub>B6</sub> and B6<sub>CAST</sub> mice. The number of islets and the insulin content per islet were not different between the groups.

(B) Relative expression of *Fxr* mRNA from isolated islets. Figure S6 shows microbiota composition for donor and transplanted communities.

\* $p < 0.05$  by Student's *t* test. Data are mean  $\pm$  SEM;  $n = 5$ .

mice had a blunted insulin response during the oGTT (Figure 3E). This attenuated response in B6<sub>CAST</sub> mice may also reflect low insulin secretion from  $\beta$  cells and/or increased insulin clearance. To determine whether the differential insulin response during the oGTT in the B6<sub>B6</sub> versus B6<sub>CAST</sub> mice resulted from altered insulin secretion, we performed *ex vivo* insulin secretion assays on isolated islets. Islets were harvested from B6-GF mice 1 month after successful colonization with either CAST-CR or B6-CR cecum-derived microbiota (Figure S5).

The isolated islets partially recapitulated the reduced insulin secretion observed in the CAST-colonized mice *in vivo* (Figure 3E). The comparison between the B6-GF mice receiving B6 versus CAST microbiota allowed us to estimate the contribution of the microbiota to the strain difference in insulin secretion (Figure 6A). Accordingly, the reduction in insulin secretion caused by CAST microbiota colonization in B6 mice was  $\sim 33\%$ .

Circulating acetate is capable of modulating insulin secretion from pancreatic islets. Specifically, recent studies have shown that acetate directly enhances glucose-stimulated insulin secretion through activation of free fatty acid receptors on  $\beta$  cells (Priyadarshini et al., 2015) and the parasympathetic nervous system (Perry et al., 2016). Therefore, we measured concentrations of SCFAs in plasma and cecum but found no differences in levels of acetate between B6<sub>B6</sub> and B6<sub>CAST</sub> mice (Figures S6A and S6B), suggesting that the divergent effects of the B6 and CAST microbiota on insulin secretion are unlikely to stem from differences in acetate.

Recent *in vitro* studies have also identified BAs as important regulators of islet function (Düfer et al., 2012; Renga et al., 2010). We investigated the plasma BA profiles in the B6<sub>B6</sub> and B6<sub>CAST</sub> mice used for insulin secretion studies (Figures S6C and S6D). B6<sub>CAST</sub> BA profiles were composed of a significantly higher percentage of hydrophilic BAs (Figure S6C). Consistent with a previous report (Sayin et al., 2013), BA profiles were dominated by taurine-conjugated species, with T $\omega$ MCA and T $\beta$ MCA being the two most abundant in both groups of animals (Figure S6D). In B6<sub>B6</sub> mice, the hydrophobic secondary BAs DCA and LCA were significantly higher than in B6<sub>CAST</sub> mice (Figure S6D).

BAs regulate insulin secretion through the activation of specific receptors in islets. For instance, BAs can directly increase insulin secretion and production through activation of farnesoid X receptor (*Fxr*) in  $\beta$  cells (Düfer et al., 2012; Renga et al., 2010). Expression of *Fxr* is increased in an agonist-dependent manner (Lee et al., 2006). Remarkably, we found that expression of *Fxr* was significantly higher in B6<sub>B6</sub> islets compared with B6<sub>CAST</sub> islets (Figure 6B). These results suggest that the gut microbiota modulate BA-dependent signaling in pancreatic islets.

## DISCUSSION

The collective genetic variance of the eight CC strains is roughly equivalent to that of the entire human population, with the three wild-derived strains (WSB, CAST, and PWK) accounting for  $\sim 75\%$  of the genetic diversity within the cohort (Roberts et al., 2007). Remarkably, these three wild-derived strains captured the full scope of dietary responsiveness observed across the panel (Figures 1 and S1). HF/HS feeding had no effect on any of the phenotypes measured in CAST mice, whereas it resulted in weight gain, glucose intolerance, and insulin resistance in B6 mice. Additionally, the diet caused a simultaneous increase in weight and glucose in NZO mice. We also identified significant differences in the gut microbiota composition among strains and between diets. All animals were obtained from the same facility and subject to the same environmental conditions throughout the study, and genetic differences among the mice are the only known variable. Together, these results support a role for host genetics to regulate the composition of the microbiota. However, it is important to note that, although large population studies have identified highly heritable taxa, the genetic architecture underlying these taxa is highly complex with relatively small effect sizes that are difficult to replicate (Benson, 2016).

From the CC founder panel, we identified B6 and CAST as the two strains with the most divergent phenotypes. Previous studies have exploited the differential response to diet-induced metabolic disease between B6 and CAST to identify genetic loci associated with metabolic disease (Mehraban et al., 1998,

2000). In these studies, the gut microbiome may have contributed to the metabolic differences between strains.

In order to dissect the contribution of the microbiome of B6 and CAST to their contrasting metabolic profiles, we resorted to fecal transplantation experiments. B6-GF mice colonized with the CAST microbiota were less affected by chronic HF/HS feeding relative to B6-GF mice colonized with the B6 microbiota. The mice receiving the CAST microbiota secreted far less insulin in response to a glucose challenge but were still able to maintain normal blood glucose levels.

We consistently identified microbial taxa in both the CC founders and transplant recipients associated with metabolic traits. Clostridiaceae showed the strongest positive correlation with plasma insulin levels and weight gain (Figure 2A). Clostridiaceae also had a strong positive correlation with AUC<sub>insulin</sub>, a proxy for pancreas function. OTUs within the Clostridiaceae family have previously been both positively and negatively associated with metabolic traits (Ussar et al., 2015; Karlsson et al., 2013), and a recent study showed a positive correlation between an increase in BMI and an increase of SCFA-producing Clostridia species in Danish infants (Bergström et al., 2014). In contrast to the elevated Clostridiaceae in mice with a B6 microbiota, Bacteroidaceae was significantly higher in CAST-CR and B6<sub>CAST</sub> mice (Figures S3C and 4C). Bacteroidaceae was negatively correlated with body weight, circulating insulin, and AUC<sub>insulin</sub> (Figure 2A). A previous report found that daily oral administration of *Bacteroides uniformis*, a member of the Bacteroidaceae family, ameliorated metabolic dysfunction resulting from a high-fat diet (Gauffin Cano et al., 2012). This species also evoked a reduction in hepatic triglyceride levels, consistent with our observations that B6<sub>CAST</sub> mice have lower hepatic lipid levels compared to B6<sub>B6</sub> mice. Fecal abundance of members of the Bacteroidaceae family, including *Bacteroides vulgatus*, has also been reported to be lower in humans with T2D (Wu et al., 2010). Despite the high abundance of Bacteroidaceae in B6<sub>CAST</sub> mice, we did not observe complete protection from diet-induced metabolic disease that we observed in CAST-CR mice, suggesting that host factors, or taxa that failed to colonize transplanted mice (e.g., Verrucomicrobiaceae), contribute to the metabolotype differences.

Vitamin B12 is exclusively produced by microbes (Martens et al., 2002), and several members of the Bacteroidaceae family transport, metabolize, and produce vitamin B12 analogs (Goodman et al., 2009; Degnan et al., 2014; Wu et al., 2015). Metagenomic analysis of the microbial communities from mice with the CAST microbiota revealed microbial functional enrichment for pathways involved in the biosynthesis of vitamin B12 (Figure S4A), which is necessary for DNA synthesis, neurological function, hematopoiesis, epigenetic modifications, and propionate metabolism (Kibirige and Mwebaze, 2013). Importantly, deficiencies in vitamin B12 are commonly observed in individuals with T2D and gestational diabetes (Kibirige and Mwebaze, 2013; Krishnaveni et al., 2009), and B12 therapy improves insulin resistance and endothelial function in patients with metabolic syndrome by mechanisms that are not fully elucidated (Setola et al., 2004).

Our metagenomic analysis also revealed that genes involved in LPS production are enriched in the CAST-transplanted microbiome (Figures 4E and S4B). This finding was surprising given that increased levels of LPS have been causally linked to the

development of metabolic disease, yet B6<sub>CAST</sub> mice are partially protected from the effects of HF/HS feeding relative to B6<sub>B6</sub> animals (Figure 3). Taxonomic evaluation of the metagenomic data indicated that the Bacteroidetes phylum is the major contributor to the increased abundance of genes from this pathway (Table S4). This is relevant because unrelated bacteria generate structurally distinct LPS molecules with varying capacity to elicit an innate immune response (Whitfield and Trent, 2014). Notably, a recent study showed that LPS derived from *E. coli* generates a strong inflammatory signal, whereas LPS derived from members of the Bacteroidetes phylum inhibited the host immune response (Vatanen et al., 2016). The differential ability of LPS sub-types to modify host physiology may explain why LPS has been shown to both stimulate (Nguyen et al., 2014) and attenuate (Amyot et al., 2012) insulin secretion. Studies aimed at testing the roles of LPS derived from phylogenetically diverse taxa on metabolic disease and insulin secretion are warranted to further clarify how structural differences in this molecule affect host metabolism.

In addition to LPS, gut microbes produce SCFAs, which are important energy and signaling molecules implicated in metabolic disease. For instance, butyrate has been shown to improve whole-body insulin sensitivity (Gao et al., 2009) and patients with T2D have reduced levels of butyrate-producing bacteria (Qin et al., 2010). SCFAs are also elevated in individuals with diet-induced obesity, which is consistent with the elevated cecal SCFA levels in B6<sub>B6</sub> mice (Turnbaugh et al., 2008). Interestingly, SCFAs are also known regulators of insulin sensitivity and secretion. Acetate can modulate insulin secretion from  $\beta$  cells either directly through FFAR2 or via parasympathetic activation (Priyadarshini et al., 2015; Perry et al., 2016). However, we did not observe differences in concentrations of plasma or cecal acetate in the transplanted animals (Figures 4F, S6A, and S6B). Therefore, it is unlikely that the differences in insulin secretion could be attributed to SCFAs and consequentially implies there are multiple pathways through which the gut microbiota can modulate insulin secretion from  $\beta$  cells.

Gut microbes are responsible for the production of the highly hydrophobic secondary BAs DCA and LCA through the dehydroxylation of the primary BAs, CA and CDCA, in the colon. Removal of glycine/taurine BA conjugates via BSH enzymes is a prerequisite for  $7\alpha/\beta$ -dehydroxylation of primary BAs into secondary BAs (Batta et al., 1990). Interestingly, there were 13 predicted BSH genes that were more abundant in the B6 metagenome but only two in the CAST metagenome. One possible interpretation of this result is that there may be more bacterial species present in the B6 microbiome that are able to deconjugate BA. Consistent with this, B6<sub>B6</sub> mice had significantly higher levels of secondary BA as well as hydrophobic BA species than B6<sub>CAST</sub> mice (Figures 5B, 5C, S6C, and S6D), both of which are elevated in humans and mice with insulin resistance (Ryan et al., 2014; Prawitt et al., 2011). Furthermore, DCA has been positively associated with higher levels of Firmicutes (Islam et al., 2011). This is consistent with our findings as B6-CR founders and B6<sub>B6</sub> had a significantly greater relative abundance of Firmicutes and fecal DCA than CAST-CR and B6<sub>CAST</sub> (Figures S3B and 5B). Conversely, B6<sub>CAST</sub> had a higher abundance of hydrophilic BAs and the majority of the BA pool was comprised of the mouse primary BA,  $\beta$ MCA (Figures 5B and 5C).

The BA receptor FXR is expressed in pancreatic  $\beta$  cells, and its activation via BAs enhances insulin secretion (Kumar et al., 2012; Renga et al., 2010). Hydrophobic BAs, such as CDCA, DCA, LCA, and their taurine conjugates, are known ligands of FXR. The hydrophobic tauro-chenodeoxycholic acid (TCDCA) increases insulin production and secretion through an FXR-dependent regulation of  $K_{ATP}$  channels (Düfer et al., 2012). Moreover,  $\beta$  cell FXR activation in diabetic-leptin-receptor-deficient (db/db) mice and NOD mice increases insulin secretion and delays the development of diabetes (Renga et al., 2010; Zhang et al., 2006). We detected higher levels of LCA and DCA in the feces and plasma of B6<sub>B6</sub> mice relative to B6<sub>CAST</sub> mice (Figures 5B, S6C, and S6D), along with increased expression of *Fxr* in pancreatic islets from B6<sub>B6</sub> mice (Figure 6B). Altogether, this suggests that the gut microbiota and BA composition could modulate pancreatic function and insulin secretion.

We have highlighted four examples of microbial-derived products, vitamin B12, SCFAs, LPS, and BAs, as plausible mediators of the microbiome effect on insulin secretion. However, there are thousands of other metabolites that were not characterized in our study and could also play an important role in regulating host metabolism. Future experiments using gnotobiotic mice colonized with defined communities that have different metabolic capabilities will provide mechanistic insights into the communication between gut microbes and the host.

## EXPERIMENTAL PROCEDURES

### Mouse Husbandry

Animal care and study protocols were approved by the University of Wisconsin-Madison Animal Care and Use Committee.

### CC Mouse Husbandry

Mice were housed on a 12-hr light:dark cycle. CC founder strains were obtained from The Jackson Laboratory and were bred at University of Wisconsin, Madison. Mice were group housed by strain (two mice/cage) and diet under temperature- and humidity-controlled conditions and received ad libitum access to water and food. After 4 weeks of age, mice were maintained on either a control (TD.08810; Envigo Teklad; 16.8% kcal fat; 60.9% carbohydrate; 22.3% protein) or a high-fat, high-sucrose diet (TD.08811; Envigo Teklad; 44.6% kcal fat; 40.6% carbohydrate; 14.8% protein; Table S1). Strains were housed within the same vivarium throughout the duration of the study.

### Gnotobiotic Mouse Husbandry

C57BL/6J germ-free mice were bred and housed in the Microbial Sciences Building vivarium at University of Wisconsin-Madison to generate mice used in this study. B6-CR and B6-GF mice were housed in separate plastic flexible vinyl gnotobiotic isolators under temperature- and humidity-controlled conditions (12-hr light:dark). Fresh cecal contents were collected from 15-week-old conventional B6-CR and CAST-CR mice maintained on the HF/HS diet ( $n = 2-3$  mice per donor cecal microbiota samples). Cecal contents from B6 and CAST donor mice were resuspended in rich medium (1:100 w/vol) inside an anaerobic chamber. Suspensions were transferred into anaerobic sealed tubes and moved into gnotobiotic isolators. Nine-week-old B6-GF male mice were inoculated via a single oral gavage with  $\sim 0.2$  mL of cecal inocula (Turnbaugh et al., 2009). Each group of mice was housed in a controlled environment in separate plastic flexible vinyl gnotobiotic isolators under standard conditions. Recipient mice received sterilized water and HF/HS diet (TD.0.8811) ad libitum beginning 1 week before colonization.

### Metabolic Phenotypes

Mice were fasted for 4 hr before blood collection. Fasting levels of glucose, insulin, and triglycerides were quantified at regular intervals, along with body weight. Glucose tolerance tests were performed by administering an oral dose of glucose (2 g/kg body weight) after a 4-hr fast. Blood was collected

at 0, 5, 15, 30, 60, and 120 min to assess glucose and insulin levels. Hepatic triglycerides were extracted following the Bligh and Dyer method.

### Microbiome Sequencing and Analysis

DNA was isolated from cecal contents and feces by extraction using a bead-beating protocol (Turnbaugh et al., 2009). The V4 region of the 16S rRNA gene was amplified using barcoded primers, and sequencing was performed using the Illumina MiSeq platform. 16S rRNA sequences were analyzed using QIIME (Quantitative Insights Into Microbial Ecology) software package (Caporaso et al., 2010). LEfSe analysis was performed using standard parameters ( $p < 0.05$  and LDA score 2.0; Segata et al., 2011). For the metagenomic analysis, DNA fragments  $\sim 350-400$  bp were sequenced using Illumina Rapid HiSeq 2000. Details about metagenomic sequence analysis, 16S rRNA sequencing, and analyses are provided in Supplemental Experimental Procedures.

### Statistical Analysis

The data are expressed as mean  $\pm$  SEM and analyzed using GraphPad Prism 6.0 (GraphPad Software). Multiple groups were analyzed by one-way or two-way ANOVA followed by Bonferroni's multiple comparisons test. Significant differences between two groups were evaluated by two-tailed unpaired Student's *t* test or Mann-Whitney U test for samples that were not normally distributed. Pearson's correlations between microbiota and phenotypes and association testing were performed in R. The level of significance was set at  $p < 0.05$ ; \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and \*\*\*\* $p < 0.0001$ .

See Supplemental Information for detailed description of methods related to (1) plasma and liver measurements, (2) high-performance liquid chromatography-mass spectrometry (HPLC-MS) and GC-MS assays of plasma and cecal contents, (3) microbiome analyses, (4) islet isolation and GSIS, and (5) qRT-PCR assays.

### ACCESSION NUMBERS

The accession number for the data reported in this paper is ENA: PRJEB15120.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and five tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.01.062>.

### AUTHOR CONTRIBUTIONS

F.E.R., A.D.A., M.P.K., and J.H.K. conceived the project; J.H.K., E.I.V., M.E.R., D.S.S., K.L.S., B.S.Y., A.T.B., and W.Z. performed experiments and interpreted results; J.H.K., M.P.K., and L.L.T. prepared figures; J.H.K., M.P.K., A.D.A., and F.E.R. wrote the manuscript; and W.Z., L.L.T., M.E.R., and B.H. provided critical feedback. All authors read and agreed on the final version of the manuscript.

### ACKNOWLEDGMENTS

The authors thank the University of Wisconsin Biotechnology Center DNA Sequencing Facility for providing sequencing and support services and the University of Wisconsin Center for High Throughput Computing (CHTC) in the Department of Computer Sciences for providing computational resources, support, and assistance. We also thank the University of Wisconsin Mass Spectrometry Facility for technical assistance. Finally, we are grateful to Dr. Jeffrey Gordon for his support on this project. This work was supported in part by grants NIH DK108259 (to F.E.R.), DK101573, DK102948, R24DK091207 (to A.D.A.), RR021940 (to B.H.), and GM077336 (to B.H.) and by the Clinical and Translational Science Award (CTSA) program, through the NIH National Center for Advancing Translational Sciences (NCATS) grants UL1TR000427 and KL2TR000428 (to F.E.R.). L.L.T. is supported by a NLM Computation and Informatics in Biology and Medicine Postdoctoral

Fellowship (5T15LM007359). J.H.K. is supported by the NIH National Institute of Allergy and Infectious Diseases grant T32AI55397.

Received: August 22, 2016

Revised: December 9, 2016

Accepted: January 24, 2017

Published: February 14, 2017

## REFERENCES

- Amyot, J., Semache, M., Ferdaoussi, M., Fontés, G., and Poitout, V. (2012). Lipopolysaccharides impair insulin gene expression in isolated islets of Langerhans via Toll-like receptor-4 and NF- $\kappa$ B signalling. *PLoS One* **7**, e36200.
- Aylor, D.L., Valdar, W., Foulds-Mathes, W., Buus, R.J., Verdugo, R.A., Baric, R.S., Ferris, M.T., Frelinger, J.A., Heise, M., Frieman, M.B., et al. (2011). Genetic analysis of complex traits in the emerging Collaborative Cross. *Genome Res.* **21**, 1213–1222.
- Batta, A.K., Salen, G., Arora, R., Shefer, S., Batta, M., and Person, A. (1990). Side chain conjugation prevents bacterial 7-dehydroxylation of bile acids. *J. Biol. Chem.* **265**, 10925–10928.
- Benson, A.K. (2016). The gut microbiome—an emerging complex trait. *Nat. Genet.* **48**, 1301–1302.
- Bergström, A., Skov, T.H., Bahl, M.I., Roager, H.M., Christensen, L.B., Ejlerkov, K.T., Mølgaard, C., Michaelsen, K.F., and Licht, T.R. (2014). Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants. *Appl. Environ. Microbiol.* **80**, 2889–2900.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336.
- Carmody, R.N., Gerber, G.K., Luevano, J.M., Jr., Gatti, D.M., Somes, L., Svensson, K.L., and Turnbaugh, P.J. (2015). Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host Microbe* **17**, 72–84.
- Churchill, G.A., Airey, D.C., Allayee, H., Angel, J.M., Attie, A.D., Beatty, J., Beavis, W.D., Belknap, J.K., Bennett, B., Berrettini, W., et al.; Complex Trait Consortium (2004). The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* **36**, 1133–1137.
- Clemente, J.C., Ursell, L.K., Parfrey, L.W., and Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell* **148**, 1258–1270.
- Degnan, P.H., Taga, M.E., and Goodman, A.L. (2014). Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab.* **20**, 769–778.
- den Besten, G., van Eunen, K., Groen, A.K., Venema, K., Reijngoud, D.J., and Bakker, B.M. (2013). The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J. Lipid Res.* **54**, 2325–2340.
- Deutscher, J., Francke, C., and Postma, P.W. (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 939–1031.
- Düfer, M., Hörth, K., Wagner, R., Schittenhelm, B., Prowald, S., Wagner, T.F., Oberwinkler, J., Lukowski, R., Gonzalez, F.J., Krippel-Dreus, P., and Dreus, G. (2012). Bile acids acutely stimulate insulin secretion of mouse  $\beta$ -cells via farnesoid X receptor activation and K(ATP) channel inhibition. *Diabetes* **61**, 1479–1489.
- Gao, Z., Yin, J., Zhang, J., Ward, R.E., Martin, R.J., Lefevre, M., Cefalu, W.T., and Ye, J. (2009). Butyrate improves insulin sensitivity and increases energy expenditure in mice. *Diabetes* **58**, 1509–1517.
- Gauffin Cano, P., Santacruz, A., Moya, Á., and Sanz, Y. (2012). *Bacteroides uniformis* CECT 7771 ameliorates metabolic and immunological dysfunction in mice with high-fat-diet induced obesity. *PLoS One* **7**, e41079.
- Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* **6**, 279–289.
- Handelsman, Y. (2011). Role of bile acid sequestrants in the treatment of type 2 diabetes. *Diabetes Care* **34**, S244–S250.
- Hartstra, A.V., Bouter, K.E., Bäckhed, F., and Nieuwdorp, M. (2015). Insights into the role of the microbiome in obesity and type 2 diabetes. *Diabetes Care* **38**, 159–165.
- Hildebrandt, M.A., Hoffmann, C., Sherrill-Mix, S.A., Keilbaugh, S.A., Hamady, M., Chen, Y.Y., Knight, R., Ahima, R.S., Bushman, F., and Wu, G.D. (2009). High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology* **137**, 1716–1724.
- Houten, S.M., Watanabe, M., and Auwerx, J. (2006). Endocrine functions of bile acids. *EMBO J.* **25**, 1419–1425.
- Islam, K.B.M.S., Fukiya, S., Hagio, M., Fujii, N., Ishizuka, S., Ooka, T., Ogura, Y., Hayashi, T., and Yokota, A. (2011). Bile acid is a host factor that regulates the composition of the cecal microbiota in rats. *Gastroenterology* **141**, 1773–1781.
- Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C.J., Fagerberg, B., Nielsen, J., and Bäckhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103.
- Kars, M., Yang, L., Gregor, M.F., Mohammed, B.S., Pietka, T.A., Finck, B.N., Patterson, B.W., Horton, J.D., Mittendorfer, B., Hotamisligil, G.S., and Klein, S. (2010). Tauroursodeoxycholic acid may improve liver and muscle but not adipose tissue insulin sensitivity in obese men and women. *Diabetes* **59**, 1899–1905.
- Kawamata, Y., Fujii, R., Hosoya, M., Harada, M., Yoshida, H., Miwa, M., Fukusumi, S., Habata, Y., Itoh, T., Shintani, Y., et al. (2003). A G protein-coupled receptor responsive to bile acids. *J. Biol. Chem.* **278**, 9435–9440.
- Kibirige, D., and Mwebaze, R. (2013). Vitamin B12 deficiency among patients with diabetes mellitus: is routine screening and supplementation justified? *J. Diabetes Metab. Disord.* **12**, 17.
- Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* **165**, 1332–1345.
- Krishnaveni, G.V., Hill, J.C., Veena, S.R., Bhat, D.S., Wills, A.K., Karat, C.L., Yajnik, C.S., and Fall, C.H. (2009). Low plasma vitamin B12 in pregnancy is associated with gestational ‘diabesity’ and later diabetes. *Diabetologia* **52**, 2350–2358.
- Kuipers, F., Bloks, V.W., and Groen, A.K. (2014). Beyond intestinal soap—bile acids in metabolic control. *Nat. Rev. Endocrinol.* **10**, 488–498.
- Kumar, D.P., Rajagopal, S., Mahavadi, S., Mirshahi, F., Grider, J.R., Murthy, K.S., and Sanyal, A.J. (2012). Activation of transmembrane bile acid receptor TGR5 stimulates insulin secretion in pancreatic  $\beta$  cells. *Biochem. Biophys. Res. Commun.* **427**, 600–605.
- Lee, F.Y., Lee, H., Hubbert, M.L., Edwards, P.A., and Zhang, Y. (2006). FXR, a multipurpose nuclear receptor. *Trends Biochem. Sci.* **31**, 572–580.
- Ley, R.E., Bäckhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D., and Gordon, J.I. (2005). Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. USA* **102**, 11070–11075.
- Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228–8235.
- Makishima, M., Okamoto, A.Y., Repa, J.J., Tu, H., Learned, R.M., Luk, A., Hull, M.V., Lustig, K.D., Mangelsdorf, D.J., and Shan, B. (1999). Identification of a nuclear receptor for bile acids. *Science* **284**, 1362–1365.
- Martens, J.H., Barg, H., Warren, M.J., and Jahn, D. (2002). Microbial production of vitamin B12. *Appl. Microbiol. Biotechnol.* **58**, 275–285.
- Mehrabian, M., Wen, P.Z., Fisler, J., Davis, R.C., and Lusk, A.J. (1998). Genetic loci controlling body fat, lipoprotein metabolism, and insulin levels in a multifactorial mouse model. *J. Clin. Invest.* **101**, 2485–2496.
- Mehrabian, M., Castellani, L.W., Wen, P.Z., Wong, J., Rithaporn, T., Hama, S.Y., Hough, G.P., Johnson, D., Albers, J.J., Mottino, G.A., et al. (2000). Genetic control of HDL levels and composition in an interspecific mouse cross (CAST/Ei  $\times$  C57BL/6J). *J. Lipid Res.* **41**, 1936–1946.

- Nguyen, A.T., Mandard, S., Dray, C., Deckert, V., Valet, P., Besnard, P., Drucker, D.J., Lagrost, L., and Grober, J. (2014). Lipopolysaccharides-mediated increase in glucose-stimulated insulin secretion: involvement of the GLP-1 pathway. *Diabetes* 63, 471–482.
- O'Connor, A., Quizon, P.M., Albright, J.E., Lin, F.T., and Bennett, B.J. (2014). Responsiveness of cardiometabolic-related microbiota to diet is influenced by host genetics. *Mamm. Genome* 25, 583–599.
- Ozcan, U., Yilmaz, E., Ozcan, L., Furuhashi, M., Vaillancourt, E., Smith, R.O., Görgün, C.Z., and Hotamisligil, G.S. (2006). Chemical chaperones reduce ER stress and restore glucose homeostasis in a mouse model of type 2 diabetes. *Science* 313, 1137–1140.
- Parks, D.J., Blanchard, S.G., Bledsoe, R.K., Chandra, G., Consler, T.G., Kliewer, S.A., Stimmel, J.B., Willson, T.M., Zavacki, A.M., Moore, D.D., and Lehmann, J.M. (1999). Bile acids: natural ligands for an orphan nuclear receptor. *Science* 284, 1365–1368.
- Parks, B.W., Nam, E., Org, E., Kostem, E., Norheim, F., Hui, S.T., Pan, C., Civelek, M., Rau, C.D., Bennett, B.J., et al. (2013). Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metab.* 17, 141–152.
- Perry, R.J., Peng, L., Barry, N.A., Cline, G.W., Zhang, D., Cardone, R.L., Petersen, K.F., Kibbey, R.G., Goodman, A.L., and Shulman, G.I. (2016). Acetate mediates a microbiome-brain- $\beta$ -cell axis to promote metabolic syndrome. *Nature* 534, 213–217.
- Prawitt, J., Caron, S., and Staels, B. (2011). Bile acid metabolism and the pathogenesis of type 2 diabetes. *Curr. Diab. Rep.* 11, 160–166.
- Priyadarshini, M., Villa, S.R., Fuller, M., Wicksteed, B., Mackay, C.R., Alquier, T., Poutout, V., Mancebo, H., Mirmira, R.G., Gilchrist, A., and Layden, B.T. (2015). An acetate-specific GPCR, FFAR2, regulates insulin secretion. *Mol. Endocrinol.* 29, 1055–1066.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al.; MetaHIT Consortium (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65.
- Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60.
- Renga, B., Mencarelli, A., Vavassori, P., Brancaleone, V., and Fiorucci, S. (2010). The bile acid sensor FXR regulates insulin transcription and secretion. *Biochim. Biophys. Acta* 1802, 363–372.
- 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.
- Ridlon, J.M., Kang, D.-J., and Hylemon, P.B. (2006). Bile salt biotransformations by human intestinal bacteria. *J. Lipid Res.* 47, 241–259.
- Roberts, A., Pardo-Manuel de Villena, F., Wang, W., McMillan, L., and Threadgill, D.W. (2007). The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: implications for QTL discovery and systems genetics. *Mamm. Genome* 18, 473–481.
- Russell, D.W. (2009). Fifty years of advances in bile acid synthesis and metabolism. *J. Lipid Res.* 50, S120–S125.
- Ryan, K.K., Tremaroli, V., Clemmensen, C., Kovatcheva-Datchary, P., Myronovich, A., Karns, R., Wilson-Pérez, H.E., Sandoval, D.A., Kohli, R., Bäckhed, F., and Seeley, R.J. (2014). FXR is a molecular target for the effects of vertical sleeve gastrectomy. *Nature* 509, 183–188.
- Sayin, S.I., Wahlström, A., Felin, J., Jäntti, S., Marschall, H.U., Bamberg, K., Angelin, B., Hyötyläinen, T., Orešič, M., and Bäckhed, F. (2013). Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell Metab.* 17, 225–235.
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, R60.
- Setola, E., Monti, L.D., Galluccio, E., Palloschi, A., Fragasso, G., Paroni, R., Magni, F., Sandoli, E.P., Lucotti, P., Costa, S., et al. (2004). Insulin resistance and endothelial function are improved after folate and vitamin B12 therapy in patients with metabolic syndrome: relationship between homocysteine levels and hyperinsulinemia. *Eur. J. Endocrinol.* 151, 483–489.
- Sommer, F., and Bäckhed, F. (2013). The gut microbiota—masters of host development and physiology. *Nat. Rev. Microbiol.* 11, 227–238.
- 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.
- Turnbaugh, P.J., Bäckhed, F., Fulton, L., and Gordon, J.I. (2008). Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* 3, 213–223.
- Turnbaugh, P.J., Hamady, M., Yatsunenkov, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., et al. (2009). A core gut microbiome in obese and lean twins. *Nature* 457, 480–484.
- Ussar, S., Griffin, N.W., Bezy, O., Fujisaka, S., Vienberg, S., Softic, S., Deng, L., Bry, L., Gordon, J.I., and Kahn, C.R. (2015). Interactions between gut microbiota, host genetics and diet modulate the predisposition to obesity and metabolic syndrome. *Cell Metab.* 22, 516–530.
- Vatanen, T., Kostic, A.D., d'Hennezel, E., Siljander, H., Franzosa, E.A., Yassour, M., Kolde, R., Vlamakis, H., Arthur, T.D., Hämäläinen, A.M., et al.; DIABIMMUNE Study Group (2016). Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. *Cell* 165, 842–853.
- Vrieze, A., Holleman, F., Zoetendal, E.G., de Vos, W.M., Hoekstra, J.B., and Nieuwdorp, M. (2010). The environment within: how gut microbiota may influence metabolism and body composition. *Diabetologia* 53, 606–613.
- Whitfield, C., and Trent, M.S. (2014). Biosynthesis and export of bacterial lipopolysaccharides. *Annu. Rev. Biochem.* 83, 99–128.
- Wu, X., Ma, C., Han, L., Nawaz, M., Gao, F., Zhang, X., Yu, P., Zhao, C., Li, L., Zhou, A., et al. (2010). Molecular characterisation of the faecal microbiota in patients with type II diabetes. *Curr. Microbiol.* 61, 69–78.
- Wu, M., McNulty, N.P., Rodionov, D.A., Khoroshkin, M.S., Griffin, N.W., Cheng, J., Latreille, P., Kerstetter, R.A., Terrapon, N., Henrissat, B., et al. (2015). Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides. *Science* 350, aac5992.
- Zhang, Y., Lee, F.Y., Barrera, G., Lee, H., Vales, C., Gonzalez, F.J., Willson, T.M., and Edwards, P.A. (2006). Activation of the nuclear receptor FXR improves hyperglycemia and hyperlipidemia in diabetic mice. *Proc. Natl. Acad. Sci. USA* 103, 1006–1011.

# Clonal Expansion of Lgr5-Positive Cells from Mammalian Cochlea and High-Purity Generation of Sensory Hair Cells

Will J. McLean,<sup>1,2,3,10,11</sup> Xiaolei Yin,<sup>4,5,6,7,8,10</sup> Lin Lu,<sup>4,5</sup> Danielle R. Lenz,<sup>1,2</sup> Dalton McLean,<sup>1,2</sup> Robert Langer,<sup>4,8,9,\*</sup> Jeffrey M. Karp,<sup>5,6,7,8,12,\*</sup> and Albert S.B. Edge<sup>1,2,3,7,\*</sup>

<sup>1</sup>Department of Otolaryngology and Laryngology, Harvard Medical School, Boston, MA 02115, USA

<sup>2</sup>Eaton-Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA 02114, USA

<sup>3</sup>Program in Speech and Hearing Bioscience and Technology, Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA

<sup>4</sup>Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02142, USA

<sup>5</sup>Division of Engineering in Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA

<sup>6</sup>Department of Medicine, Harvard Medical School, Boston, MA 02115, USA

<sup>7</sup>Harvard Stem Cell Institute, Cambridge, MA 02138, USA

<sup>8</sup>Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA

<sup>9</sup>Department of Chemical Engineering, MIT, Cambridge, MA 02142, USA

<sup>10</sup>Co-first author

<sup>11</sup>Present address: Frequency Therapeutics, Woburn, MA 01801, USA

<sup>12</sup>Lead Contact

\*Correspondence: rlanger@mit.edu (R.L.), jmkarp@partners.org (J.M.K.), albert\_edge@meei.harvard.edu (A.S.B.E.)  
<http://dx.doi.org/10.1016/j.celrep.2017.01.066>

## SUMMARY

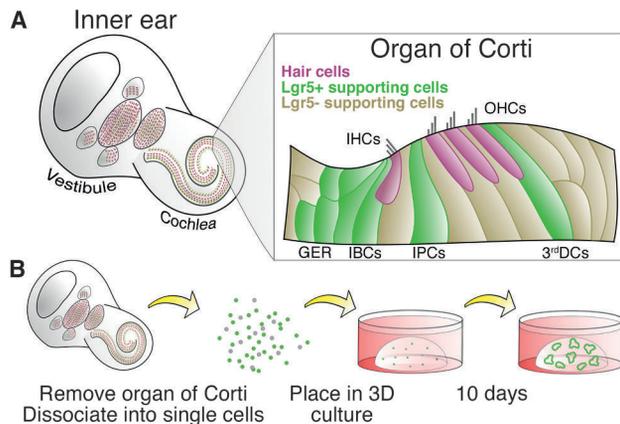
Death of cochlear hair cells, which do not regenerate, is a cause of hearing loss in a high percentage of the population. Currently, no approach exists to obtain large numbers of cochlear hair cells. Here, using a small-molecule approach, we show significant expansion (>2,000-fold) of cochlear supporting cells expressing and maintaining Lgr5, an epithelial stem cell marker, in response to stimulation of Wnt signaling by a GSK3 $\beta$  inhibitor and transcriptional activation by a histone deacetylase inhibitor. The Lgr5-expressing cells differentiate into hair cells in high yield. From a single mouse cochlea, we obtained over 11,500 hair cells, compared to less than 200 in the absence of induction. The newly generated hair cells have bundles and molecular machinery for transduction, synapse formation, and specialized hair cell activity. Targeting supporting cells capable of proliferation and cochlear hair cell replacement could lead to the discovery of hearing loss treatments.

## INTRODUCTION

Hearing impairment is a major health challenge estimated by the World Health Organization to affect over 5% of the world's population (360 million people, including 32 million children). The sensory hair cells that detect sound and transmit their signal to the brain via the auditory nerve are susceptible to damage. After

loss, the hair cells are never replaced (Cox et al., 2014; Fujioka et al., 2015), and thus, the number of cells, which is low (15,000 per ear in humans and 3,000 in mice) at the start of post-natal life, only decreases with age, and the absence of cell replacement leads to a high prevalence of acquired forms of deafness. Indeed, hair cell and auditory nerve damage, typically caused by noise exposure, ototoxic drugs, viral or bacterial infections, and aging, accounts for more than 80% of all cases of hearing loss (Davis, 1983).

Lgr5, an epithelial cell protein first discovered as a marker for intestinal stem cells and then shown to be critical for their function (Barker et al., 2007; Koo and Clevers, 2014), was recently shown to be expressed in cochlear supporting cells that surround the hair cells (Chai et al., 2012; Shi et al., 2012). These Lgr5-expressing cells could be induced to undergo limited proliferation when stimulated by Wnt in the normally post-mitotic cochlear sensory epithelium (Shi et al., 2013). Indeed, consistent with a progenitor role, supporting cells that expressed Lgr5 gave rise to new Lgr5+ cells by propagation and to hair cells that were Lgr5-, whereas supporting cells that did not express this receptor did not give rise to hair cells (Bramhall et al., 2014; Shi et al., 2012). Consistent with its role in upstream regulation of the transcription factor Atoh1 (Shi et al., 2010), which is a master regulator of hair cell differentiation (Edge and Chen, 2008; Kelley, 2006), upregulation of Wnt also increased hair cell differentiation. This combination of the ability to divide in response to Wnt signaling and the potency to differentiate into hair cells suggested that Lgr5+ cells were acting as progenitor cells of the cochlear epithelium (Shi et al., 2012). Indeed, in the newborn cochlea, Lgr5+ cells showed the capacity to regenerate spontaneously after damage (Bramhall et al., 2014; Cox et al., 2014). These data supported a role for the Lgr5+ cells as cochlear



**Figure 1. Diagram of Cochlear Lgr5+ Cell Culture**

(A) Lgr5 is expressed in a subset of supporting cells surrounding the cochlear hair cells, including the greater epithelial ridge (GER), inner border cells (IBCs), inner pillar cells (IPCs), and third Deiters cells (3<sup>rd</sup> DCs). IHC, inner hair cell; OHC, outer hair cell.

(B) Diagram of cochlear epithelial cell isolation and culture as single cells in a 3D system for 10 days.

progenitor cells, but spontaneous regeneration capacity was lost after the first postnatal week, and, indeed, no cell division or cell replacement occurs in the sensory epithelium of the adult cochlea (Bramhall et al., 2014; Cox et al., 2014; Fujioka et al., 2015; Shi et al., 2012). Supporting cell transdifferentiation can lead to some hair cell replacement (Mizutani et al., 2013), but regenerating a functional cochlea would require both stimulating these cells to divide and differentiating them to hair cells, an approach that would benefit greatly from the isolation of Lgr5+ progenitor cells to develop protocols for their expansion and differentiation to hair cells.

The limited ability of cochlear cells to regenerate is unusual compared to other epithelia, but, despite different capacities for regeneration, Lgr5+ cochlear supporting cells have several characteristics in common with Lgr5+ cells from the intestine. Activation of both the Wnt and Notch signaling pathways has been independently demonstrated to be required for the establishment of Lgr5+ cells in the cochlea during development (Shi et al., 2010, 2014; Yamamoto et al., 2011). Wnt signaling is required for hair cell differentiation (Shi et al., 2014), which is increased by concurrent inhibition of Notch (Bramhall et al., 2014; Kelley, 2006; Korrapati et al., 2013; Shi et al., 2010, 2014). This is strikingly similar to intestinal epithelia, where Wnt and Notch signaling are required for stem cell expansion, and lack of Notch signaling with active Wnt leads to differentiation to mature epithelial cell types (de Lau et al., 2011; Koo and Clevers, 2014; van Es et al., 2005; Yin et al., 2014). Some expansion of Lgr5+ cells from the cochlea could be achieved by propagation as cochlear spheres (Shi et al., 2012), but heterogeneous cell populations are obtained and the yield of hair cells upon differentiation is low. Changes in gene expression of progenitors results in a loss of sphere-forming capacity in the adult mouse cochlea (Oshima et al., 2007). Studies on progenitor cells have thus been limited by the small number of Lgr5+ cells, which comprise a subset of the few cells in the cochlear epithelium,

and inefficient protocols for their expansion. Here, by employing a cocktail of drugs and growth factors to modulate multiple pathways, we demonstrate mechanisms to clonally expand Lgr5+ cells from both newborn and normally unresponsive, adult tissue and to efficiently differentiate these colonies into nearly pure populations of hair cells in high yield. We show, furthermore, that the same drug cocktail drives the generation of hair cells and supporting cells in both healthy and damaged neonatal organ of Corti.

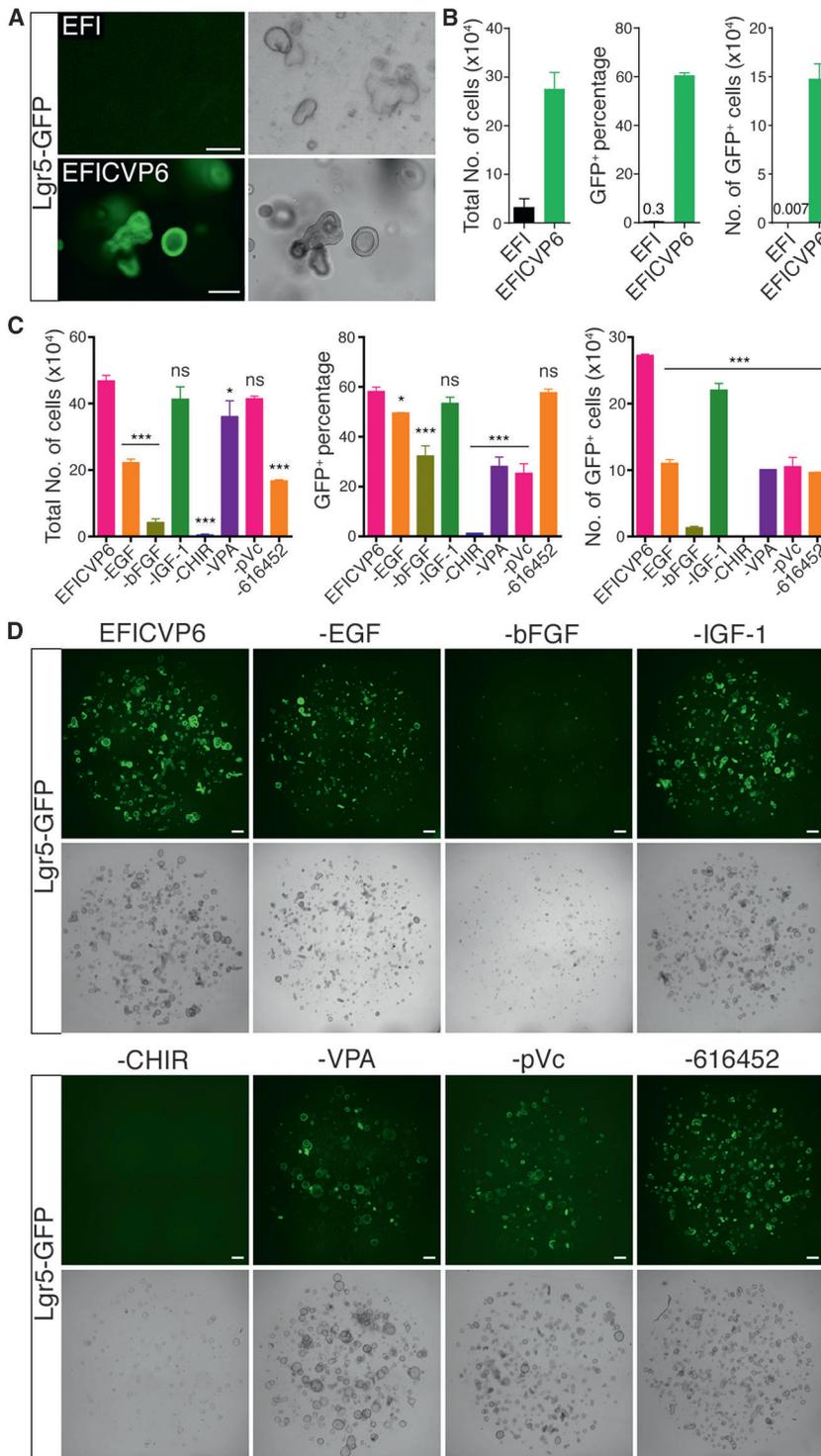
## RESULTS

### Lgr5+ Cochlear Cell Expansion

Lgr5+ cells represent a subset of supporting cells within the cochlear epithelium (Figure 1A). Using an Lgr5-GFP mouse line, we tested the activation or inhibition of multiple pathways to expand single Lgr5+ supporting cells isolated from the neonatal cochlea in a Matrigel-based 3D culture system (Figure 1B). Initially, we aimed to adapt conditions we previously developed for culture of intestinal stem cells to the inner ear progenitor cells (Yin et al., 2014). We added the glycogen synthase kinase 3 $\beta$  (GSK3 $\beta$ ) inhibitor CHIR99021 (CHIR or C) and the histone deacetylase (HDAC) inhibitor valproic acid (VPA or V) to the growth factor cocktail that was previously used for the culture of inner ear spheres, which includes epidermal growth factor (EGF), basic fibroblast growth factor (bFGF or F), and insulin like growth factor 1 (IGF-1 or I) (Li et al., 2003). In parallel, we tested conditions used for the culture of intestinal stem cells, which includes EGF (E), R-Spondin1 (R), and Noggin (N) (Sato et al., 2009). The addition of CHIR and VPA to EFI (EGF, bFGF, IGF-1) significantly increased the total number and percentage of Lgr5-GFP cells in culture, which outperformed the conditions previously used to expand intestinal Lgr5-GFP cells (ENR) (Figure S1A). The addition of CV to previously used factors led to the formation of large GFP+ colonies (Figure S1B), consistent with our previous finding in the intestine (Yin et al., 2014).

Without Wnt stimulation (EFI alone), Lgr5-GFP expression diminished, and, unlike the intestinal colonies, the number of Lgr5-GFP cells diminished after passage. Although the Lgr5-GFP cell numbers decreased after passage, further Wnt stimulation allowed Lgr5-GFP cells to be maintained in culture for extended periods of time (out to 45 days, the longest time point tested). We reasoned that other factors were needed to optimize the culture of Lgr5-GFP cells and thus performed screening to identify additional factors to enable the colonies' prolonged culture and passaging. Addition of 2-phospho-L-ascorbic acid (pVc or P), a stable form of vitamin C, increased Lgr5+ cell expansion by an additional 2- to 3-fold (Figures S2A and 2B). Addition of a transforming growth factor  $\beta$  (TGF- $\beta$ ) receptor (Alk5) inhibitor, 616452 (or 6), also increased cell expansion (by 2- to 3-fold) and was required for the passage of colonies (Figures 2C, 2D, and S2C). Collectively, the addition of small molecules (CVP6), compared to growth factors alone, increased Lgr5+ cell numbers by >2,000-fold with high consistency (Figure 2B).

To examine the relative importance of individual factors in our culture system (without passaging), we separately removed each factor from the medium and quantified cell proliferation and Lgr5 expression of inner ear epithelial cells following



**Figure 2. Characterization of Culture Conditions for Inner Ear Lgr5+ Cells**

(A) GFP fluorescence and bright-field images of Lgr5-GFP colonies obtained from inner ear epithelial cells cultured for 10 days in the presence of EGF, bFGF, IGF-1 (EFI); EFI and CHIR, VPA, pVc, 616452 (EFICVP6). Scale bars, 200  $\mu$ m. (B) Number of live cells and percentage and number of Lgr5-GFP cells from inner ear epithelia cultured for 10 days.  $n = 3$ . Error bars represent mean  $\pm$  SD. (C) Number of live cells and number and percentage of Lgr5-GFP cells from inner ear epithelia cultured for 10 days. Lgr5+ cell number and percentage were highest in cultures containing EGF, bFGF, IGF-1, CHIR, VPA, pVc, and 616452 (EFICVP6) compared to cultures from which individual factors were removed. Each condition was compared to EFICVP6.  $n = 3$ . Error bars represent mean  $\pm$  SD; \*\*\* $p < 0.001$ ; \* $p < 0.05$ ; ns, not significant ( $p > 0.05$ ). (D) GFP fluorescence and bright-field images of cultures as shown in (C). Scale bars, 400  $\mu$ m.

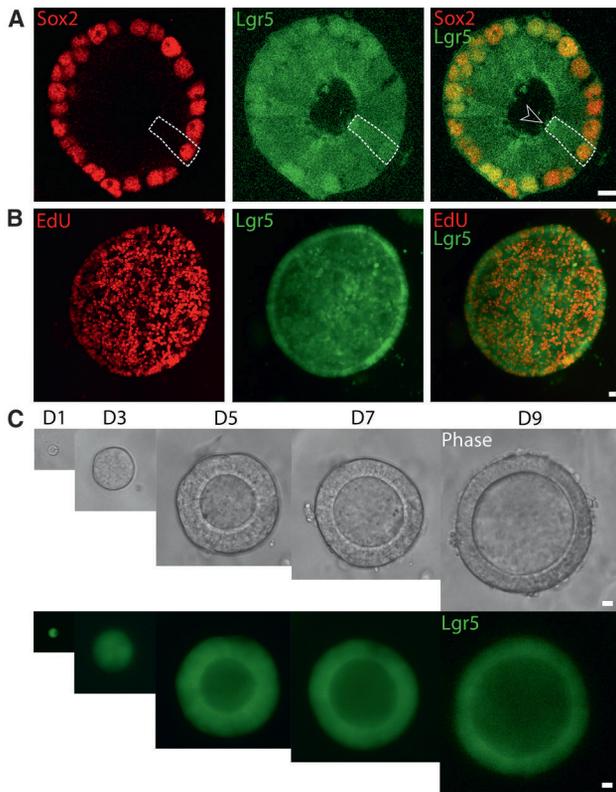
marginal beneficial effect on Lgr5 cell number and percentage. The treatment with the combined agents (EFICVP6) yielded the highest number of total cells, Lgr5+ cells, and percentage of Lgr5+ cells following 10 days of culture. These results suggest that bFGF and CHIR were most critical to Lgr5+ cell culture, while the other factors promoted maximal Lgr5 cell growth and expression. Similar results were obtained by direct visualization of GFP expression and cell growth (Figure 2D).

We further examined the potential function of individual factors. The effects of CHIR in increasing Lgr5-GFP cell number and percentage could be partially replicated with Wnt3a in combination with R-spondin1 (Figures S2D–S2G), suggesting a role of CHIR in activating the Wnt pathway. Using an Atoh1-nGFP mouse line, we found that VPA suppressed spontaneous differentiation of supporting cells into hair cells (Figure S2H), which is consistent with the role of VPA in maintaining Notch activation in intestinal stem cells (Greenblatt et al., 2007; Yin et al., 2014).

Consistent with previous reports (Chai et al., 2012; Shi et al., 2012), Lgr5+ cells expressed the supporting cell marker Sox2, and a single optical slice revealed that Lgr5-GFP colonies cultured in EFICVP6

10 days of culture (Figures 2C and 2D). Removal of CHIR or bFGF had the greatest effect on Lgr5-GFP cell number and percentage, while removal of CHIR had the greatest effect on Lgr5 expression. Removing EGF or 616452 caused a significant reduction in Lgr5-GFP cell number, while removing VPA or pVc greatly reduced Lgr5 expression. The presence of IGF-1 had a

comprised pure populations of Sox2-expressing supporting cells with nuclear localization in the basal portion of the cell (Figure 3A). Exposure of EFICVP6 cultures to ethynyldeoxyuridine (EdU) revealed that Lgr5-GFP colonies were actively proliferating (Figure 3B). Tracking of single Lgr5-GFP cells over time revealed that Lgr5-GFP colonies were formed clonally (Figure 3C).



**Figure 3. Lgr5<sup>+</sup> Supporting Cells Actively Proliferate and Form Clonal Colonies**

(A) Lgr5<sup>+</sup> colonies generated in EFICVP6 expressed the supporting cell marker Sox2 in nuclei located in the basal region of the cell. The dashed line indicates the border of a single cell with the apical surface (hollow arrowhead) facing the lumen. Image is a single optical slice. n = 4. Scale bar, 15  $\mu$ m.

(B) EdU staining of an Lgr5-GFP colony. Image is a maximum projection of a z stack. n = 6. Scale bar, 15  $\mu$ m.

(C) A single Lgr5<sup>+</sup> cell tracked over 9 days while cultured in the presence of EFICVP6. n = 7. Scale bar, 15  $\mu$ m. D, day.

### Conversion of Lgr5<sup>+</sup> Cells to Hair Cells

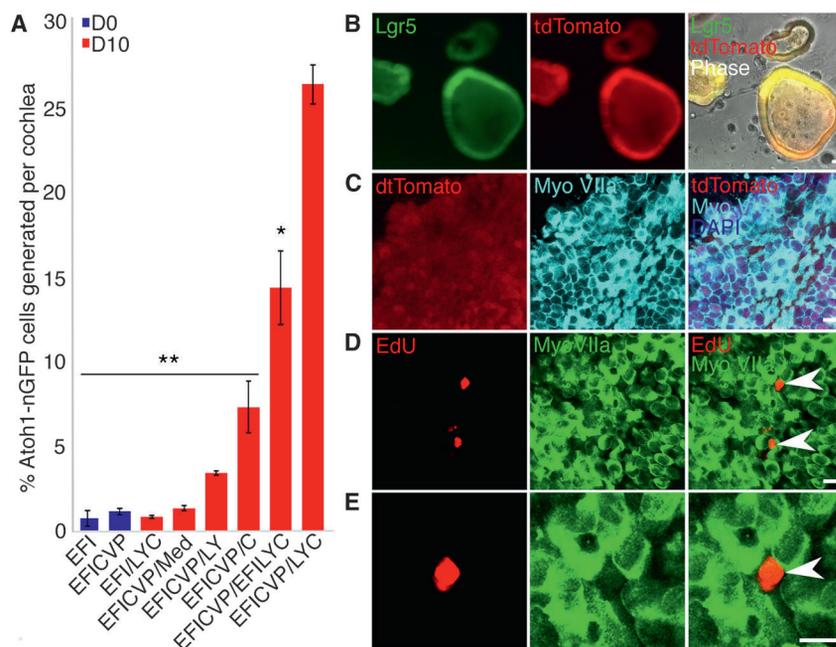
Although Notch inhibition and  $\beta$ -catenin expression were separately shown to promote hair cell differentiation in vitro from inner ear progenitor cells at a higher rate than removal of growth factors, the number of hair cells produced remained low due to the inability to expand progenitor cells and sufficiently convert them into hair cells (Jeon et al., 2011; Oshima et al., 2007; Shi et al., 2012, 2013). To test whether the expanded Lgr5<sup>+</sup> cells were able to generate higher yields of hair cells after simultaneous Notch inhibition and Wnt activation, we treated Lgr5-GFP or Atoh1-nGFP cells, expanded by the above procedures, with LY411575, a  $\gamma$ -secretase inhibitor previously used to differentiate inner ear progenitor cells (Jeon et al., 2011; Mizutani et al., 2013), and CHIR, the GSK3 $\beta$  inhibitor. Following 10 days of differentiation, the expression of Lgr5 was diminished (Figure S3A), suggesting that they were differentiated cells. Atoh1-nGFP cells were rare during the expansion phase of the culture but increased in prevalence during the differentiation step of the protocol (Figure S3B). This suggests that the combination of LY411575 and CHIR induced the differentiation of the expanded

Lgr5<sup>+</sup> cells and transformed the colonies into high-purity populations of Atoh1-nGFP hair cells (Figure S3B).

We quantified hair cell production using flow cytometry to count Atoh1-nGFP cells after expansion (day 0) and differentiation (day 10) phases of cultures originating from isolated epithelial cells from a single Atoh1-nGFP mouse cochlea (Figure 4A). We found that the addition of 616452 caused leakage of the enhancer-mediated Atoh1-nGFP (which was confirmed through co-staining with myosin VIIa during expansion). It was therefore removed from hair cell quantification assays using the Atoh1-nGFP reporter. Treatment with growth factors alone or with CHIR, VPA, and pVc produced few Atoh1-nGFP<sup>+</sup> cells (0.72% with the growth factors and 1.108% in the presence of growth factors with drugs,  $p > 0.05$ ) at the end of the 10-day expansion. Our results suggest that CHIR leads to greater differentiation than LY411575 (5.6-fold vs. 2.6-fold) when compared to removal of growth factors alone, and combining LY411575 and CHIR (LYC) further increases hair cell yield (Figure 4A). Expansion using EFICVP, followed by differentiation with LYC, produced the maximum number and purity of hair cells when multiple differentiation conditions were analyzed, suggesting that inclusion of growth factors in culture during differentiation reduces hair cell formation (Figure 4A). Expansion with EFICVP and differentiation with LYC resulted in  $\sim 26\%$  of all cells in culture expressing Atoh1-nGFP, which corresponds to  $\sim 11,600$  Atoh1-nGFP<sup>+</sup> cells per cochlea compared to the 0.72% and average of 173 cells generated per cochlea when differentiating with LYC after expanding with EFI ( $p < 0.0001$ ). Our optimal conditions generated a highly reproducible hair cell yield ( $26.3\% \pm 2.5\%$ , n = 5 independent experiments; Figure 4A) that was 67-fold greater than previous methods using only growth factors to culture supporting cells (Oshima et al., 2007) and represented a 580-fold increase of viable hair cells during culture (20 Atoh1-nGFP cells).

To determine whether Lgr5<sup>+</sup> cells were the source of hair cells, we crossed Lgr5-Cre-ER mice with Rosa26-flox-tdTomato mice and followed tdTomato expression in any new cells. When 4-hydroxytamoxifen was added to the culture at day 0 to activate the Lgr5-Cre, Lgr5<sup>+</sup> cells formed colonies in the cocktail of growth factors and drugs, in which all cells within a colony were positive for tdTomato (Figure 4B). After differentiation with LY411575 and CHIR, hair cells (marked by hair cell-specific myosin VIIa (Weil et al., 1996)) were also tdTomato-positive (Figure 4C), indicating that the myosin VIIa<sup>+</sup> cells were derived from Lgr5-expressing cells. Colonies that were tdTomato-negative did not produce myosin VIIa<sup>+</sup> cells, indicating that Lgr5-negative cells do not generate hair cells, as also demonstrated previously (Shi et al., 2012). Cultures stained for EdU given 1 day after LY411575 and CHIR administration showed that cells expressing hair cell genes did not proliferate in the differentiation conditions (Figure 4D). These data thus suggest that hair cells did not proliferate during colony formation or differentiation.

Further analyses showed that differentiating expanded supporting cells with LYC resulted in large colonies that were almost uniformly positive for myosin VIIa and contained actin-rich protrusions within the inner lumen (Figure 5A; Movie S1). Closer inspection of hair cell colonies revealed that myosin VIIa<sup>+</sup> cells contained CtBP2<sup>+</sup> ribbon synapse-like puncta in the basal region, where ribbon synapses are found in native hair cells



**Figure 4. Hair Cell Colonies Are Generated from Lgr5+ Colonies**

(A) Flow cytometry was performed for quantification of Atoh1+ cells in multiple expansion (blue bars) and differentiation (red bars) conditions in cultures originating from Atoh1-nGFP mice. Inner ear cell culture in growth factors (EFI) or growth factors with VPA and CHIR (EFICVP) did not change the percentage of Atoh1-nGFP cells after 10 days of expansion (shown as day 0 of the experiment;  $p > 0.05$ ). A combination of LY411575 and CHIR (LYC) was the most effective for differentiation of Atoh1-nGFP cells from EFICVP-expanded cells and was therefore compared to each condition.  $n = 5$ . Error bars represent mean  $\pm$  SD. \*\* $p < 0.0001$ ; \* $p < 0.05$ ; medium without growth factors or drugs (Med). D, day. (B) Lgr5-Cre-tdTomato cells were cultured with EFICVP for 10 days. 4-Hydroxytamoxifen was added to the culture at day 0 of expansion. Expression of Lgr5-GFP and tdTomato is shown. Scale bar, 15  $\mu$ m;  $n = 9$ . (C) Immunocytochemical staining of Lgr5-Cre-tdTomato cells for myosin VIIa following 10 days of culture in LYC. Scale bar, 15  $\mu$ m;  $n = 5$ . (D) EdU and myosin VIIa cells following 10 days of differentiation. Scale bar, 15  $\mu$ m;  $n = 5$ . (E) High-power view of the EdU+ cell in (D). EdU was added at day 1. The arrowheads in (D) and (E) refer to EdU-positive cells.

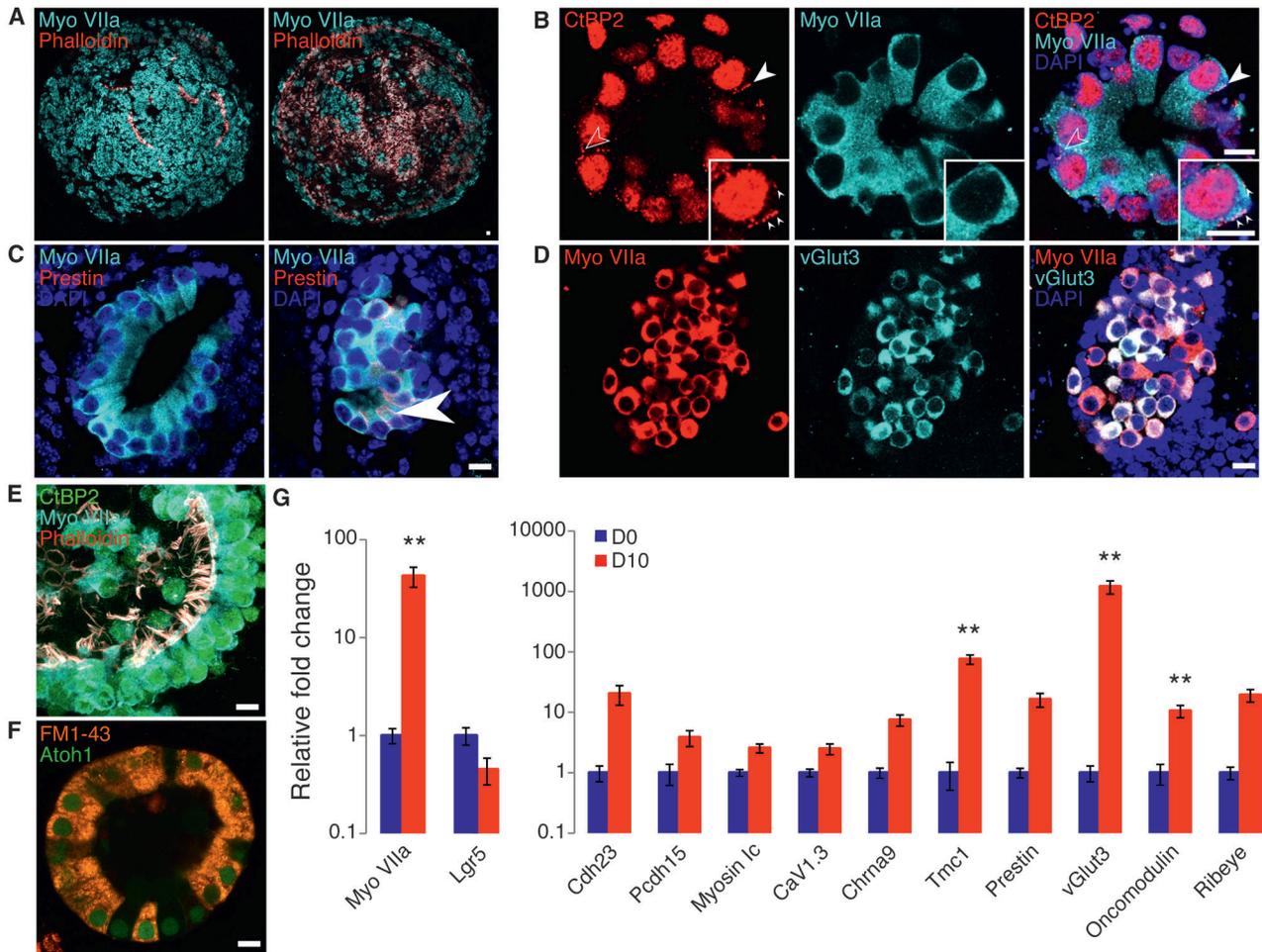
(Figure 5B, arrowheads). Hair cell colonies were either negative (Figure 5C, left) or positive (Figure 5C, right, arrowhead) for prestin, a motor protein located in the membrane of outer hair cells, identifying a subset of the differentiated cells as outer hair cells (Dallos et al., 2008). Colonies of new hair cells also expressed vesicular glutamate transporter3 (vGlut3), an inner hair cell marker (Seal et al., 2008), which was only found in colonies that did not express prestin (Figure 5D). The staining also revealed that the actin-rich protrusions comprised several individual stereocilia on the cells' apical surface (Figure 5E). The new colonies of hair cells rapidly accumulated the dye FM1-43, which enters hair cells through active transduction channels (Meyers et al., 2003) (Figure 5F).

qPCR studies to determine gene expression profiles before and after differentiation using the optimal conditions for Atoh1-nGFP quantification (EFICVP/LYC) revealed that myosin VIIa was upregulated between day 0 and day 10 after expansion, while Lgr5 expression decreased (Figure 5G). The differentiated cells expressed the tip-link genes cadherin23 and protocadherin15 (Hudspeth, 2008). The transduction adaptation component myosin Ic (Hudspeth, 2008), the synapse-associated calcium channel CaV1.3, and the ribbon synapse component ribeye were also upregulated (Khimich et al., 2005). The  $\alpha$ 9 acetylcholine receptor Chrna9 and the transduction channel component transmembrane channel 1 (Tmc1) (Pan et al., 2013a) both had increased expression. Prestin, the motor protein, and oncomodulin, a calcium modulator, both of which are found in outer hair cells, as well as the inner hair cell calcium modulator vesicular glutamate transporter 3, also showed increased expression. Our differentiation conditions thus gener-

ated inner and outer cochlear hair cell types that contained components of synaptic specializations, the transduction apparatus receptors, and ion channels of hair cells that were identified by both staining for specific markers and real-time qPCR.

#### Expansion and Hair Cell Differentiation of Lgr5-Expressing Cells from Adult Inner Ear Tissue

Previous studies have documented a decline in proliferative capacity and stem cell properties in the inner ear after the early postnatal period (Oshima et al., 2007; White et al., 2006). Since the drug combination applied here enhanced proliferation of neonatal cells compared to previous techniques, we next tested whether the compounds could be used to expand and differentiate otherwise quiescent adult cells into hair cells. Given the low numbers of Lgr5+ cells available from the adult mouse cochlea, we applied the cocktail of agents that we established for passaging neonatal cells, EFICVP6, to generate clonal colonies of adult cells positive for Lgr5 (Figure 6A). After the expansion, the cultures were treated with LY411575 and CHIR to differentiate the Lgr5+ cells. The colonies initiated expression of myosin VIIa only after differentiation (Figures 6B and 6C), indicating that adult Lgr5+ cells expanded and differentiated into cells that expressed hair cell markers. Myosin VIIa expression varied between colonies, with some colonies expressing the protein more robustly (Figures 6B and 6C). Cells isolated from mice at postnatal day 30 ( $4.67 \pm 0.28$  cells) and postnatal day 60 ( $6.75 \pm 1.53$  cells) formed similar sized colonies (average colony size across both ages of  $6.18 \pm 0.13$  cells; Figure 6D). Colonies from postnatal day 30 ( $4.33 \pm 0.28$  cells) and postnatal day 60 ( $3.88 \pm 1.48$  cells) generated a similar number of myosin



### Figure 5. Colonies of Inner Ear Progenitor Cells Generate Hair Cells In Vitro

(A) The combination of CHIR and LY411575 (LYC) converted progenitor colonies into high-purity populations of myosin VIIa<sup>+</sup> cells (left, surface view) with actin-rich protrusions projecting into the lumen (right, section through the colony). n = 11. Scale bar, 15  $\mu$ m.

(B) Myosin VIIa<sup>+</sup> cells had CtBP2<sup>+</sup> puncta at the basal end of the cell near the membrane (arrowheads). n = 4. Scale bar, 15  $\mu$ m.

(C) Myosin VIIa<sup>+</sup>/prestin<sup>-</sup> colonies, indicative of inner hair cells and myosin VIIa<sup>+</sup>/prestin<sup>+</sup> (arrowhead) colonies, indicative of outer hair cells, were distinct. n = 4. Scale bar, 15  $\mu$ m.

(D) Myosin VIIa<sup>+</sup>/vGlut3<sup>+</sup> cells were also produced, indicative of inner hair cells. n = 5. Scale bar, 15  $\mu$ m.

(E) Myosin VIIa<sup>+</sup> cells had actin-rich bundles on the apical surface comprising several individual stereocilia. n = 11. Scale bar, 15  $\mu$ m.

(F) Atoh1-nGFP colonies incorporated FM1-43 dye. Image is a single optical slice. n = 6. Scale bar, 15  $\mu$ m.

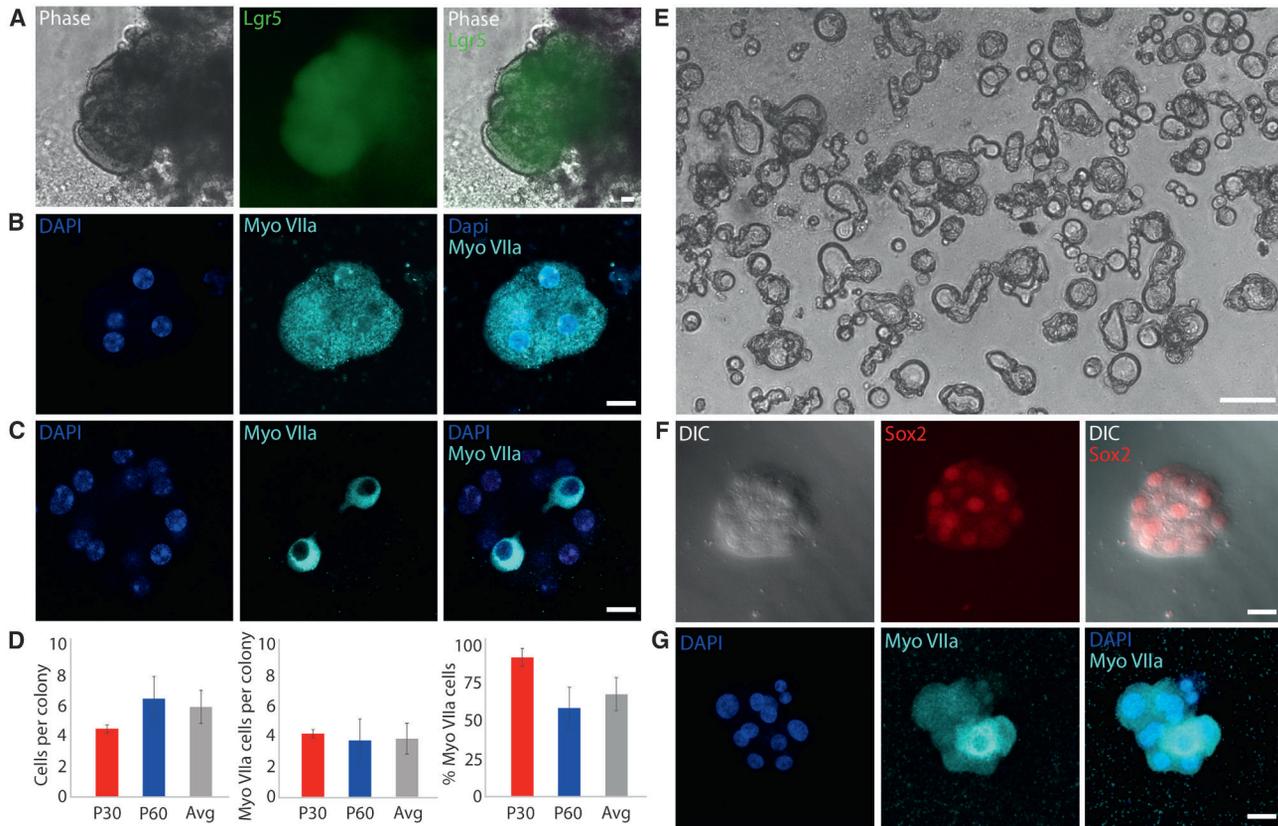
(G) Key hair cell genes were compared by real-time qPCR at days 0 and 10 of differentiation. Myosin VIIa expression increased while Lgr5 expression decreased between days 0 and 10. Hair cell genes (CaV1.3, Ribeye, Chrna9, Tmc1, Pcdh15, Cdh23, myosin 1c, prestin, oncomodulin, and vGlut3), which include the markers measured by antibody staining (A–D), were strongly upregulated. n  $\geq$  5 independent samples per gene. Error bars represent mean  $\pm$  SEM;  $p < 0.05$  for all genes presented; \*\* $p < 0.005$ . D, day.

VIIa<sup>+</sup> cells per colony (average across ages of  $4.00 \pm 1.06$  myosin VIIa<sup>+</sup> cells per colony; Figure 6D). The myosin VIIa<sup>+</sup> cells were obtained in colonies generated from postnatal day 30 ( $93.3\% \pm 5.8\%$  cells) and postnatal day 60 ( $59.6\% \pm 13.7\%$  cells); (average across ages of  $68.82\% \pm 0.28\%$  myosin VIIa cells; Figure 6D). No significant differences were seen across ages.

We next tested whether the expansion and differentiation condition could be applied to non-human primates. Inner ear epithelial cells were isolated from adult rhesus macaques and cultured with EFICVP6. These preliminary results indicated that the cells

formed clonal colonies (Figure 6E). However, differentiation to hair cells was not achieved due to repeated contamination likely caused by non-sterile conditions encountered during the temporal bone isolation.

We further tested the conditions using one sample of healthy human inner ear tissue isolated from a 40-year-old male patient undergoing a labyrinthectomy to access a tumor on the brain. The inner ear tissue was microdissected to remove bone, debris, and nerve tissue. The tissue was then treated identically to the mouse tissue to isolate single cells for culture. The single cells formed clonal colonies after 12 days under EFICVP6 conditions,



**Figure 6. Expansion and Differentiation of Progenitor Cells from Adult Mouse, Rhesus Macaque, and Human Inner Ear**

(A) Adult mouse Lgr5-GFP cells formed colonies in EFICVP6.  $n = 4$ . Scale bar, 15  $\mu\text{m}$ .

(B) Upon differentiation with LYC, subsets of colonies contained high purity populations of cells expressing hair cell gene, myosin VIIa (Myo VIIa).  $n = 4$ . Scale bar, 15  $\mu\text{m}$ .

(C) Other colonies showed a smaller percentage of myosin VIIa+ cells.  $n = 4$ . Scale bar, 15  $\mu\text{m}$ .

(D) Left: cells isolated from 30-day and 60-day old (p30 and p60) animals formed similar sized colonies ( $p > 0.05$ ). Average (Avg) is also shown. Middle: colonies from 30- and 60-day-old animals generated a similar number of myosin VIIa+ cells per colony ( $p > 0.05$ ). Average (Avg) is also shown. Right: myosin VIIa+ cells were represented in similar proportions in colonies from 30- and 60-day-old animals ( $p > 0.05$ ). Average (Avg) is also shown. p30  $n = 3$ ; p60  $n = 8$ . Error bars represent mean  $\pm$  SD. P, postnatal day.

(E) Cells isolated from adult rhesus macaque inner ear epithelia generated clonal colonies in EFICVP6 for 7 days.  $n = 4$ . Scale bar, 50  $\mu\text{m}$ .

(F) Cells isolated from human inner ear epithelia from a 40-year-old male generated clonal colonies that stained for Sox2 after a 12-day EFICVP6 treatment.  $n = 1$ . Scale bar, 15  $\mu\text{m}$ . DIC, differential interference contrast.

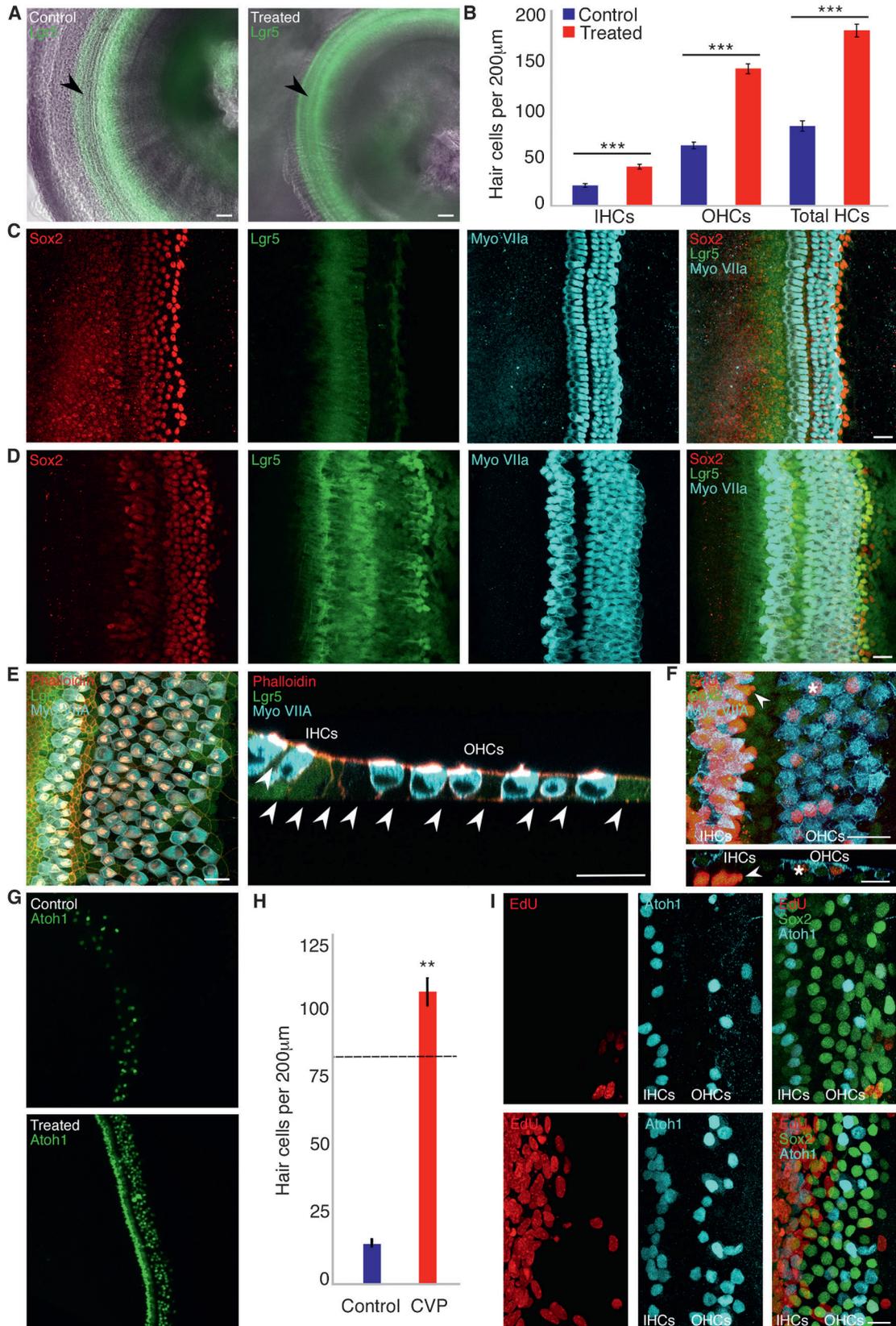
(G) LYC treatment of human inner ear colonies generated populations of hair cell-like cells (Myo VIIa) with few myosin VIIa- cells (arrow). Colony size was  $7.25 \pm 1.74$  cells. The number of myosin VIIa+ cells per colony was  $5.25 \pm 2.21$ . The proportion of myosin VIIa+ cells was  $66.7\% \pm 18.0\%$ . Scale bar, 15  $\mu\text{m}$ .

although expansion was not as robust as that seen for neonatal cells (Figure 6F). The colonies stained for Sox2, a known marker of inner ear progenitor cells (Figure 6F). After 12 days of expansion, the cultures were treated with LY411575 and CHIR for 10 days to differentiate the colonies. The colonies stained positively for the hair cell marker myosin VIIa (Figure 6G), suggesting that sensory epithelium from adult human inner ear can also give rise to hair cell progenitors.

#### Hair Cell Generation In Situ in Cochlear Explants

The drugs that were critical for expansion and differentiation of Lgr5+ cells (HDAC inhibitors, GSK3 $\beta$  inhibitors, and  $\gamma$ -secretase inhibitors) have all been used clinically for other indications and could potentially be candidates for clinical development. To investigate their effects in a more clinically relevant tissue, we

applied the drugs to cochlear explants. Supporting cells play a key role in cochlear function and homeostasis. Therefore, we treated intact and hair cell-damaged explants from postnatal day 2 mice with small molecules from the expansion conditions (CVP) rather than the differentiation conditions in an attempt to maintain a supporting cell population and permit spontaneous differentiation (Figure S4). We performed these cultures without growth factors in the presence of the surrounding tissue. These tests resulted in extensive proliferation of supporting cells and differentiation to hair cells. Whereas Lgr5-GFP was absent in a control cochlea in the region between the third Deiters cell and inner pillar cells (i.e., outer pillar cells, first and second Deiters cells), treatment with CVP for 3 days caused upregulation of Lgr5-GFP in all supporting cells (Figures 7A, 7C, and 7D). There was a highly significant ( $p < 0.001$ )  $\sim$ 2-fold increase in myosin



(legend on next page)

VIIa+ inner and outer hair cells after 3 days of drug treatment (Figures 7B and 7D) as compared to control cochlea (Figure 7C). Addition of 616452 to VPA and CHIR did not increase the generation of hair cells. The new hair cells had morphology similar to the intact cochlea, with phalloidin+ stereociliary bundles and hair cells that were separated by intact supporting cells, suggesting that the treatment caused proliferation and subsequent differentiation (Figure 7E). Supporting cells had incorporated EdU, suggesting that they had divided (Figure 7F), and some of the hair cells, identified as “new” based on their expression of Sox2 (Bramhall et al., 2014; Kempfle et al., 2016), had transdifferentiated from supporting cells that had taken up EdU, similar to the division of supporting cells stimulated by Wnt signaling (Shi et al., 2013).

Hair cell regeneration was achieved in cochlear explants treated with gentamicin, which causes hair cell death in the basal portion of the organ of Corti, where transduction channels are active in the neonate (Figures 7G–7I). Gentamicin caused extensive hair cell death in the base of the cochlea, but after 3 days of treatment with CVP, new Atoh1-nGFP hair cells appeared (Figures 7G and 7H). The number of hair cells was close to normal after treatment and 7-fold greater than that observed for control-treated cochlea (Figure 7H). Supporting cells were EdU+, indicating that supporting cell division was a part of the mechanism for hair cell replacement (Figure 7I). Thus, the treatment with CVP that expanded Lgr5+ cells from the cochlea after their isolation and placement into a 3D culture was also able to expand supporting cells *in situ* and force the generation of new hair cells.

## DISCUSSION

Lgr5+ stem cells have been identified in epithelial cells of a number of tissues, including the intestine, colon, stomach, and liver (Barker et al., 2007; Koo and Clevers, 2014). Lgr5+ cells from these tissues can be induced to form organoids when cultured in the presence of Wnt pathway activators, including R-spondin 1, and contain a heterogeneous population of cells. Previously, we identified Lgr5 as a marker for progenitor cells in the newborn mouse cochlea (Chai et al., 2012; Shi et al., 2012, 2013). The cells that expressed Lgr5 were supporting cells that surround the hair cells of the cochlea. Similar to the Lgr5-expressing stem cells in

the gut, these cells were Wnt responsive and could be stimulated to divide and differentiate to some extent by forced activation of Wnt signaling (Shi et al., 2013), even though the postnatal mammalian cochlea is normally quiescent. Further, although these previous studies showed Wnt stimulation could induce division in Lgr5+ cells, the limited potential for propagation and conversion to hair cells suggested that other pathways might be required to increase the stem cell capacity of Lgr5+ cells in the cochlea. Here, we show that Lgr5+ cells from the inner ear can also be extensively expanded with a GSK3 $\beta$  inhibitor to activate the Wnt signaling pathway combined with an HDAC inhibitor to activate Notch signaling. When provided with additional cues, specifically, 2-phospho-L-ascorbic acid, which was previously shown to facilitate induced pluripotent stem cell (iPSC) generation (Esteban et al., 2010), and the TGF- $\beta$  inhibitor 616452, which regulates cell senescence (Hua and Thompson, 2001), neonatal cells could be passaged and clonal colonies of adult murine, primate, and human progenitor cells could be generated (Figures 6A, 6E, and 6F). Differentiating these cells by simultaneously activating Wnt and inhibiting Notch enabled conversion of progenitor cells into high purity populations of hair cells. Moreover, treatment of cochlear tissue with small molecules to simultaneously activate Wnt and Notch led to upregulation of Lgr5 in all supporting cells and increased numbers of Lgr5+ cells and hair cells (Figure 7). The increase in hair cell number was achieved even in cochlear tissue that had been depleted of hair cells by exposure to an aminoglycoside antibiotic. The effect of this drug combination on the cochlea suggests that small molecules activating Wnt and Notch could be useful as a therapeutic option to restore hair cells without loss of the supporting cells, which are important for cochlear homeostasis and mechanics.

In our previous work, we have shown that simultaneously providing Wnt and Notch signaling synergistically maintains self-renewal of Lgr5+ cells from the mouse small intestine, stomach, colon, and human small intestine. The expanded Lgr5+ cells could be used to generate mature intestinal epithelial cells, including Paneth cells, goblet cells, and enterocytes (Yin et al., 2014). Intestine, colon, stomach, and liver epithelia are actively renewed or activated upon injury, whereas the cochlear cells do not regenerate tissue spontaneously. However, like the stem cells from the intestine (de Lau et al., 2011; Koo and

### Figure 7. Increase in Hair Cell Numbers after Treatment of Cochlear Explants with GSK3 $\beta$ and HDAC Inhibitors

- (A) Cells between the third Deiters and inner border cells (arrowhead) had increased Lgr5-GFP expression in a cochlea treated with CHIR, VPA, and pVc (CVP). n = 9. Scale bar, 25  $\mu$ m.
- (B) Increased numbers of inner hair cells, outer hair cells, and total hair cells (IHCs, OHCs, and total HCs) were observed in treated as compared to control cochleae by myosin VIIa (Myo VIIa) expression. n = 4 each. Error bars represent mean  $\pm$  SD; \*\*\*p < 0.001.
- (C) Control cochleae had typical Lgr5-GFP expression, one row of inner hair cells, and three rows of outer hair cells. n = 3. Scale bar, 15  $\mu$ m.
- (D) A cochlear epithelium had an increased number of hair cell after CVP treatment. n = 4. Scale bar, 15  $\mu$ m.
- (E) Treated cochlear explant (left) had extra hair cells. The new hair cells possessed microvillar bundles in an orthogonal view (right). Supporting cells remained between new hair cells (arrowheads) as outlined by phalloidin staining. n = 5. Scale bars, 15  $\mu$ m.
- (F) Treated cochlear explant (top) showed supporting cells (Sox2+, arrowheads), hair cells (myosin VIIa+, asterisks), and EdU. Staining for EdU was visible in both supporting cells and hair cells (orthogonal view; bottom). n = 3. Scale bar, 25  $\mu$ m.
- (G) A cochlea damaged by gentamicin following a 3-day treatment with CVP had an increased number of Atoh1+ cells in the inner and outer hair cell regions. n = 3 each. Scale bar, 25  $\mu$ m.
- (H) Treatment with CVP increased hair cell numbers after gentamicin exposure. Dashed line represents hair cell counts in a healthy mouse cochlea. n = 3 each. Error bars represent mean  $\pm$  SD. \*\*p < 0.01.
- (I) EdU incorporation into a gentamicin-treated cochlear explant (top) compared to a gentamicin and CVP-treated cochlear explant (bottom). EdU and CVP were added at 16 hr. n = 4. Scale bars, 15  $\mu$ m.

Clevers, 2014), the *Lgr5*-expressing cells in the cochlea show Wnt-dependent cell division (Shi et al., 2012). Wnt signaling and Notch inhibition also stimulate *Atoh1*-dependent differentiation to mature cells, similar to intestinal epithelium (Jacques et al., 2012; Shi et al., 2012, 2014; Yang et al., 2001). *Lgr5* is a potentiator of the Wnt pathway through interaction with its ligand, R-spondin, and is uniquely driven by Wnt signals (de Lau et al., 2011; Koo and Clevers, 2014). R-spondin binding to *Lgr5*, along with Wnt binding to the Frizzled receptor, potentiates Wnt activity by as much as 100-fold (de Lau et al., 2011). The up-regulation of *Lgr5* in organoids from the cochlea after treatment with the GSK3 $\beta$  inhibitor was consistent with a role of Wnt signaling upstream of *Lgr5* (Barker et al., 2007). Expansion of *Lgr5*<sup>+</sup> intestinal cells is dependent on Notch (van Es et al., 2005), which also expands progenitor cells in the cochlea (Jeon et al., 2011; Kelley, 2006; Pan et al., 2013b; Shi et al., 2010, 2014). In both the intestine and the cochlea, Notch plays a dual role, requiring cell-specific downregulation to establish the fate of supporting cells and hair cells in the patterning of the mature organ (Yamamoto et al., 2011). Wnts are released in response to injury, and this comprises one of the mechanisms through which stem cells may be activated in tissue repair (Liu et al., 2013; Whyte et al., 2012). Wnt signaling also stimulated the growth of *Lgr5*<sup>+</sup> cells from the pancreas, which like the cochlea is a non-regenerative tissue (Huch et al., 2013). This study shows the generation of *Lgr5*<sup>+</sup> organoids from non-endodermal epithelium. In the CNS, where a similar restricted role of stem cells has been noted, the stem cell compartments are Wnt dependent (Hochedlinger et al., 2005; Seib et al., 2013), and Wnt activity and responsiveness have been used to identify stem cell compartments (Denayer et al., 2008; Hochedlinger et al., 2005; Seib et al., 2013). These pathways provided a basis for the approach studied here for expanding *Lgr5*-expressing cells from the cochlea.

Although we have shown that inner ear cells could be isolated and cultured under neurosphere conditions and differentiated into hair cells and neural cells (Li et al., 2003; Martinez-Monedero et al., 2008; Oshima et al., 2007), the overall yield of hair cells was low. Since only a few thousand hair cells exist in the healthy adult mammalian cochlea and since the tissue is difficult to obtain and harvest from humans, a readily available source of bona fide hair cells would represent a major advance that would enable physiological studies as well as genetic screens using drugs, small interfering RNA (siRNA), or gene overexpression. Here, we demonstrated that *Lgr5*<sup>+</sup> supporting cells from neonatal and adult cochlea responded to the protocols using Wnt stimulation and HDAC inhibition, giving rise to colonies with a high capacity for mature tissue cell differentiation. While the extent of expansion of *Lgr5*<sup>+</sup> cells from adult mouse tissue was less than that from neonatal cells, the normally quiescent supporting cells from the adult mouse, rhesus, and human inner ear responded to the small-molecule cocktail used on neonatal tissue. However, additional molecules will likely be required to enhance the expansion of *Lgr5*<sup>+</sup> cells from the adult cochlea. Importantly, hair cells produced from the protocols described here typify cochlear hair cells in both morphology and gene expression and possess key components necessary for proper function and communication with neurons. These hair cells possess ster-

eo-ciliary bundles that express transduction-associated genes (cadherin 23, protocadherin 15, myosin 1c, and transmembrane channel1), synaptic and neurotransmitter genes (CaV1.3, ribeye,  $\alpha$ 9-acetylcholine receptor, vesicular glutamate transporter 3, and oncomodulin), and other genes required for inner (vesicular glutamate transporter 3) and outer (prestin) hair cell function. We anticipate that *Lgr5*<sup>+</sup> cells will need to be therapeutically targeted in situ to stimulate functional regeneration of hair cells, because cochlear mechanics rely on the precise anatomy of the organ and are unlikely to be recreated by disruptive cell transplantation approaches; thus, our identification of small molecules that can expand these cells is of particular significance. To date, difficulty in generating large numbers of primary hair cell progenitors has limited genetic and physiological studies for advancement of potential therapies. The small-molecule and growth factor cocktails used here enable the clonal expansion of *Lgr5*<sup>+</sup> cells from mammalian cochlea and generation of sensory hair cells that will be useful for biological studies and will serve to generate a source of cells for both drug and genetic screens. Further, this work suggests that a small-molecule approach to activate Wnt and Notch could be a viable therapeutic route to restore hair cells.

## EXPERIMENTAL PROCEDURES

### Mouse Strains

*Lgr5*-EGFP-IRES-Cre-ER mice (The Jackson Laboratory, strain 8875) (Barker et al., 2007) were used to analyze the effects of small molecules on cochlear stem cell expansion. The same mice were crossed with Rosa26-td-Tomato reporter mice (The Jackson Laboratory, strain 7909) (Madisen et al., 2010) to create a mouse line that enabled lineage tracing of the cells that resulted from differentiated *Lgr5*-expressing cells. *Atoh1*-nGFP mice (Lumpkin et al., 2003) (provided by Dr. Jane Johnson) were used to identify differentiated hair cells.

### Isolation of Stem Cells from the Inner Ear

All animal studies were conducted under an approved institutional protocol according to National Institutes of Health guidelines. For experiments with neonatal mice (postnatal days 1–3), the cochleae were dissected in Hank's balanced salt solution (HBSS), and the organ of Corti (sensory epithelium) was separated from the stria vascularis (ion transport epithelium) and the modiolus (nerve tissue). The organs of Corti were then treated with Cell Recovery Solution (Corning) for 1 hr to separate cochlear epithelium from the underlying mesenchyme. Epithelia were then collected and treated with TrypLE (Life Technologies) for 15–20 min at 37°C. Single cells obtained by mechanical trituration were filtered (40  $\mu$ m) and suspended in a Matrigel (Corning) dome for 3D culture. Laminin (0.015 mg/mL) was added to the Matrigel where indicated. For adult tissue, the stria vascularis was removed, but the epithelium was not removed from the underlying mesenchyme due to the limited amount of intact cochlea that could be extracted. Rhesus macaque tissue was acquired from the New England Primate Research Center. Temporal bones from adult rhesus macaques were dissected to obtain the inner ear and micro-dissected to remove bone, debris, and nerve tissue. The tissue was then processed and treated identically to the mouse cells to make a 3D culture of single cells. Human tissue was acquired after surgical removal in accordance with the policy of the Mass Eye and Ear Institutional Review Board for research use of discarded tissue. The tissue was then microdissected to isolate the sensory epithelium.

### Expansion of *Lgr5*<sup>+</sup> Cells

Cells were cultured in a 3D system and bathed in a serum free 1:1 mixture of DMEM and F12, supplemented with Glutamax (GIBCO), N2, B27 (Invitrogen), EGF (50 ng/mL; Chemicon), bFGF (50 ng/mL; Chemicon), IGF-1 (50 ng/mL; Chemicon), and small molecules, including CHIR99021 (3  $\mu$ M), VPA (1 mM),

pVc (100  $\mu$ g/mL), and 616452 (2  $\mu$ M) (Table S1). Media were changed every other day. 4-Hydroxytamoxifen (20 ng/mL) was added to cultures on day 0 for lineage tracing studies.

### Differentiation of Lgr5+ Progenitor Cells

To differentiate stem cell colonies, the expansion media was removed and colonies remained in 3D culture. A serum-free 1:1 mixture of DMEM and F12, supplemented with Glutamax (GIBCO), N2, and B27, (Invitrogen) was added with various combinations of drugs or growth factors, as indicated in Figure 2. Small molecules were added to the culture to test their effect on differentiation. The optimal differentiation conditions included CHIR99021 (3  $\mu$ M) and LY411575 (5  $\mu$ M).

### Quantification of Cells In Vitro

Lgr5+ cells were quantified after 10 days in culture in multiple conditions. Cell colonies were dissociated into single cells using TrypLE (GIBCO). Cell counting was performed with a hemocytometer. The cells were then stained with propidium iodide (PI) and analyzed using a flow cytometer for Lgr5-GFP expression. The number of GFP+ cells was calculated by multiplying the total number of cells by the percentage of GFP+ cells. Tukey's post hoc test was used to analyze Lgr5-GFP expansion conditions.

To quantify hair cell production, cochlear epithelial cells were isolated as described above and seeded at one cochlea per well of a 24-well plate. This corresponded to a seeding density that was  $\sim$ 130,000 viable cells/mL, or 3,900 viable cells/well, which was determined using a hemocytometer and counting DAPI+ and DAPI- cells. More specifically, seeding at one cochlea per well corresponded to an Lgr5-GFP cell density of 23,750 cells/mL (710 cells/well) and an Atoh1-nGFP+ cell density of 740 cells/mL (20 cells/well).

Atoh1-nGFP+ cells were quantified at day 0 and day 10 of differentiation treatment to determine the number of Atoh1+ cells. Cell colonies were incubated in Cell Recovery Solution (Corning) to release the colonies from Matrigel and dissociated into single cells using TrypLE. The total number and percentage of GFP+ cells were quantified using fluorescence-activated cell sorting (FACS) of multiple culture conditions. ANOVA was used to compare means across conditions, and the two-tailed Student's t test was used to compare each condition to the treatment with the highest yield.

### Cochlear Explant Studies

Cochleae were isolated from 2-day-old Lgr5-GFP or Atoh1-nGFP mice and transferred to HBSS (Invitrogen). The organ of Corti was isolated from the otic capsule, and the nerve tissue and stria vascularis were removed. The basal hook portion of the organ of Corti was removed to allow for optimal plating of the organ. The organ of Corti was plated on a glass coverslip that had been coated with a 1:10 mixture of serum-free 1:1 DMEM/F12 and Matrigel to promote attachment to the coverslip. One cochlea from each animal served as a control, and the other was treated. Cochlear explants were cultured in a serum-free 1:1 mixture of DMEM and F12, supplemented with Glutamax, N2, and B27. For the treated cochlea, small molecules were added to this culture medium without growth factors present, while the control cochlea was bathed in culture media with DMSO at the same concentration (0.03%–0.13%) used in the treatments.

The number of cells expressing a hair cell protein within a 200- $\mu$ m segment of the midbasal region of the cochlea was counted. The organs were treated with small molecules for 3 days. Lgr5-GFP cochlea and myosin VIIa staining were used for analysis of the undamaged cochlea.

To test hair cell generation after ablation, cochlear explants were prepared as described above using cochlear tissue isolated from Atoh1-nGFP mice. 1 hr after dissection, the organ of Corti explants were treated with 50  $\mu$ M gentamicin (Sigma) for 16 hr to ablate hair cells as described previously (Bramhall et al., 2014). The organs were then treated for 3 days under control conditions that contained no growth factors or small molecules or with the small molecules used in our proliferation conditions. Atoh1-nGFP cells were counted in the basal portion of the cochlea using a 200- $\mu$ m window.

### Immunohistochemistry

Colonies or cochlear explants were fixed at room temperature in 4% paraformaldehyde/PBS for 15–20 min and then washed in PBS. Permeabilization of

the cellular membrane (0.3% Triton X-100) was followed by blocking solution (15% heat inactivated donkey serum in PBS for 1 hr). Diluted primary antibody (0.1% Triton X-100 and 10% heat inactivated donkey serum in PBS) was applied for 4 hr at room temperature or overnight at 4°C. Primary antibody dilutions are listed in Table S2. Secondary antibodies (Alexa 488, Alexa Fluor 568, and Alexa Fluor 647 conjugated; Invitrogen) were used at 1:500 dilutions. Nuclei were visualized with DAPI (Vector Laboratories).

EdU studies were performed using the Click-iT EdU imaging kit (Thermo Fisher Scientific). For Lgr5+ cell proliferation studies, EdU (5  $\mu$ M) was pulsed for 1.5 hr, and the cultures were washed and stained immediately after exposure. To test proliferation of hair cells, EdU was added 1 day after initial LYC exposure, and cultures were stained after 10 days. Explant cultures were exposed to EdU together with the expansion conditions for a total of 3 days.

For transduction assays, 5  $\mu$ M FM1-43X (Invitrogen) was administered to cells for 30 s and then washed three times in 1 $\times$  HBSS prior to fixation and further staining. Staining was visualized with confocal microscopy (TCD, Leica).

### RNA Extraction and Real-Time qPCR

Colonies of expanded and differentiated cells were treated for 45–60 min with Cell Recovery Solution (Corning) to extract the cells from Matrigel and subsequently frozen in RLT buffer. RNA was extracted using the RNeasy Micro Kit (QIAGEN), and cDNA was generated using ImProm-II Reverse Transcription Kit (Promega). SYBR green real-time qPCR was performed in duplicates using SYBR Select Master Mix (Applied Biosystems) on a StepOne Real-Time PCR machine (Applied Biosystems). Specific primers were designed for all genes with mouse Gapdh as reference standard. A list of primers can be found in Table S3. The Student's t test was used for analysis of gene expression, and at least five biologically distinct samples were analyzed for each condition.

### Statistical Analysis

Results are reported as mean  $\pm$  SD, where n represents the number of independent experiments conducted in the same manner. For quantification of cells in vitro, ANOVA and Tukey's post hoc tests were used to compare means across conditions, while the two-tailed Student's t test was used to compare each condition to the treatment with the highest yield. The Student's t test was also used to compare hair cell counts from control and treated cochlear explants and to analyze gene expression levels at the beginning and end of in vitro differentiation. A p value < 0.05 was considered significant.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, three tables, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.01.066>.

### AUTHOR CONTRIBUTIONS

W.J.M., X.Y., R.L., J.M.K., and A.S.B.E. designed research; W.J.M., X.Y., L.L., D.R.L., and D.M. performed research; W.J.M., X.Y., D.R.L., J.M.K., and A.S.B.E. analyzed data; and W.J.M., X.Y., D.R.L., J.M.K., and A.S.B.E. wrote the paper.

### CONFLICTS OF INTEREST

J.M.K., R.S.L., X.Y., and W.J.M. hold equity in Frequency Therapeutics, a company that has an option to license IP generated by J.M.K., R.S.L., and X.Y. and that may benefit financially if the IP is licensed and further validated. W.J.M. is an employee of Frequency Therapeutics. The interests of J.M.K., R.S.L., and X.Y. were reviewed and are subject to a management plan overseen by their institutions in accordance with their conflict of interest policies.

### ACKNOWLEDGMENTS

We thank Lisa Ogawa and Marco Petrillo for contributions to this work. This work was supported by grants from the NIH (DE-013023, DC-007174,

DC-013909, and RR-00168), a cooperative grant from the European Commission (FP7 Health 2013 Innovation), a Harvard-MIT IDEA2 award, the Shulsky Foundation, David H. Koch, and Robert Boucai.

Received: December 30, 2015

Revised: January 5, 2017

Accepted: January 25, 2017

Published: February 21, 2017

## REFERENCES

- 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.
- Bramhall, N.F., Shi, F., Arnold, K., Hochedlinger, K., and Edge, A.S. (2014). *Lgr5*-positive supporting cells generate new hair cells in the postnatal cochlea. *Stem Cell Reports* 2, 311–322.
- Chai, R., Kuo, B., Wang, T., Liaw, E.J., Xia, A., Jan, T.A., Liu, Z., Taketo, M.M., Oghalai, J.S., Nusse, R., et al. (2012). Wnt signaling induces proliferation of sensory precursors in the postnatal mouse cochlea. *Proc. Natl. Acad. Sci. USA* 109, 8167–8172.
- Cox, B.C., Chai, R., Lenoir, A., Liu, Z., Zhang, L., Nguyen, D.H., Chalasani, K., Steigelman, K.A., Fang, J., Rubel, E.W., et al. (2014). Spontaneous hair cell regeneration in the neonatal mouse cochlea in vivo. *Development* 141, 816–829.
- Dallos, P., Wu, X., Cheatham, M.A., Gao, J., Zheng, J., Anderson, C.T., Jia, S., Wang, X., Cheng, W.H., Sengupta, S., et al. (2008). Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. *Neuron* 58, 333–339.
- Davis, A.C. (1983). Hearing disorders in the population: first phase findings of the MRC National Study of Hearing. In *Hearing Science and Hearing Disorders*, M.E. Lutman and M.P. Haggard, eds. (Academic Press), pp. 35–60.
- de Lau, W., Barker, N., Low, T.Y., Koo, B.K., Li, V.S., Teunissen, H., Kujala, P., Haegebarth, A., Peters, P.J., van de Wetering, M., et al. (2011). *Lgr5* homologues associate with Wnt receptors and mediate R-spondin signalling. *Nature* 476, 293–297.
- Denayer, T., Locker, M., Borday, C., Deroo, T., Janssens, S., Hecht, A., van Roy, F., Perron, M., and Vleminckx, K. (2008). Canonical Wnt signaling controls proliferation of retinal stem/progenitor cells in postembryonic *Xenopus* eyes. *Stem Cells* 26, 2063–2074.
- Edge, A.S., and Chen, Z.Y. (2008). Hair cell regeneration. *Curr. Opin. Neurobiol.* 18, 377–382.
- Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al. (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* 6, 71–79.
- Fujioka, M., Okano, H., and Edge, A.S. (2015). Manipulating cell fate in the cochlea: a feasible therapy for hearing loss. *Trends Neurosci.* 38, 139–144.
- Greenblatt, D.Y., Vaccaro, A.M., Jaskula-Sztul, R., Ning, L., Haymart, M., Kunimallyayan, M., and Chen, H. (2007). Valproic acid activates notch-1 signaling and regulates the neuroendocrine phenotype in carcinoid cancer cells. *Oncologist* 12, 942–951.
- Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* 121, 465–477.
- Hua, X., and Thompson, C.B. (2001). Quiescent T cells: actively maintaining inactivity. *Nat. Immunol.* 2, 1097–1098.
- 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. (2013). Unlimited in vitro expansion of adult bi-potent pancreas progenitors through the *Lgr5/R-spondin* axis. *EMBO J.* 32, 2708–2721.
- Hudspeth, A.J. (2008). Making an effort to listen: mechanical amplification in the ear. *Neuron* 59, 530–545.
- Jacques, B.E., Puligilla, C., Weichert, R.M., Ferrer-Vaquero, A., Hadjantonakis, A.K., Kelley, M.W., and Dabdoub, A. (2012). A dual function for canonical Wnt/ $\beta$ -catenin signaling in the developing mammalian cochlea. *Development* 139, 4395–4404.
- Jeon, S.J., Fujioka, M., Kim, S.C., and Edge, A.S.B. (2011). Notch signaling alters sensory or neuronal cell fate specification of inner ear stem cells. *J. Neurosci.* 31, 8351–8358.
- Kelley, M.W. (2006). Regulation of cell fate in the sensory epithelia of the inner ear. *Nat. Rev. Neurosci.* 7, 837–849.
- Kempfle, J.S., Turban, J.L., and Edge, A.S.B. (2016). Sox2 in the differentiation of cochlear progenitor cells. *Sci. Rep.* 6, 23293.
- Khimich, D., Nouvian, R., Pujol, R., Tom Dieck, S., Egner, A., Gundelfinger, E.D., and Moser, T. (2005). Hair cell synaptic ribbons are essential for synchronous auditory signalling. *Nature* 434, 889–894.
- Koo, B.K., and Clevers, H. (2014). Stem cells marked by the R-spondin receptor *LGR5*. *Gastroenterology* 147, 289–302.
- Korrapati, S., Roux, I., Glowatzki, E., and Doetzlhofer, A. (2013). Notch signaling limits supporting cell plasticity in the hair cell-damaged early postnatal murine cochlea. *PLoS ONE* 8, e73276.
- Li, H., Liu, H., and Heller, S. (2003). Pluripotent stem cells from the adult mouse inner ear. *Nat. Med.* 9, 1293–1299.
- Liu, B., Hunter, D.J., Rooker, S., Chan, A., Paulus, Y.M., Leucht, P., Nusse, Y., Nomoto, H., and Helms, J.A. (2013). Wnt signaling promotes Müller cell proliferation and survival after injury. *Invest. Ophthalmol. Vis. Sci.* 54, 444–453.
- Lumpkin, E.A., Collisson, T., Parab, P., Omer-Abdalla, A., Haeberle, H., Chen, P., Doetzlhofer, A., White, P., Groves, A., Segil, N., and Johnson, J.E. (2003). *Math1*-driven GFP expression in the developing nervous system of transgenic mice. *Gene Expr. Patterns* 3, 389–395.
- Madisen, L., Zwingman, T.A., Sunken, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* 13, 133–140.
- Martinez-Monedero, R., Yi, E., Oshima, K., Glowatzki, E., and Edge, A.S. (2008). Differentiation of inner ear stem cells to functional sensory neurons. *Dev. Neurobiol.* 68, 669–684.
- Meyers, J.R., MacDonald, R.B., Duggan, A., Lenzi, D., Standaert, D.G., Corwin, J.T., and Corey, D.P. (2003). Lighting up the senses: FM1-43 loading of sensory cells through nonselective ion channels. *J. Neurosci.* 23, 4054–4065.
- Mizutani, K., Fujioka, M., Hosoya, M., Bramhall, N., Okano, H.J., Okano, H., and Edge, A.S.B. (2013). Notch inhibition induces cochlear hair cell regeneration and recovery of hearing after acoustic trauma. *Neuron* 77, 58–69.
- Oshima, K., Grimm, C.M., Corrales, C.E., Senn, P., Martinez Monedero, R., Géléoc, G.S., Edge, A., Holt, J.R., and Heller, S. (2007). Differential distribution of stem cells in the auditory and vestibular organs of the inner ear. *J. Assoc. Res. Otolaryngol.* 8, 18–31.
- Pan, B., Géléoc, G.S., Asai, Y., Horwitz, G.C., Kurima, K., Ishikawa, K., Kawashima, Y., Griffith, A.J., and Holt, J.R. (2013a). TMC1 and TMC2 are components of the mechanotransduction channel in hair cells of the mammalian inner ear. *Neuron* 79, 504–515.
- Pan, W., Jin, Y., Chen, J., Rottier, R.J., Steel, K.P., and Kiernan, A.E. (2013b). Ectopic expression of activated notch or SOX2 reveals similar and unique roles in the development of the sensory cell progenitors in the mammalian inner ear. *J. Neurosci.* 33, 16146–16157.
- 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.
- Seal, R.P., Akil, O., Yi, E., Weber, C.M., Grant, L., Yoo, J., Clause, A., Kandler, K., Noebels, J.L., Glowatzki, E., et al. (2008). Sensorineural deafness and seizures in mice lacking vesicular glutamate transporter 3. *Neuron* 57, 263–275.
- Seib, D.R., Corsini, N.S., Ellwanger, K., Plaas, C., Mateos, A., Pitzer, C., Niehrs, C., Celikel, T., and Martin-Villalba, A. (2013). Loss of Dickkopf-1 restores

- neurogenesis in old age and counteracts cognitive decline. *Cell Stem Cell* 12, 204–214.
- Shi, F., Cheng, Y.F., Wang, X.L., and Edge, A.S. (2010). Beta-catenin up-regulates Atoh1 expression in neural progenitor cells by interaction with an Atoh1 3' enhancer. *J. Biol. Chem.* 285, 392–400.
- Shi, F., Kempfle, J.S., and Edge, A.S. (2012). Wnt-responsive Lgr5-expressing stem cells are hair cell progenitors in the cochlea. *J. Neurosci.* 32, 9639–9648.
- Shi, F., Hu, L., and Edge, A.S. (2013). Generation of hair cells in neonatal mice by  $\beta$ -catenin overexpression in Lgr5-positive cochlear progenitors. *Proc. Natl. Acad. Sci. USA* 110, 13851–13856.
- Shi, F., Hu, L., Jacques, B.E., Mulvaney, J.F., Dabdoub, A., and Edge, A.S. (2014).  $\beta$ -Catenin is required for hair-cell differentiation in the cochlea. *J. Neurosci.* 34, 6470–6479.
- van Es, J.H., van Gijn, M.E., Riccio, O., van den Born, M., Vooijs, M., Begthel, H., Cozijnsen, M., Robine, S., Winton, D.J., Radtke, F., and Clevers, H. (2005). Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* 435, 959–963.
- Weil, D., Levy, G., Sahly, I., Levi-Acobas, F., Blanchard, S., El-Amraoui, A., Crozet, F., Philippe, H., Abitbol, M., and Petit, C. (1996). Human myosin VIIA responsible for the Usher 1B syndrome: a predicted membrane-associated motor protein expressed in developing sensory epithelia. *Proc. Natl. Acad. Sci. USA* 93, 3232–3237.
- White, P.M., Doetzlhofer, A., Lee, Y.S., Groves, A.K., and Segil, N. (2006). Mammalian cochlear supporting cells can divide and trans-differentiate into hair cells. *Nature* 441, 984–987.
- Whyte, J.L., Smith, A.A., and Helms, J.A. (2012). Wnt signaling and injury repair. *Cold Spring Harb. Perspect. Biol.* 4, a008078.
- Yamamoto, N., Chang, W., and Kelley, M.W. (2011). Rbpj regulates development of prosensory cells in the mammalian inner ear. *Dev. Biol.* 353, 367–379.
- Yang, Q., Bermingham, N.A., Finegold, M.J., and Zoghbi, H.Y. (2001). Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* 294, 2155–2158.
- 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.

# Temporal Tracking of Microglia Activation in Neurodegeneration at Single-Cell Resolution

Hansruedi Mathys,<sup>1,2</sup> Chinnakkaruppan Adaikkan,<sup>1,2</sup> Fan Gao,<sup>1,2</sup> Jennie Z. Young,<sup>1,2</sup> Elodie Manet,<sup>1,2</sup> Martin Hemberg,<sup>3</sup> Philip L. De Jager,<sup>4,6</sup> Richard M. Ransohoff,<sup>5</sup> Aviv Regev,<sup>6,7,8</sup> and Li-Huei Tsai<sup>1,2,6,9,\*</sup>

<sup>1</sup>Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>2</sup>Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>3</sup>Department of Cellular Genetics, Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton CB10 1SA, UK

<sup>4</sup>Center for Translational and Systems Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY, USA

<sup>5</sup>Biogen, 225 Binney Street, Cambridge, MA 02142, USA

<sup>6</sup>Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

<sup>7</sup>Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02140, USA

<sup>8</sup>Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>9</sup>Lead Contact

\*Correspondence: lhstai@mit.edu

<https://doi.org/10.1016/j.celrep.2017.09.039>

## SUMMARY

Microglia, the tissue-resident macrophages in the brain, are damage sensors that react to nearly any perturbation, including neurodegenerative diseases such as Alzheimer's disease (AD). Here, using single-cell RNA sequencing, we determined the transcriptome of more than 1,600 individual microglia cells isolated from the hippocampus of a mouse model of severe neurodegeneration with AD-like phenotypes and of control mice at multiple time points during progression of neurodegeneration. In this neurodegeneration model, we discovered two molecularly distinct reactive microglia phenotypes that are typified by modules of co-regulated type I and type II interferon response genes, respectively. Furthermore, our work identified previously unobserved heterogeneity in the response of microglia to neurodegeneration, discovered disease stage-specific microglia cell states, revealed the trajectory of cellular reprogramming of microglia in response to neurodegeneration, and uncovered the underlying transcriptional programs.

## INTRODUCTION

Microglia are tissue-resident macrophages in the brain and spinal cord (Prinz and Priller, 2014). They clear apoptotic cells and are involved in both the elimination and maintenance of synapses for proper neural circuit wiring (Aguzzi et al., 2013). As the damage sensors for the CNS, microglia have been found to respond to nearly any CNS perturbation (Fourgeaud et al., 2016). Indeed, a growing body of evidence based on genome-wide association studies, transcriptomic, and epigenomic analyses, as well as experimental evidence in mouse models, implicates immunological mechanisms and their cellular component,

microglia, in the pathogenesis of Alzheimer's disease (AD) (Gjoneska et al., 2015; Mosher and Wyss-Coray, 2014; Neumann and Daly, 2013; Wang et al., 2015; Zhang et al., 2013). In brain tissue taken at autopsy from individuals with AD, microglia surround A $\beta$  plaques, and their altered morphology indicates that these cells are responding to challenge (Bouvier et al., 2016; Heppner et al., 2015).

However, much remains to be learned about the molecular changes underlying the response of microglia in the AD brain. In particular, the signal that triggers the initial microglial response in the brain undergoing neurodegeneration remains to be determined. Genome-wide transcriptional profiling in microglia has revealed widespread changes in gene expression in mouse models of AD (Orre et al., 2014; Wang et al., 2015). However, ensemble-based approaches that measure gene expression from bulk populations of microglia cells in AD brains can only report population averages that may not reflect the response of individual cells or reveal cell subsets. Furthermore, these studies characterized the reactive microglia phenotype at only one time point, late in the progression of neurodegeneration. Therefore, it remains to be determined how the transcriptional programs in microglia cells change over time as they transition from their initial homeostatic state in the healthy brain to the reactive phenotypes seen in the neurodegenerating brain. Because the microglia phenotype may change drastically over the course of neurodegeneration, tracking with a fine temporal resolution is needed to capture the full spectrum of microglia cell states.

Here, we use single-cell RNA sequencing to examine the phenotypic heterogeneity of microglial cells in the healthy brain and in a mouse model of severe neurodegeneration with AD-like phenotypes. We identified multiple disease stage-specific microglia cell states that are almost exclusively observed in the diseased but not in the healthy brain. We delineate early- versus late-response gene modules and find that microglia proliferation is an early response to neuronal insult. We further identify two distinct reactive microglia phenotypes that arise at a later stage of neurodegeneration and are typified by modules of co-regulated type I and type II interferon response genes,

respectively. Our work identifies previously unobserved heterogeneity in the response of microglia to neurodegeneration, discovers microglia cell states, reveals the trajectory of cellular reprogramming of microglia in response to neurodegeneration, and uncovers the underlying transcriptional programs.

## RESULTS

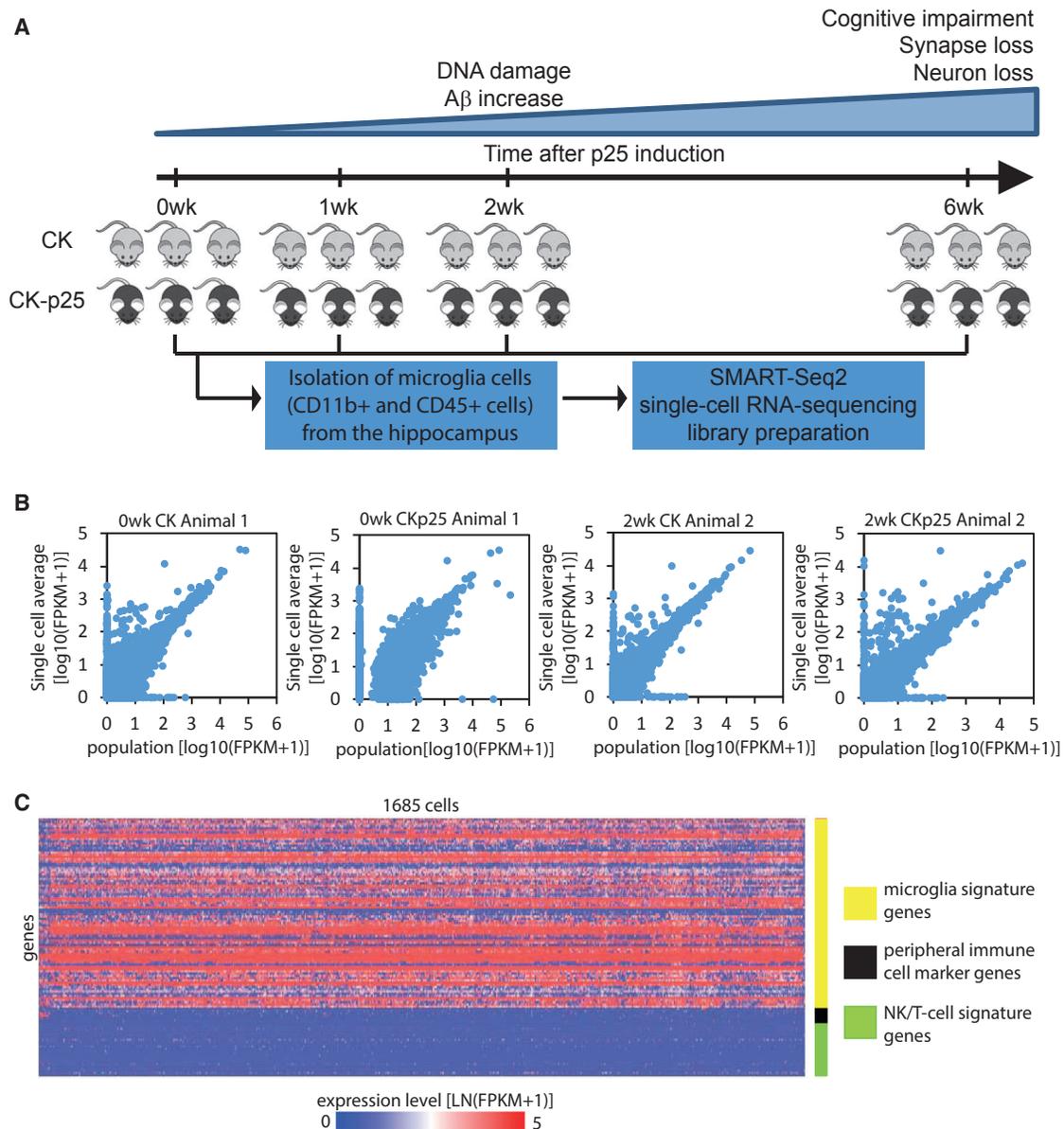
### RNA Sequencing Profiling of Single Microglial Cells Isolated from the Hippocampus

To determine the phenotypic heterogeneity and the transcriptional dynamics of microglia cells during the progression of neurodegeneration, we used the CK-p25 inducible mouse model of severe neurodegeneration. In CK-p25 mice, the expression of p25, the calpain cleavage product of Cdk5 activator p35, is under the control of the *CamKII* promoter and can be switched on by withdrawing doxycycline from the animal's diet (Cruz et al., 2003; Cruz and Tsai, 2004; Fischer et al., 2005; Patrick et al., 1999). Although the CK-p25 model is not based on any genetic mutations associated with familial AD, it develops many aspects of AD-like pathology. A recent study reported that transcriptional profiles from CK-p25 mice and 5XFAD mice, a commonly used AD mouse model, show moderate but similar concordance with human AD brain signatures (Hargis and Blalock, 2017). Neurodegeneration in CK-p25 mice occurs in a temporally compressed but highly predictable manner (Cruz et al., 2003, 2006; Fischer et al., 2005). At 2 weeks after p25 induction, CK-p25 mice exhibit DNA damage and increased amyloid- $\beta$  levels, followed by progressive neuronal and synaptic loss with cognitive impairment, which is severe by 6 weeks of p25 induction (Cruz et al., 2003, 2006; Fischer et al., 2005; Kim et al., 2008). Additional AD-like phenotypes found in 6-week induced CK-p25 mice include forebrain atrophy, astrogliosis, aberrant APP processing, and hyperphosphorylation of tau (Cruz et al., 2003, 2006). CK-p25 mice also display neurofibrillary tangle-like pathology after 27 weeks of p25 induction (Cruz et al., 2003). Importantly, as the p25 transgene is strictly expressed only in excitatory neurons (Figures S1A–S1D), this model permits us to examine the microglial response to neuronal cell death. We isolated cells expressing the microglia markers CD11b and CD45 from the hippocampus of three to four CK-p25 mice and three CK control littermates at each of the following four time points during the progression of neurodegeneration: before p25 induction and 1, 2, and 6 weeks after p25 induction (henceforth abbreviated as 0wk, 1wk, 2wk, and 6wk, respectively) (Figure 1A; Table S1). Individual CD11b-positive and CD45-positive microglial cells were sorted directly into RNA lysis buffer in 96-well plates using fluorescence-activated cell sorting (FACS) (Figure S2). To verify that our sorting protocol resulted in single cells in each well, we mixed microglia cells from two individually distinguishable mouse strains (wild-type and a *Cx3cr1*-knockout strain expressing TdTomato in microglia) and sorted single cells into the wells of a 96-well plate. We then measured the level of *Cx3cr1* and TdTomato mRNA in each well using qPCR. In the large majority of wells, we detected either *Cx3cr1* or TdTomato—but not both—mRNA species, confirming that the large majority of wells contained a single cell (Figures S1E–S1G). We prepared a total of 2,183 single-cell RNA

sequencing libraries using a modified version of the Smart-Seq2 protocol (Picelli et al., 2014) and sequenced the libraries to a depth of 251,353 (median) mapped reads per cell. The average expression profile across single cells and the matching population profile were tightly correlated (Figures 1B and S1H), with an average Pearson product-moment correlation coefficient ( $r = 0.86$ ) comparable with that observed in previous studies (Gaublomme et al., 2015; Shalek et al., 2014). We removed 498 cells on the basis of quality metrics (see Experimental Procedures), retaining 1,685 cells for further analyses. To verify that the cells isolated were indeed microglia cells, we compared the average expression profile of cells isolated from CK control mice with previously published expression profiles of specific brain cell types and with the expression profiles of monocytes and tissue-resident macrophages (Zhang et al., 2014; Lavin et al., 2014). We found that the average expression profile of cells isolated from CK control mice clustered with previously published microglia profiles (Figures S1I and S1J). Comparison with the set of 86 microglia signature genes defined by Butovsky et al. (2014), showed that microglial marker genes (including *Csf1r*, *Tmem119*, *P2ry12*, *Hexb*, and *Sall1*) were expressed at a relatively uniform level in the large majority of cells. In contrast, marker genes of peripheral immune cells (*Cdc20*, *Ccr2*, *Cd163*, and *Ly6c1*), and natural killer cell and T cell signature genes (including *Zap70*, *Skap1*, and *Cd247*) (Bezman et al., 2012) were expressed in only a small subset of the cells (Figures 1C and S1K).

### Non-linear Dimensionality Reduction Reveals Multiple Distinct and Disease Stage-Specific Microglia Cell States

We analyzed 1,685 transcriptomes of single cells expressing the microglia markers CD11b and CD45, isolated from the hippocampus of 0wk, 1wk, 2wk, and 6wk CK-p25 and CK control littermates. Non-linear dimensionality reduction with t-distributed stochastic neighbor embedding (t-SNE) followed by density clustering revealed multiple distinct cell populations (Figure 2A) (van der Maaten and Hinton, 2008). Cluster 2 predominantly contained cells isolated from the CK control and 0wk CK-p25 mice (Figures 2B and 2C). Clusters 3 and 7 mainly contained cells isolated from 1wk CK-p25 mice, and cluster 6 was composed almost exclusively of cells isolated from 2wk and 6wk CK-p25 mice (Figures 2B and 2C). The large majority of cells isolated from the CK control mice at all four time points and from 0wk CK-p25 grouped together in cluster 2 (Figures 2B–2D). Cells isolated from 1wk CK-p25 mice were found in cluster 2 (42%), cluster 3 (44%), and cluster 7 (10%) (Figures 2C and 2D). By contrast, the large majority of cells isolated from 2wk (79%) and 6wk (86%) CK-p25 mice grouped together in cluster 6 (Figures 2C and 2D). We obtained similar results after using a normalization technique to reduce the effects of confounders (Gaublomme et al., 2015) (Figures S3A and S3B), when we used a different RNA sequencing transcript quantification program, RSEM (Figures S3C and S3D), and when we used the GENECODE release M9 gene annotation (data not shown). We saw a similar disease stage-specific response of microglia to neurodegeneration using linear dimensionality reduction (data not shown) and found comparable results using the Single



**Figure 1. Single-Cell RNA Sequencing of Microglia Cells Isolated from the Hippocampus of CK-p25 Mice and CK Control Littermates**

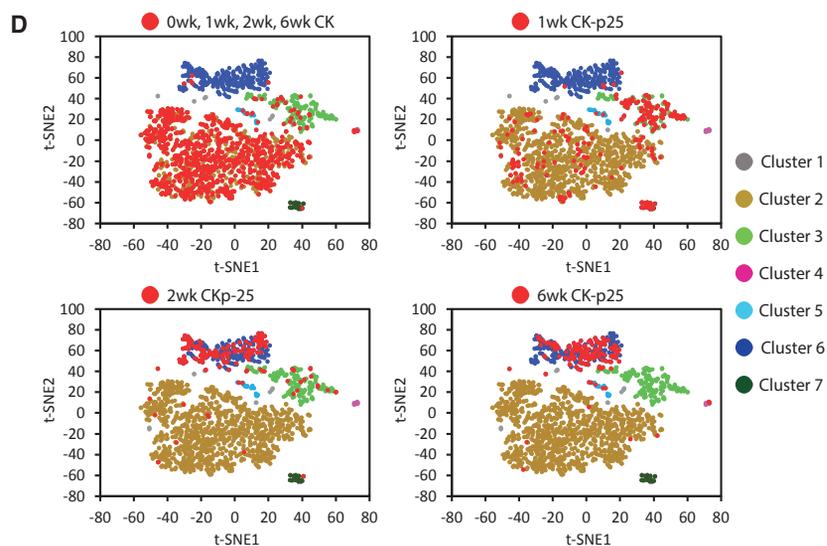
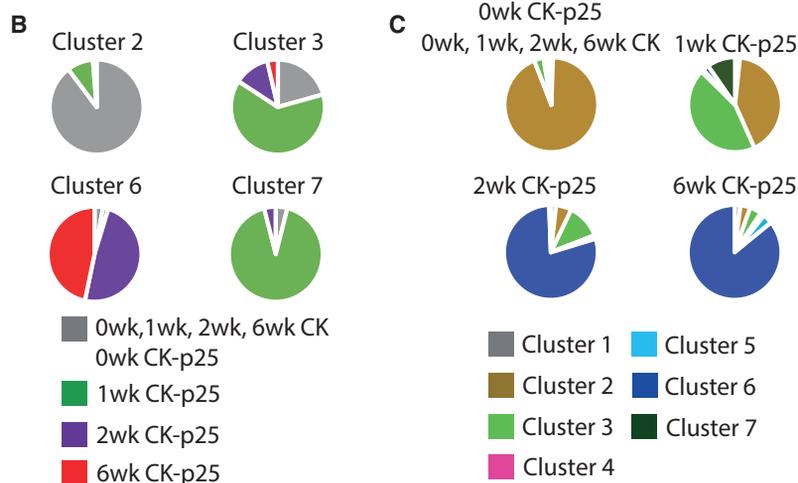
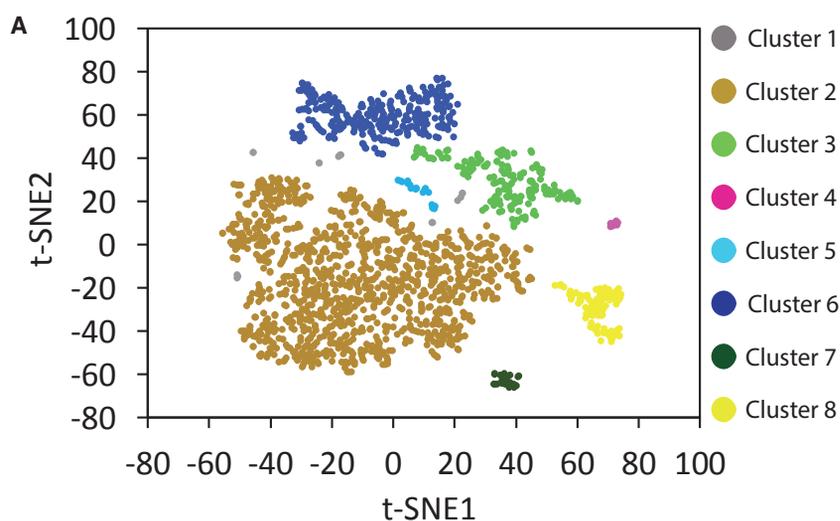
(A) Workflow diagram for single-cell RNA sequencing of microglia cells isolated from the hippocampus of CK-p25 mice and CK control littermates at four different time points after p25 induction.

(B) Quality of single-cell RNA sequencing. Scatterplots compare transcript expression ( $\log_{10}$ [FPKM+1]) between the average of 95 single-cells and a bulk population of 200 cells. The data from four representative animals are shown.

(C) Heatmap showing the expression level of 86 microglia signature genes (yellow), genes preferentially expressed in peripheral immune cells (black), and natural killer cell and T cell signature genes (green) across the 1,685 CD11b- and CD45-positive cells analyzed in this study.

Cell Consensus Clustering (SC3) algorithm as an alternative approach to partitioning cellular subpopulations (Kiselev et al., 2016) (Figures S3E and S3F). Thus, our analysis revealed multiple microglia cell states in the progression of neurodegeneration that clustered separately from most of the cells isolated from the healthy brain (homeostatic microglia): an early-response state composed mainly of cells from 1wk CK-p25 mice and a late-response state composed of the large majority of cells isolated from 2wk and 6wk CK-p25 mice.

The t-SNE analysis also revealed clusters 4 and 5 (Figure 2A), which contained only a small number of cells (9 and 21 cells, respectively) and expressed peripheral immune cell marker genes at higher levels relative to all other clusters, suggesting that they represent bone marrow-derived peripheral immune cells (Table S2). Cluster 8 (Figure 2A) contained cells that came exclusively from one out of three biological replicates (Table S3), indicating that it likely resulted from technical confounders and this cluster was



**Figure 2. Non-linear Dimensionality Reduction Reveals Multiple Distinct and Disease Stage-Specific Microglia Cell States**

(A) Clustering of 1,685 CD11b and CD45 double-positive cells isolated from the hippocampus into eight populations. The t-SNE plot shows a two-dimensional representation of global gene expression profile relationships among 1,685 cells.

(B) Pie charts showing the composition of some of the clusters identified in (A).

(C) Pie charts showing the distribution of each group of cells indicated across the clusters identified in (A) (excluding cluster 8). Cells are grouped by genotype and time point.

(D) t-SNE plots as shown in (A). Cells isolated from CK control mice (at all four time points) and cells from CK-p25 mice (1, 2, and 6 weeks after p25 induction (0wk, 1wk, 2wk, and 6wk, respectively) are highlighted in red in individual panels.

therefore excluded from further analyses (see Experimental Procedures).

### Single-Cell Differential Expression Analysis Reveals Hundreds of Genes Regulated during Cellular Reprogramming of Microglia in Response to Neurodegeneration

We next examined how microglia respond over time in the neurodegenerating brain environment by comparing homeostatic microglia (cluster 2) with early-response state (clusters 3 and 7) versus late-response state microglia (cluster 6). We used the single-cell differential expression (SCDE) software package (Kharchenko et al., 2014) to reveal hundreds of significantly up- and downregulated genes in the early and late-response microglia cell states, compared with homeostatic microglia (Figures S4A–S4C; Table S4). All future references to up- and downregulated gene expression refer to comparison with homeostatic microglia cluster 2. Gene Ontology (GO) term enrichment analysis revealed that cell cycle genes and genes involved in DNA replication and repair were over-represented among the genes upregulated in the early-response cluster 3 (Figures 3A and 3D). Cell cycle-related genes were also found to be upregulated in the early-response cluster 7. The top five GO terms enriched among the genes upregulated in cluster 7 were all related to cell division (Figures 3B and 3E). Plotting the average expression of G1/S and G2/M genes (Tirosh et al., 2016) revealed an approximate circle (Figure S4D), presumably reflecting progression along the cell cycle. Cells from Cluster 3 and Cluster 7 were separated by G1/S and G2/M phase scores, indicating they are in different phases of the cell cycle (Figure S4D). To measure the proliferative activity of microglia, we injected CK and CK-p25 mice with the thymidine analog, 5-ethynyl-2'-deoxyuridine (EdU), every second day during the first 2 weeks after p25 induction. Fluorescent labeling of EdU followed by flow cytometry analysis using the microglia markers CD11b and CD45 revealed that 3% of the microglia from CK control mice had incorporated EdU, in contrast to the significantly larger fraction (22%) seen in CK-p25 mice (Figure S4E). To test whether this increased proliferative activity is reflected in higher microglia density, we performed immunostaining of hippocampal sections and imaged across the major subregions in CK and CK-p25 mice. In 1wk CK-p25 mice, we observed that the density of Iba1-positive microglia was unchanged in the CA1 subregion (Figure S4F) but already significantly increased in CA3 (Figure S4G) and dentate gyrus (Figures S4H and S4K). In 2wk and 6wk CK-p25 mice, microglia density was significantly increased in all three hippocampal subregions examined (Figures S4F–S4K). These findings are consistent with our analysis of single-cell RNA sequencing (Figures 3A, 3B, 3D, and 3E) and EdU incorporation data (Figure S4E), indicating that hippocampal microglia show increased proliferation early in the response to neurodegeneration.

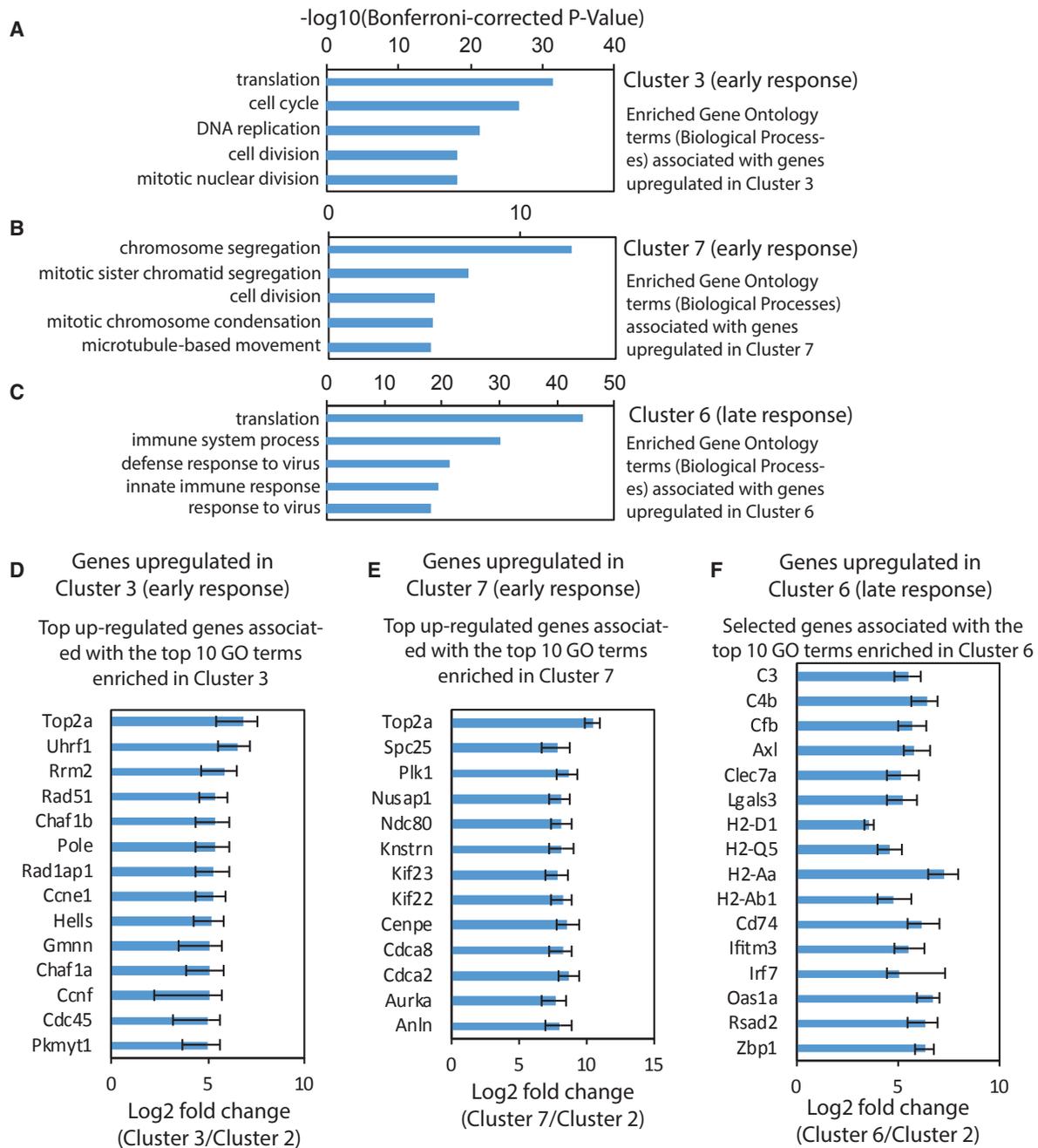
Among the genes upregulated in the late-response cluster 6, we did not see cell cycle genes but rather saw over-representation of immune response-related genes instead (Figure 3C). The top five GO terms included immune system process, defense response to virus, and innate immune response (Figures 3C and 3F). The list of genes included those encoding major histocompatibility complex (MHC) class 1 (H2-D1, H2-Q5) and class II

(H2-Aa, H2-Ab1, Cd74) components. We also identified many interferon response genes (*Irfm3*, *Irf7*) and genes associated with the GO term “defense response to virus” (*Oas1a*, *Rsad2*, *Zbp1*). Binding motifs of the interferon-regulatory factor (IRF) family were significantly enriched within a 1 kb window around the transcription start site of genes upregulated in cluster 6, providing further evidence for an enrichment of interferon response-related genes in late-response state microglia (Table S5). Overall, these data indicate that microglia mount a pronounced immune response during the later stages of neurodegeneration.

### Temporally Distinct Transcriptional Dynamics among Immune Response-Related Genes during Microglia Activation

To further dissect the gene expression changes underlying cellular reprogramming of microglia, we analyzed the pairwise correlations between the expression levels of upregulated genes (see Experimental Procedures). Hierarchical clustering of the resulting correlation matrix of genes upregulated in the early-response cluster 3 cells revealed at least two major modules of co-regulated genes (Figure S5A). GO term enrichment analysis suggested that these two modules contained functionally distinct sets of genes. Whereas genes related to translation, metabolic processes, and immune response were over-represented in module 1, genes involved in cell cycle and DNA metabolic processes were over-represented in module 2 (Figure S5B). A similar separation into two functionally distinct modules was observed for the genes upregulated in the early-response cluster 7 (Figure S5C). Genes involved in cell cycle and mitosis were over-represented in module 1, while genes with a role in the immune response and metabolic processes were over-represented in module 2 (Figure S5D). Modules of co-regulated genes upregulated in the late-response cluster 6 also contained functionally distinct sets of genes (Figure S5E). Whereas genes involved in the immune response were over-represented in module 1, genes related to translation and metabolic processes were over-represented in module 2 (Figure S5F). Thus, this analysis identified functionally distinct modules of co-regulated genes in microglia responding to neurodegeneration. Interestingly, the GO terms enriched among module 1 genes upregulated in cluster 3 were also found to be the most highly enriched among the genes upregulated in cluster 6. Indeed, 199 of the 255 module 1 genes upregulated in cluster 3 were also found to be upregulated in cluster 6, indicating a partial overlap in the gene expression signatures that distinguish cells of the early and the late-response cluster from cells in the healthy brain (data not shown).

To more precisely examine the relationship between the early and late response of microglia cells, we next focused only on cluster 3 cells that were isolated from 1wk CK-p25 mice (Figure 2D). We found that a substantial fraction (191) of the genes significantly upregulated in late-response cluster 6 cells were also significantly upregulated (with a certainty of 95%) in CK-p25 microglia after only 1 week of p25 induction (Figure 4A). Of the two early-response clusters, we focused our analysis on cluster 3 because of relatively higher cell numbers, but a similar overlap in gene expression signatures could be seen in cluster 7, as 38 of the 57 module 2 genes upregulated in Cluster 7 were also found to be upregulated in cluster 6 (data not shown). When we plotted



**Figure 3. Single-Cell Differential Expression Analysis Reveals Hundreds of Genes Regulated during Cellular Reprogramming of Microglia in Response to Neurodegeneration**

(A–C) Bar plots showing the top 5 Gene Ontology (GO) terms (biological processes) associated with genes upregulated in (A) cluster 3 compared with cluster 2, (B) cluster 7 compared with cluster 2, and (C) cluster 6 compared with cluster 2.

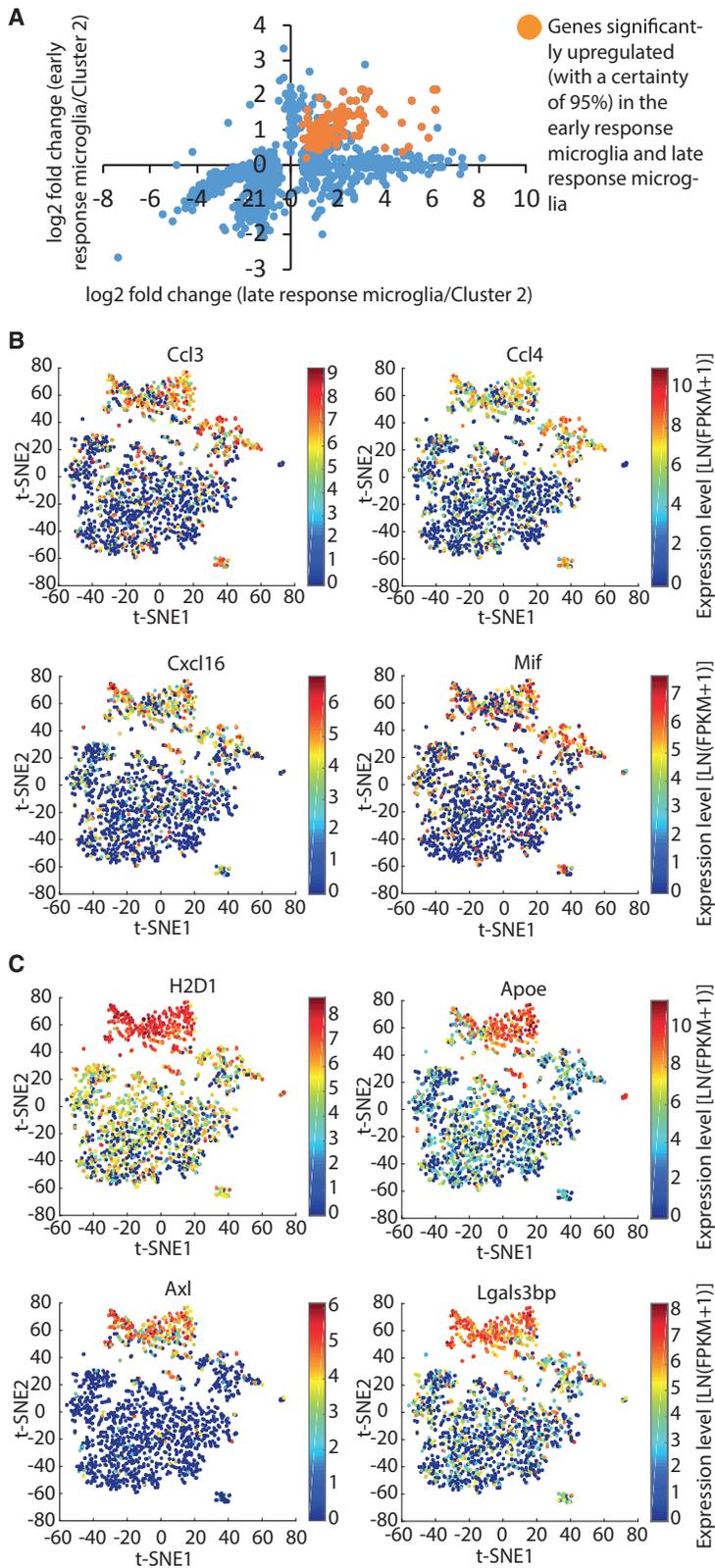
(D and E) Bar graphs showing the fold change in gene expression of the top upregulated genes associated with the top 10 GO terms in cells of cluster 3 (D) and cluster 7 (E) compared with cells of cluster 2.

(F) Bar graph showing the fold change in gene expression of selected genes associated with the top 10 GO terms in cells of cluster 6 compared with cells of cluster 2.

For all panels, error bars show the 95% confidence interval.

the fold change of gene expression in 1wk CK-p25 cluster 3 versus the fold change of gene expression in late-response cluster 6, we found roughly two groups of differentially expressed

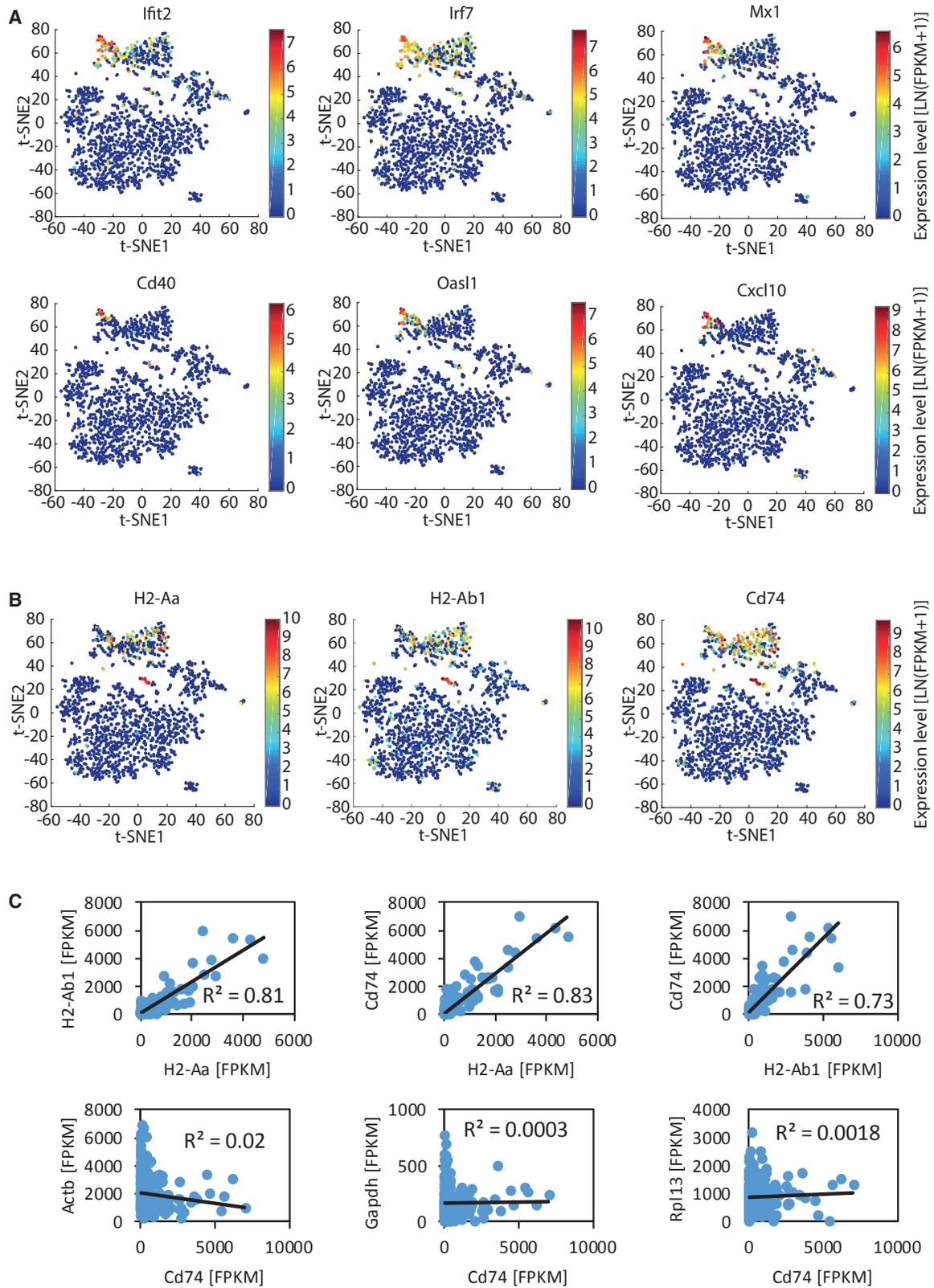
genes: one in which the early and late expression changes appeared to be correlated and a second group that was exclusively upregulated only in late-response microglia but not in



**Figure 4. Temporally Distinct Transcriptional Dynamics among Immune Response-Related Genes during Microglia Activation**

(A) Scatterplot comparing the fold change in gene expression in early- and late-response microglia for genes significantly differentially expressed. The genes significantly upregulated (with a certainty of 95%) in early- and late-response microglia are shown in orange. All other genes are shown in blue. Definition of early-response microglia: cells of cluster 3 that were isolated from CK-p25 mice 1 week after p25 induction (1wk CK-p25). Definition of late-response microglia: cells of cluster 6.

(B and C) t-SNE plots as shown in Figure 2A. Selected genes showing early and consistent upregulation (B) and selected genes exclusively upregulated in the late-response cluster 6 (C) are shown. Data points are colored by the expression levels of the genes indicated.



(legend on next page)

early-response microglia (Figure 4A). Among the set of 191 early and consistently upregulated genes, 38 genes were significantly upregulated with a more stringent cutoff ( $p < 0.001$ ) in cells of the early-response cluster 3 (data not shown). This latter set included genes encoding lysosomal cysteine proteases (*Ctsb*, *Ctsz*), the cysteine protease inhibitor *Cst7*, the inflammatory cytokine *Mif*, chemokines (*Ccl12*, *Ccl3*, *Ccl4*, *Cxcl16*), other immune response-related genes (*Tlr2*, *Lilrb4*), and genes involved in glycolysis (*Pkm*, *Pgk1*, *Gapdh*, *Pgam1*). These genes were also significantly upregulated (with a certainty of 95%) in the microglia of the early-response cluster 7 (Table S4). The set of genes that was exclusively upregulated during the later stages of neurodegeneration included many genes we had identified to be more than 10-fold upregulated in the cells of cluster 6, including complement components (*C3*, *C4b*, and *Cfb*), MHC class I (*H2-D1*, *H2-Q4*, *H2-Q5*) and class II (*H2-Aa*, *H2-Ab1*, *Cd74*) components, genes involved in the interferon response (*Ifitm3*, *Irf7*) and genes associated with the GO term “defense response to virus” (*Oas1a*, *Rsad2*, *Zbp1*). Overall, our analysis revealed a dichotomy in the transcriptional dynamics among genes significantly upregulated in late-response microglia and point to the possibility that the early-response cluster represents a transient, intermediate activation state of microglia.

Next, we explored the transcriptional dynamics of these genes during the time course of neurodegeneration in individual microglia. This analysis provided evidence that a subset of immune response-related genes in late-response cluster 6, including chemokines *Ccl3*, *Ccl4*, and *Cxcl16* and inflammatory cytokine *Mif*, were already upregulated in a subset of microglia 1 week after p25 induction (Figure 4B). In contrast, the expression of other immune-related genes, such as *H2-D1*, *Axl*, *ApoE*, and *Lgals3bp*, was increased in the majority of cells of the late-response cluster 6 but not in 1wk CK-p25 Cluster 3 cells (Figure 4C). In summary, our analysis of transcriptional dynamics during disease progression uncovered temporally distinct subsets of immune response-related genes during microglia activation in a neurodegenerative context.

### Heterogeneous Late Response of Microglia to Neurodegeneration Is Typified by Different Modules of Co-regulated Genes

We next focused solely on the transcriptional dynamics in late-response microglia and discovered small subsets of single cells characterized by upregulation of specific gene sets, such as antiviral and interferon response genes (Figure 5A) and components of the MHC class II pathway (Figure 5B). To quantify the activation of these two gene sets within individual cells, we computed a module score as previously described (Shalek et al., 2014; see Experimental Procedures). We computed the module scores for a set of 132 co-regulated antiviral and interferon response genes. Most 2wk and 6wk CK-p25 microglia

showed activation of the antiviral and interferon response gene expression module (Figures 6A and 6C). However, there was substantial variation across individual cells, with a small subset of cells exhibiting a fold induction module score that was at least an order of magnitude higher than the average (Figures 6A and 6C). We also computed the module scores for the four genes encoding components of MHC class II. Whereas some 2wk and 6wk CK-p25 microglia showed no detectable activation of the MHC class II module, other cells exhibited a fold induction value at least an order of magnitude higher than the average (Figures 6B and 6D). In contrast to the variability seen in antiviral and interferon response genes and MHC class II components, the induction scores of another module of co-regulated genes, containing mainly ribosomal protein-encoding genes exhibited a much narrower distribution (Figure 6E). Moreover, for the MHC class II components, although the expression of these genes was highly variable between cells, it was tightly correlated within individual cells and the expression of the MHC class II-related gene *CD74* was not correlated with the expression of the housekeeping genes *Actb*, *Gapdh*, and *Rpl13* (Figure 5C). These observations indicate that the extensive variability in the expression of MHC class II components measured across cells reflects true biological differences, rather than technical noise inherent to single-cell RNA sequencing experiments. There was no correlation observed between the induction of the antiviral and interferon response module and the MHC class II module, indicating that the cells expressing high levels of antiviral and interferon response genes and the cells expressing high levels of MHC class II genes are not necessarily identical (Figure 6F). Thus, these data indicate that there are at least two distinct reactive microglia phenotypes in neurodegeneration.

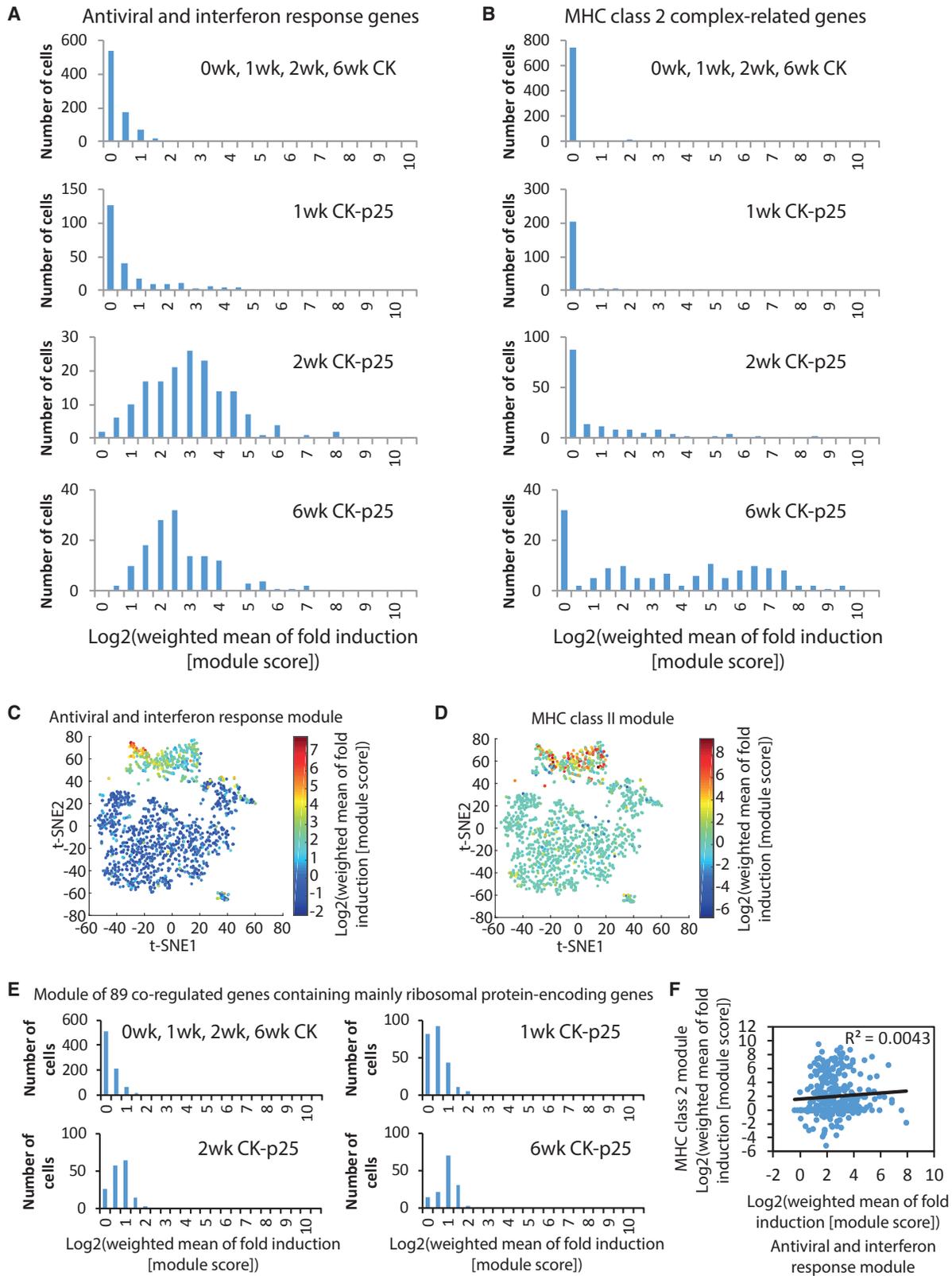
To corroborate the observed heterogeneity among microglia with an alternative approach, we performed immunostaining of hippocampal sections from CK and CK-p25 mice. To examine the expression of co-regulated antiviral and interferon response genes, we stained the sections with antibodies recognizing the protein products of *Cd40* and *Cd69* and for the microglia marker *Iba1*. We observed an increase in the number of *CD40*-expressing cells in all three hippocampal subregions of 6wk CK-p25 mice compared with CK control mice (dentate gyrus: Figures 7A, 7C, and S6A; CA1 and CA3: Figure S6B). Importantly, the immunostaining revealed that only a subset of the *Iba1*-positive cells detectably express *CD40* (dentate gyrus: Figures 7A, 7C, and S6A; CA1 and CA3: Figure S6B). Similarly, we found an increased number of *CD69*-positive cells in all three hippocampal subregions of CK-p25 mice 6 weeks after p25 induction compared with CK control mice (dentate gyrus: Figures 7B, 7D, and S6C; CA1 and CA3: Figure S6D). However, again, only a subset of *Iba1*-positive cells detectably expressed *CD69* in CK-p25 mice (dentate gyrus: Figures 7D and S6C; CA1 and

**Figure 5. Heterogeneous Late Response to Neurodegeneration of Microglia Is Typified by Different Modules of Co-regulated Genes**  
t-SNE plots as shown in Figure 2A.

(A) Data points are colored by the expression level of selected antiviral and interferon response genes as indicated.

(B) Data points are colored by the expression level of the genes encoding MHC class II components as indicated.

(C) Scatterplots showing the correlation of the expression level of MHC class II related genes (*H2-Aa*, *H2-Ab1*, *Cd74*) and of *Cd74* and the housekeeping genes *Actb*, *Gapdh*, and *Rpl13* across the cells of cluster 6.



(legend on next page)

CA3: Figure S6D). We next stained the sections with an antibody recognizing the MHC class II component CD74. In the hippocampus of CK control mice, only very few cells expressing the MHC class II component CD74 were detected (dentate gyrus: Figures 7E, 7F, and S6E; CA1 and CA3: Figure S6F). In contrast, the number of CD74-positive cells was dramatically increased in the hippocampus of CK-p25 mice 6 weeks after p25 induction (dentate gyrus: Figures 7E, 7F, and S6E; CA1 and CA3: Figure S6F). Importantly, the overall proportion of CD74-expressing microglia was limited to only a subset of cells, confirming the heterogeneity of the hippocampal microglia population in the response to neurodegeneration (dentate gyrus: Figures 7E, 7F, and S6E; CA1 and CA3: Figure S6F). Taken together, our immunostaining data show that the expression of two components of a module of co-regulated antiviral and interferon response genes (CD69 and CD40) as well as the MHC class II component CD74 are strongly upregulated in subsets of microglia in the hippocampus of CK-p25. Thus, our immunostaining data confirm a previously unobserved heterogeneity in the response of microglia to neurodegeneration.

#### Gene Sets Induced in the CK-p25 Mouse Model of Neurodegeneration Also Tend to Be Upregulated in Microglia of Aged Human Individuals

Finally, we wondered whether the gene sets induced in the CK-p25 mouse model of neurodegeneration might also be relevant for the biology of human microglia. To this end, we compared our data with a recently published transcriptomic analysis of purified human cortical microglia (Galatro et al., 2017). Specifically, we asked whether the gene sets induced in the CK-p25 mouse model of neurodegeneration also tend to be upregulated in microglia of aged human individuals compared with microglia of younger individuals. Indeed, we observed that the expression of gene sets that are human orthologs of the genes significantly upregulated in clusters 3, 6, and 7 was positively correlated with age (Figures S7A, S7C, and S7E). In contrast, the expression of genes downregulated in clusters 3, 6, and 7 was not significantly correlated with age (Figures S7B, S7D, and S7F). We also found that the expression of human orthologs of the modules of co-regulated antiviral and interferon response genes and MHC class II components was positively correlated with age (Figures S7G and S7H). Specifically, the antiviral and interferon response genes *Isg15*, *Oasl*, *Ifitm3*, *Irf7*, and *Ifi30* were found among the genes most significantly correlated with age in human microglia (Table S6). Thus, these data indicate that the gene sets induced in the CK-p25 mouse model of neurodegeneration also tend to be upregulated in microglia of aged human individuals.

## DISCUSSION

In this study we comprehensively surveyed the transcriptome of microglia cells as they progressed from healthy to neurodegeneration states in the brain at single-cell resolution. We found a remarkable phenotypic heterogeneity of microglia: from early-response states, characterized by marked proliferation, to late-response states of mounting immune response. In this latter group, we discovered a further heterogeneity that was typified by functionally distinct modules of co-regulated genes.

#### Temporal Changes of Microglia Response to Neurodegeneration

We were interested to decipher the initial response of microglia to insult and to track the temporal changes of these cells over the course of neurodegeneration. We used the CK-p25 mouse model of severe neurodegeneration, as it displays key pathological hallmarks of AD in a temporally predictable manner, to study the response of microglia at fine temporal- and single-cell resolution. Our single-cell RNA sequencing analysis identified two neurodegeneration-associated microglia cell states that are distinct from the microglia state in the healthy brain. Importantly, we found that microglia isolated at an early stage of neurodegeneration (early-response microglia cell state) are distinct from microglia isolated at a late stage of neurodegeneration (late-response microglia cell state). However, the early- and late-response microglia cell states are not completely unrelated, as we found that a substantial fraction of genes that are upregulated in late-response microglia are already increased in early-response microglia. Moreover, 1wk CK-p25 cells were distributed across homeostatic microglia cluster 2 and early-response cluster 3, and share characteristics with both microglia in the healthy brain and fully activated microglia at the later stages of neurodegeneration. These observations suggest that the early-response microglia cell state represents a transient intermediate activation state that constitutes part of the trajectory of cellular reprogramming of homeostatic microglia in response to neurodegeneration.

Although we cannot exclude the possibility that some of the cells analyzed are peripheral immune cells, our data suggest that the large majority of cells examined are microglia cells. First, the large majority of cells analyzed expressed microglia marker genes but not peripheral immune cells signature genes (Figures 1C and S1K). Second, only a very small fraction of cells expressed high levels of CD45, a characteristic of peripheral immune cells (Figure S2). Third, t-SNE analysis revealed two small clusters (clusters 4 and 5) that expressed peripheral immune cell marker genes, showing that our methods have sufficient power

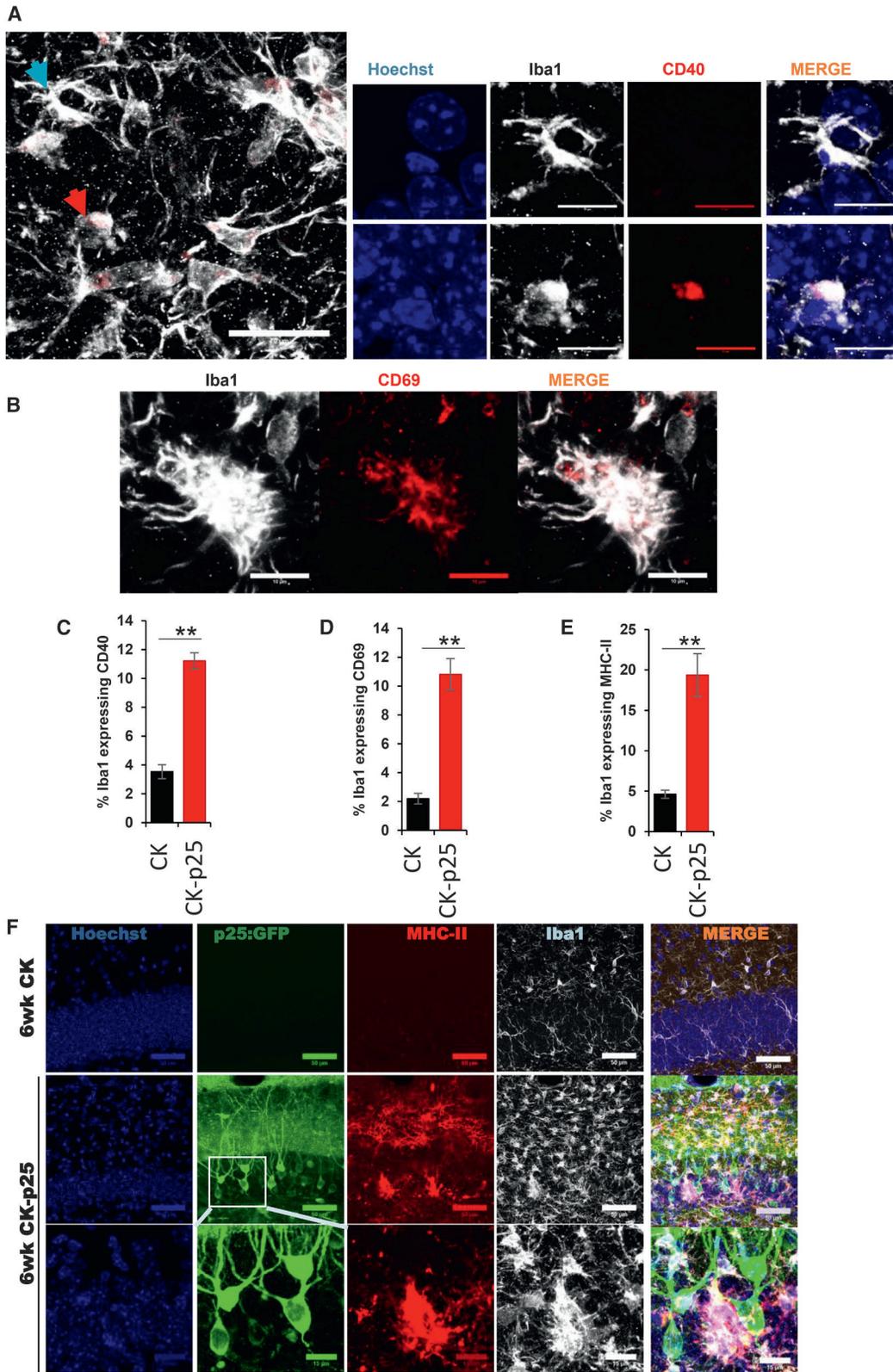
#### Figure 6. Two Distinct Reactive Microglia Phenotypes in Late Response to Neurodegeneration

(A and B) Histograms showing the distribution of the weighted fold induction of a module of 132 antiviral and interferon response genes (A) and four MHC class II complex-related genes (B) across the cells of the groups indicated.

(C and D) t-SNE plots as shown in Figure 2A. Data points are colored by the weighted fold induction of (C) a module of 132 antiviral and interferon response genes and (D) a module of four MHC class II complex-related genes.

(E) Histograms showing the distribution of the weighted fold induction of a module of co-regulated genes mainly containing ribosomal protein-encoding genes across the cells of the groups indicated.

(F) Scatterplot showing the correlation between the induction of the antiviral and interferon response module and the MHC class II module across the cells isolated from CK-p25 mice 2 and 6 weeks after p25 induction.



(legend on next page)

to detect the subtle differences in gene expression between microglia and peripheral immune cells (Figure 2A).

### Microglia Proliferate and Increase the Expression of a Small Set of Cytokines at an Early Stage of Neurodegeneration

Our previous study had revealed an increased level of transcripts encoding immune response-related genes in hippocampal tissue of CK-p25 mice (Gjoneska et al., 2015). However, it was unclear whether this increase was due to gene expression changes in microglia or to an expansion of the microglia population as a result of proliferation. Our data clearly demonstrate that microglia respond to neurodegeneration with both a dramatic reprogramming of their transcriptome and increased proliferation. We identified a small set of cytokines, including the pro-inflammatory cytokines MIF and TNF, that are upregulated in the late-response microglia cell state but already show upregulation 1 week after p25 induction. Interestingly, the pro-inflammatory cytokine TNF has recently been identified as one of the factors that are necessary and sufficient to induce a neurotoxic astrocyte cell state (Liddelow et al., 2017). This raises the possibility that early-response microglia may contribute to the initiation of a cascade that ultimately leads to neuronal death during the later stages of neurodegeneration. Given the pleiotropic effects of TNF reported (Probert, 2015), it is possible that increased expression of TNF in microglia serves multiple distinct functions that remain to be explored.

### Two Distinct Reactive Microglia Phenotypes in Neurodegeneration

We discovered small subsets of microglia in the hippocampus of CK-p25 mice late in neurodegeneration typified by strong upregulated expression of selected gene sets, such as antiviral and interferon response genes and MHC class II components. Interestingly, these two subsets do not overlap entirely, indicating that not only are there at least two distinct reactive microglia phenotypes in the heterogeneous cell population but that there are likely multiple triggers of microglia activation. Interestingly, we detected many genes that are induced by type I interferon among the module of co-regulated antiviral and interferon response genes. In contrast, MHC class II genes have been shown to be induced by interferon-gamma (type II interferon) (O'Keefe et al., 2001). This raises the intriguing possibility that the two distinct subsets of reactive microglia we observed in CK-p25 mice could reflect exposure to type I interferon versus type II interferon.

What might be the trigger to induce expression of the co-regulated antiviral and interferon response gene module? We found

that a subset of microglia with activated type I interferon pathway are in close proximity to neurons harboring DNA damage (Figures S6G and S6H). It is well established that DNA damage can lead to the induction of type I interferons (Härtlova et al., 2015). Thus, DNA damage might be one factor that contributes to the induction of antiviral and interferon response genes in microglia. Consistent with this idea, we found that the module of co-regulated antiviral and interferon response genes was significantly upregulated in microglia of the DNA-repair deficient mouse model *Ercc1Δ/-* compared with wild-type control mice (Figures S7I and S7J).

In the peripheral immune system, MHC class II molecules are responsible for presenting peptides derived from extracellular pathogens to T cells and are expressed by a subset of antigen-presenting cells, such as macrophages, dendritic cells, and B cells (Ting and Trowsdale, 2002). Whether the subset of microglia expressing high levels of MHC class II molecules indeed functions as antigen-presenting cells is an intriguing hypothesis that remains to be addressed. Alternatively, MHC class II molecules could serve a T cell-independent function in microglia. Supporting this idea, it has been shown that the presence of MHC class II molecules exacerbated both neurodegenerative symptoms and neuropathology in murine globoid leukodystrophy, most likely in a T cell-independent manner (Matsushima et al., 1994).

Recently, Keren-Shaul et al. (2017) reported the identification of a microglia type associated with neurodegenerative diseases, termed disease-associated microglia (DAM). The authors carried out single-cell RNA sequencing on microglia isolated from a mouse model of AD that expresses five human familial AD gene mutations (5XFAD) and from a mouse model of amyotrophic lateral sclerosis (*mSOD1* [G93A] mice). The DAM microglia phenotype was observed only in the 5XFAD and *mSOD1* (G93A) mice but not in the corresponding wild-type mice. Using a different mouse model of neurodegeneration (CK-p25), we identified a microglia cell state that was almost exclusively observed at the later stages of disease progression (late-response microglia). Interestingly, late-response microglia express increased levels of many genes that were also observed to be upregulated in DAM (e.g., *Cd9*, *Itgax*, *Clec7a*, *Cd63*, *Spp1*, *Fth1*, *Axl*, *Lpl*, *Cst7*, *Ctsb*, *Apoe*). Of the 278 genes significantly upregulated in DAM, 202 genes were also significantly upregulated in late-response microglia, suggesting a substantial similarity between the expression profiles of DAM and late-response microglia. This observation is consistent with the idea that the DAM program may be a primed set of genes that is expressed in response to varied conditions of altered homeostasis (Keren-Shaul et al.,

### Figure 7. Immunostaining Reveals Heterogeneous Late Response of Microglia to Neurodegeneration

Immunostaining in the dentate gyrus of CK-p25 mice 6 weeks after p25 induction.

(A) Immunohistochemistry with anti-CD40 (red) and anti-Iba1 (white) antibodies. Cells indicated with blue and red arrows are shown at higher magnification on the right.

(B) Immunohistochemistry with anti-CD69 (red) and anti-Iba1 (white) antibodies.

(C–E) Quantification of the CD40 (C), CD69 (D), MHC2 (E), and Iba1 immunostaining. Values are percentages of Iba1-positive cells expressing CD40, CD69, and MHC2, respectively.

(F) Immunohistochemistry with anti-GFP (green), anti-CD74 (MHC2, red), and anti-Iba1 (white) antibodies. For all graphs, quantification is based on immunostaining in the dentate gyrus of seven CK mice and four CK-p25 mice, with two sections per animal.

Error bars show SEM. \*\*p < 0.01; \*\*\*p < 0.001.

2017). However, we found that many antiviral and interferon response genes were significantly upregulated in late-response microglia but not in DAM, suggesting that there may also be potential differences between these two microglia cell states. Keren-Shaul et al. (2017) proposed that DAM are generated through a two-step activation process. Homeostatic microglia first transition to an intermediate stage (stage 1 DAM) in a Trem2-independent manner, followed by a second, Trem2-dependent, transition to stage 2 DAM. We saw significant differences in the expression of both stage 1 DAM and stage 2 DAM characteristic genes in late-response microglia but much less so in early-response microglia (Figure S7K). This observation may indicate that early-response microglia are in an even more naive activation state than stage 1 DAM. Whether the induction of a DAM-like expression program in microglia of the CK-p25 model of neurodegeneration is protective, neutral, or deleterious remains to be determined.

In summary, our work identified previously unobserved heterogeneity in the response of microglia to neurodegeneration, including the discovery of microglia cell states, and uncovered transcriptional programs underlying the trajectory of cellular reprogramming of microglia in response to neurodegeneration. Our analysis also indicates that many of the gene expression regulatory events identified in our mouse model are conserved in aged human microglia. Thus, these insights into the molecular programs underlying microglia activation may pave the way for designing rational and efficient strategies to treat AD and other neurodegenerative diseases.

## EXPERIMENTAL PROCEDURES

### Animals

All animal work was approved by the Committee for Animal Care of the Division of Comparative Medicine at the Massachusetts Institute of Technology.

### Isolation of Microglia from the Hippocampus

Hippocampal tissue was enzymatically digested using the Neural Tissue Dissociation Kit (P) (catalog number 130-092-628; Miltenyi Biotec) according to the manufacturer's protocol, with minor modifications. The resulting cell suspension was then stained using allophycocyanin (APC)-conjugated CD11b mouse clone M1/70.15.11.5 (130-098-088; Miltenyi Biotec) and phycoerythrin (PE)-conjugated CD45 antibody (553081; BD Pharmingen) according to the manufacturer's (Miltenyi Biotec) recommendations. FACS was then used to purify CD11b and CD45 positive microglial cells.

### Single-Cell RNA Sequencing Library Preparation

Single-cell RNA sequencing libraries were generated on the basis of the SMART-Seq2 protocol (Picelli et al., 2014) with minor modifications. Libraries were tagged using the Nextera XT DNA Library Preparation Kit (catalog number FC-131-1096; Illumina) and the Nextera XT Index Kit version 2 Sets A, B, C, and D according to the manufacturer's instructions with minor modifications. Specifically, reactions were run at one fourth the recommended volume, the tagging step was extended to 10 min, and the extension time during the PCR step was increased from 30 s to 60 s.

### Click-iT Plus EdU Cell Proliferation Assay

Animals were intraperitoneally injected with 50 mg/kg EdU every second day during the first 2 weeks after p25 induction. Two weeks after p25 induction, the animals were transcardially perfused with ice-cold PBS, and microglia were isolated using the Neural Tissue Dissociation Kit (P) as described above. EdU incorporation was then detected using the Click-iT Plus EdU Pacific Blue

Flow Cytometry Assay Kit (catalog number C10636; Thermo Fisher Scientific) according to the manufacturer's instructions.

### Immunohistochemistry

Immunohistochemistry experiments were performed as described previously (Iaccarino et al., 2016).

### DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA sequencing data reported in this paper is GEO: GSE103334.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and six tables and can be found with this article online at <https://doi.org/10.1016/j.celrep.2017.09.039>.

### AUTHOR CONTRIBUTIONS

H.M., C.A., P.L.D.J., R.M.R., A.R., and L.-H.T. designed experiments. H.M. performed the majority of experiments with help from E.M. C.A. performed the immunohistochemistry experiments. H.M. and F.G. performed the bioinformatics analysis with help from M.H. H.M., J.Z.Y., R.M.R., and L.-H.T. wrote the manuscript.

### ACKNOWLEDGMENTS

We are grateful to E. Demmons and J.J. Trombetta for technical assistance. We thank members of the Tsai laboratory, J. Penney, H. Meharena, W. Ralvenius, and E. Gjoneska, for discussions and comments on the paper. We thank P. Wisniewski, C. Zollo, C. Johns, and P. Autissier of the Whitehead flow cytometry core facility for assistance with FACS. This work is supported in part by the JBP Foundation and NIH grant RF1 AG054321 to L.-H.T. H.M. was supported by an Early Postdoc Mobility fellowship from the Swiss National Science Foundation (P2BSP3\_151885).

Received: February 15, 2017

Revised: August 8, 2017

Accepted: September 11, 2017

Published: October 10, 2017

### REFERENCES

- Aguzzi, A., Barres, B.A., and Bennett, M.L. (2013). Microglia: scapegoat, saboteur, or something else? *Science* 339, 156–161.
- Bezman, N.A., Kim, C.C., Sun, J.C., Min-Oo, G., Hendricks, D.W., Kamimura, Y., Best, J.A., Goldrath, A.W., and Lanier, L.L.; Immunological Genome Project Consortium (2012). Molecular definition of the identity and activation of natural killer cells. *Nat. Immunol.* 13, 1000–1009.
- Bouvier, D.S., Jones, E.V., Quesseveur, G., Davoli, M.A., A Ferreira, T., Quirion, R., Mechawar, N., and Murai, K.K. (2016). High resolution dissection of reactive glial nets in Alzheimer's disease. *Sci. Rep.* 6, 24544.
- Butovsky, O., Jedrychowski, M.P., Moore, C.S., Cialic, R., Lanser, A.J., Gabriely, G., Koeglsperger, T., Dake, B., Wu, P.M., Doykan, C.E., et al. (2014). Identification of a unique TGF- $\beta$ -dependent molecular and functional signature in microglia. *Nat. Neurosci.* 17, 131–143.
- Cruz, J.C., and Tsai, L.-H. (2004). A Jekyll and Hyde kinase: roles for Cdk5 in brain development and disease. *Curr. Opin. Neurobiol.* 14, 390–394.
- Cruz, J.C., Tseng, H.C., Goldman, J.A., Shih, H., and Tsai, L.H. (2003). Aberrant Cdk5 activation by p25 triggers pathological events leading to neurodegeneration and neurofibrillary tangles. *Neuron* 40, 471–483.
- Cruz, J.C., Kim, D., Moy, L.Y., Dobbin, M.M., Sun, X., Bronson, R.T., and Tsai, L.H. (2006). p25/cyclin-dependent kinase 5 induces production and intraneuronal accumulation of amyloid beta in vivo. *J. Neurosci.* 26, 10536–10541.

- Fischer, A., Sananbenesi, F., Pang, P.T., Lu, B., and Tsai, L.H. (2005). Opposing roles of transient and prolonged expression of p25 in synaptic plasticity and hippocampus-dependent memory. *Neuron* 48, 825–838.
- Fourgeaud, L., Través, P.G., Tufail, Y., Leal-Bailey, H., Lew, E.D., Burrola, P.G., Callaway, P., Zagórska, A., Rothlin, C.V., Nimmerjahn, A., and Lemke, G. (2016). TAM receptors regulate multiple features of microglial physiology. *Nature* 532, 240–244.
- Galatro, T.F., Holtman, I.R., Lerario, A.M., Vainchtein, I.D., Brouwer, N., Sola, P.R., Veras, M.M., Pereira, T.F., Leite, R.E.P., Möller, T., et al. (2017). Transcriptomic analysis of purified human cortical microglia reveals age-associated changes. *Nat. Neurosci.* 20, 1162–1171.
- Gaublomme, J.T., Yosef, N., Lee, Y., Gertner, R.S., Yang, L.V., Wu, C., Pandolfi, P.P., Mak, T., Satija, R., Shalek, A.K., et al. (2015). Single-cell genomics unveils critical regulators of Th17 cell pathogenicity. *Cell* 163, 1400–1412.
- Gjoneska, E., Pfenning, A.R., Mathys, H., Quon, G., Kundaje, A., Tsai, L.H., and Kellis, M. (2015). Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* 518, 365–369.
- Hargis, K., and Blalock, E. (2017). Transcriptional signatures of brain aging and Alzheimer's disease: What are our rodent models telling us? *Behav Brain Res.* 322, 311–328.
- Härtlova, A., Ertmann, S.F., Raffi, F.A., Schmalz, A.M., Resch, U., Anugula, S., Lienenklaus, S., Nilsson, L.M., Kröger, A., Nilsson, J.A., et al. (2015). DNA damage primes the type I interferon system via the cytosolic DNA sensor STING to promote anti-microbial innate immunity. *Immunity* 42, 332–343.
- Heppner, F.L., Ransohoff, R.M., and Becher, B. (2015). Immune attack: the role of inflammation in Alzheimer disease. *Nat. Rev. Neurosci.* 16, 358–372.
- Iaccarino, H.F., Singer, A.C., Martorell, A.J., Rudenko, A., Gao, F., Gillingham, T.Z., Mathys, H., Seo, J., Kritskiy, O., Abdurrob, F., et al. (2016). Gamma frequency entrainment attenuates amyloid load and modifies microglia. *Nature* 540, 230–235.
- Keren-Shaul, H., Spinrad, A., Weiner, A., Matcovitch-Natan, O., Dvir-Szternfeld, R., Ulland, T.K., David, E., Baruch, K., Lara-Astaiso, D., Toth, B., et al. (2017). A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* 169, 1276–1290.e17.
- Kharchenko, P.V., Silberstein, L., and Scadden, D.T. (2014). Bayesian approach to single-cell differential expression analysis. *Nat. Methods* 11, 740–742.
- Kim, D., Frank, C.L., Dobbin, M.M., Tsunemoto, R.K., Tu, W., Peng, P.L., Guan, J.S., Lee, B.H., Moy, L.Y., Giusti, P., et al. (2008). Deregulation of HDAC1 by p25/Cdk5 in neurotoxicity. *Neuron* 60, 803–817.
- Kiselev, V.Y., Kirschner, K., Schaub, M.T., Andrews, T., Yiu, A., Chandra, T., Natarajan, K.N., Reik, W., Barahona, M., Green, A.R., Hemberg, M., et al. (2016). SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods* 14, 483–486.
- Lavin, Y., Winter, D., Blecher-Gonen, R., David, E., Keren-Shaul, H., Merad, M., Jung, S., and Amit, I. (2014). Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. *Cell* 159, 1312–1326.
- LiddeLOW, S.A., Guttenplan, K.A., Clarke, L.E., Bennett, F.C., Bohlen, C.J., Schirmer, L., Bennett, M.L., Münch, A.E., Chung, W.S., Peterson, T.C., et al. (2017). Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* 541, 481–487.
- Matsushima, G.K., Taniike, M., Glimcher, L.H., Grusby, M.J., Frelinger, J.A., Suzuki, K., and Ting, J.P. (1994). Absence of MHC class II molecules reduces CNS demyelination, microglial/macrophage infiltration, and twitching in murine globoid cell leukodystrophy. *Cell* 78, 645–656.
- Mosher, K.I., and Wyss-Coray, T. (2014). Microglial dysfunction in brain aging and Alzheimer's disease. *Biochem. Pharmacol.* 88, 594–604.
- Neumann, H., and Daly, M.J. (2013). Variant *TREM2* as risk factor for Alzheimer's disease. *N. Engl. J. Med.* 368, 182–184.
- O'Keefe, G.M., Nguyen, V.T., Ping Tang, L.L., and Benveniste, E.N. (2001). IFN-gamma regulation of class II transactivator promoter IV in macrophages and microglia: involvement of the suppressors of cytokine signaling-1 protein. *J. Immunol.* 166, 2260–2269.
- Orre, M., Kamphuis, W., Osborn, L.M., Jansen, A.H., 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.
- Patrick, G.N., Zukerberg, L., Nikolic, M., de la Monte, S., Dikkes, P., and Tsai, L.H. (1999). Conversion of p35 to p25 deregulates Cdk5 activity and promotes neurodegeneration. *Nature* 402, 615–622.
- Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9, 171–181.
- Prinz, M., and Priller, J. (2014). Microglia and brain macrophages in the molecular age: from origin to neuropsychiatric disease. *Nat. Rev. Neurosci.* 15, 300–312.
- Probert, L. (2015). TNF and its receptors in the CNS: the essential, the desirable and the deleterious effects. *Neuroscience* 302, 2–22.
- Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaublomme, J.T., Yosef, N., et al. (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 510, 363–369.
- Ting, J.P.-Y., and Trowsdale, J. (2002). Genetic control of MHC class II expression. *Cell* 109 (Suppl), S21–S33.
- Tirosh, I., Venteicher, A.S., Hebert, C., Escalante, L.E., Patel, A.P., Yizhak, K., Fisher, J.M., Rodman, C., Mount, C., Filbin, M.G., et al. (2016). Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma. *Nature* 539, 309–313.
- van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.
- 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.
- Zhang, B., Gaiteri, C., Bodea, L.G., Wang, Z., McElwee, J., Podtelezchnikov, 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.
- Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'Keefe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* 34, 11929–11947.

# $\beta$ -Hydroxybutyrate Deactivates Neutrophil NLRP3 Inflammasome to Relieve Gout Flares

Emily L. Goldberg,<sup>1,2</sup> Jennifer L. Asher,<sup>1</sup> Ryan D. Molony,<sup>2</sup> Albert C. Shaw,<sup>3</sup> Caroline J. Zeiss,<sup>1</sup> Chao Wang,<sup>4</sup> Ludmilla A. Morozova-Roche,<sup>4</sup> Raimund I. Herzog,<sup>5</sup> Akiko Iwasaki,<sup>2,6</sup> and Vishwa Deep Dixit<sup>1,2,7,8,\*</sup>

<sup>1</sup>Section of Comparative Medicine, Yale School of Medicine, New Haven, CT 06520, USA

<sup>2</sup>Department of Immunobiology, Yale School of Medicine, New Haven, CT 06520, USA

<sup>3</sup>Section of Infectious Diseases, Yale School of Medicine, New Haven, CT 06520, USA

<sup>4</sup>Department of Medical Biochemistry and Biophysics, Umeå University, 901 87 Umeå, Sweden

<sup>5</sup>Section of Endocrinology and Metabolism, Yale School of Medicine, New Haven, CT 06520, USA

<sup>6</sup>Howard Hughes Medical Institute, Chevy Chase, MD 20815-6789, USA

<sup>7</sup>Yale Center for Research on Aging, Yale School of Medicine, New Haven, CT 06520, USA

<sup>8</sup>Lead Contact

\*Correspondence: vishwa.dixit@yale.edu

<http://dx.doi.org/10.1016/j.celrep.2017.02.004>

## SUMMARY

Aging and lipotoxicity are two major risk factors for gout that are linked by the activation of the NLRP3 inflammasome. Neutrophil-mediated production of interleukin-1 $\beta$  (IL-1 $\beta$ ) drives gouty flares that cause joint destruction, intense pain, and fever. However, metabolites that impact neutrophil inflammasome remain unknown. Here, we identified that ketogenic diet (KD) increases  $\beta$ -hydroxybutyrate (BHB) and alleviates urate crystal-induced gout without impairing immune defense against bacterial infection. BHB inhibited NLRP3 inflammasome in S100A9 fibril-primed and urate crystal-activated macrophages, which serve to recruit inflammatory neutrophils in joints. Consistent with reduced gouty flares in rats fed a ketogenic diet, BHB blocked IL-1 $\beta$  in neutrophils in a NLRP3-dependent manner in mice and humans irrespective of age. Mechanistically, BHB inhibited the NLRP3 inflammasome in neutrophils by reducing priming and assembly steps. Collectively, our studies show that BHB, a known alternate metabolic fuel, is also an anti-inflammatory molecule that may serve as a treatment for gout.

## INTRODUCTION

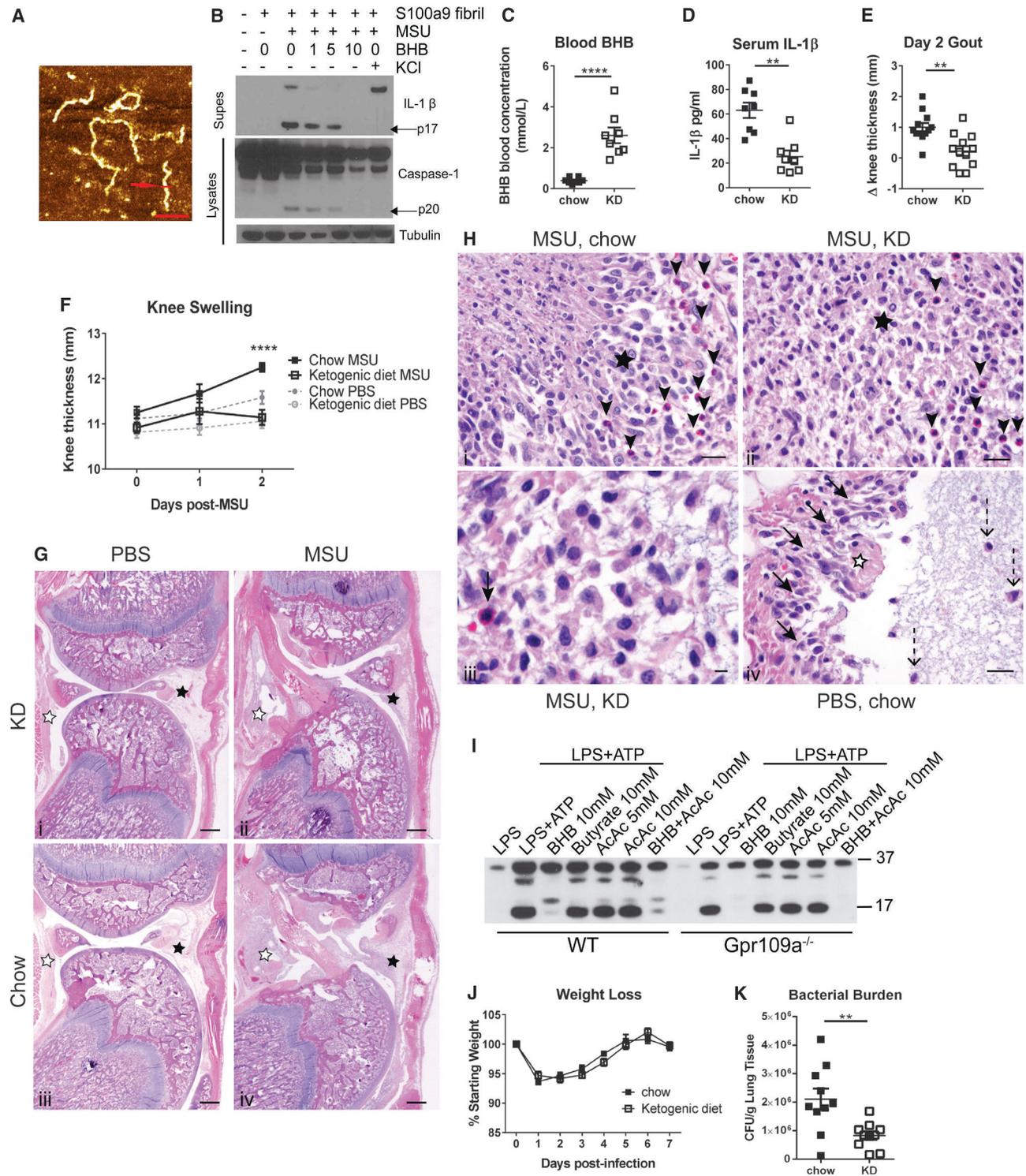
Gout is a debilitating chronic inflammatory arthritis that afflicts 4% of adults in the United States and is caused by the deposition of monosodium urate (MSU) crystals in the joints (Garrod, 1876; Seegmiller and Howell, 1962). It is widely known that aging and diet-induced lipotoxicity predispose for the development of gout, as the prevalence of gout in elderly persons above the age of 60 years is  $\sim$ 10% (Joosten et al., 2010; Li et al., 2013; Roddy and Choi, 2014). MSU crystal-induced gouty flares are characterized by interleukin-1 $\beta$  (IL-1 $\beta$ )-driven acute inflammation, fever, and intense pain caused by monocyte-mediated

neutrophil accumulation and activation in joints (Duff et al., 1983). Mechanistically, the inflammatory gouty flares caused by MSU crystals are mediated via the activation of the NLRP3 inflammasome in myeloid cells that causes the release of bioactive IL-1 $\beta$  and IL-18 (Martinon et al., 2006).

The NLRP3 inflammasome is a carefully regulated inflammatory complex that responds to both endogenous cellular abnormalities and microbial components. Activation of the complex requires two signals: (1) a priming signal (signal 1) that licenses nuclear factor  $\kappa$ B (NF- $\kappa$ B)-dependent expression of inflammasome complex proteins and (2) a secondary signal (signal 2) that promotes assembly of the complex leading to caspase-1 activation. MSU crystals provide signal 2 of NLRP3 activation, and NLRP3-deficient mice are protected from MSU-induced inflammation (Martinon et al., 2006). Therefore, gout patients face a vicious feed-forward loop in which the chronic deposition and presence of MSU crystals enables constant reactivation of acute inflammatory responses, known as gouty flares, which are associated with intense pain and fever due to high systemic IL-1 $\beta$  levels (Dalbeth et al., 2016; Duff et al., 1983).

The long-term prophylactic treatment of gout hinges on reducing hyperuricemia. Intriguingly, a major side effect of all current urate-lowering drugs such as xanthine oxidase inhibitors (allopurinol and febuxostat), recombinant uricase (pegloticase), and uricosurics (probenecid and benzbromarone) is the induction of gouty flares (Dalbeth et al., 2016). Despite progress in the field, the relief from inflammation during acute gouty flares currently relies on non-NLRP3-specific therapeutic approaches such as the use of adrenocorticotropic hormone (ACTH), corticosteroids, or nonsteroidal anti-inflammatory drugs (NSAIDs) (Dalbeth et al., 2016 and references therein). Thus, elucidating the endogenous metabolites that regulate NLRP3 inflammasome activation and resolution of inflammation remain a priority for effective management of gout.

Interestingly, metabolic interventions such as caloric restriction (CR) or moderate carbohydrate restriction reduces gout (Dessein et al., 2000). Of note, a classical feature of adaptive



**Figure 1. Ketogenic Diet Protects Rats from MSU-Induced Gouty Flare**

(A) Atomic force microscopy (AFM) height image of S100a9 amyloid fibrils; scale bar, 120 nm.

(B) Western blot of caspase-1 and IL-1β activation in BMDMs.

(C) Blood BHB levels after 1 week of ketogenic diet (KD) feeding, prior to injection with MSU.

(D) Serum IL-1β 48 hr post-MSU injection.

(E) Change in knee thickness 48 hr post-MSU injection.

(F) Knee thickness was measured daily.

(legend continued on next page)

starvation response during negative energy balance is the induction of fatty acid oxidation and production of ketone bodies  $\beta$ -hydroxybutyrate (BHB) and acetoacetate (AcAc) that serve as major substrates for ATP production to support the function of heart and brain (Cahill, 2006). Intriguingly, CR and BHB reduce inflammation and extend lifespan in animals (Edwards et al., 2014; Mitchell et al., 2016), suggesting immune-metabolic interactions driven by BHB may serve as pseudostarvation signals or CR mimetic that could be harnessed against acute inflammatory diseases such as gouty flares.

We made the surprising discovery that myeloid cells also express ketogenic and ketolytic machinery and that BHB blocks the NLRP3 inflammasome in macrophages (Youm et al., 2015). This suggests that ketones may function as regulatory metabolites, serving as endogenous regulators of inflammation (Goldberg and Dixit, 2015). Given that neutrophils are key instigators of inflammation-induced gouty flares, we hypothesized that upregulation of BHB-mediated signals in neutrophils serve as a key immunometabolic checkpoint against gout. Here, we report that BHB prevents NLRP3 inflammasome activation in both mouse and human neutrophils irrespective of the aging process. We found that BHB inhibits the signals that control the priming and assembly of the NLRP3 inflammasome in primary neutrophils and protects against gout.

## RESULTS

### Elevated BHB Protects against Acute Gout

MSU crystals are the hallmark characteristic of gout (McCarty and Hollander, 1961); the proposed mechanism of a gouty flare is that activation of macrophages by urate crystals recruits neutrophils in an inflammasome-dependent manner, and the accumulation and activation of neutrophils cause substantial pain and swelling (Duff et al., 1983; Martinon et al., 2006). The priming of inflammasome is also a critical step in the process of gouty flare. Interestingly, cytoplasmic S100A8/9 protein heterodimers are highly expressed in neutrophils (Edgeworth et al., 1991) and increase in synovial fluid (Holzinger et al., 2014) and plasma (Ryckman et al., 2003) during a gout flare. Furthermore, S100A8 can activate TLR4 (Vogl et al., 2007), and MSU stimulates the secretion of S100A8/9 from neutrophils (Ryckman et al., 2004). We found that BHB prevented caspase-1 activation and IL-1 $\beta$  secretion in MSU-treated S100A9-primed bone-marrow-derived macrophages (BMDMs) (Figures 1A and 1B). While S100A9 fibrils provided a priming signal in BMDMs (Figure S1E), they do not function as a danger-associated molecular pattern (DAMP) to provide signal 2 for inflammasome activation in lipopolysaccharide (LPS)-primed BMDM (Figure S1F). Thus, S100A8/9

proteins may represent one of the endogenous sources of priming signal for NLRP3 inflammasome in vivo.

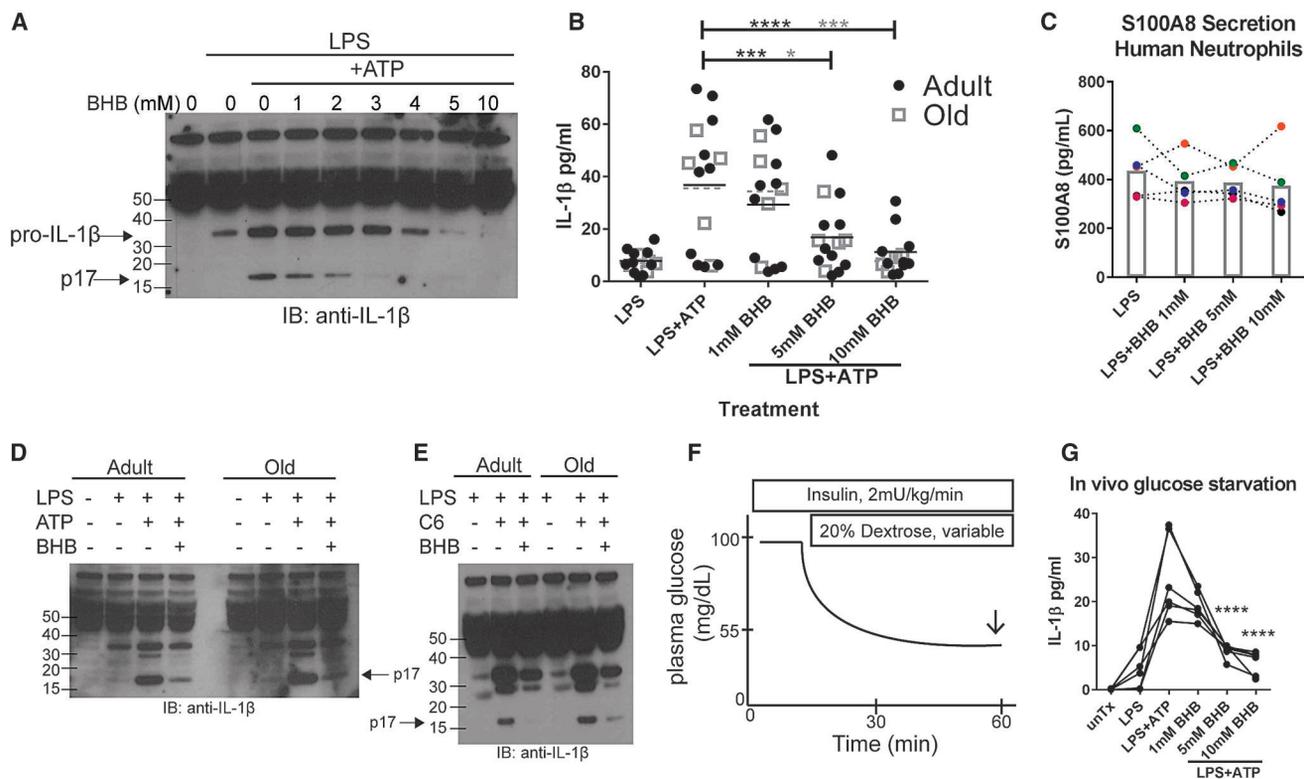
BHB serves as an alternate metabolic fuel during starvation state or in absence of insulin when glucose cannot be utilized. Ketogenic diets (KDs) that are rich in fat and low in carbohydrates are routinely employed to induce BHB or nutritional ketosis, because sustained elevation of BHB through infusion of sodium salts of BHB is contraindicated due to adverse effects on blood acid-base balance. Thus, KD has been safely employed to treat drug-resistant epilepsy (Cahill, 2006). Although IL-1R antagonists have shown benefit in small clinical trials (So et al., 2007, 2010), the high cost of these biologicals and their potential detrimental impact on host defense have limited their use in treatment of gouty flares. Therefore, we next developed a model of gouty arthritis and investigated the induction of nutritional ketosis by feeding KD as a potential therapy against gout. Humans, but not rodents, are susceptible to gout due to loss of the uricase enzyme, which normally functions to prevent high concentrations of uric acid. Therefore, we developed an in vivo gout model in outbred Sprague-Dawley rats by intra-articular injection of MSU crystals in the knee. One week of high-fat, low-carbohydrate KD feeding induced endogenous BHB production (Figure 1C). Rats fed KD were protected from the MSU-induced elevated serum IL-1 $\beta$  (Figure 1D) and knee swelling observed in the chow-fed rats (Figures 1E and 1F). Pathology analysis of H&E-stained sections of the joints showed that MSU-injected rats displayed combined lesions of intra-articular exudate and synovial soft tissue inflammation (Figure 1G). Intra-articular exudate was characterized by masses of fibrin, amorphous granular foreign material, and clusters of macrophages (Figure 1H). The extent of both intra-articular exudate and synovial inflammation was less severe in KD-fed animals than in those on chow diet. Overall, KD reduced the severity of the inflammatory reactions in MSU-injected knees. Foci of frank necrosis were present in animals on the control diet, but not in those on the ketone-rich diet (Figure 1H). Notably, KD results in elevation of blood ketone bodies BHB as well as AcAc. Our data suggest that the anti-inflammatory effects of ketones are limited to BHB, as neither AcAc nor the microbiota-derived short chain fatty acid butyrate prevented IL-1 $\beta$  secretion in inflammasome-activated BMDM (Figure 1I). Importantly, reduced inflammatory responses during KD did not increase disease severity in mice infected with *Staphylococcus aureus* (Figures 1J and S1G–S1I) and surprisingly reduced bacterial burdens in the lungs of infected mice (Figure 1K). Taken together, our data show that elevated BHB levels protect against acute gouty flare without compromising the host-defense functions of the immune system.

(G) Representative sections of the femoro-tibial joint showing local tissue reaction (black asterisks) and intrasynovial exudate (white asterisks); scale bar, 500  $\mu$ m. (H) Representative images of synovial inflammation. (i and ii) Regions of macrophage infiltration (black asterisk) and neutrophil infiltration (black arrows); scale bar, 20  $\mu$ m. (iii) Higher magnification of MSU-induced inflammatory response, where the black arrow points to an infiltrating neutrophil; scale bar, 10  $\mu$ m. (iv) Local tissue reaction in a control PBS-injected joint showing focal fibrin exudate (white asterisk) with surrounding macrophages (black arrows) and macrophages within granular amorphous injected material (dashed arrows); scale bar, 20  $\mu$ m.

(I) Western blot of IL-1 $\beta$  secretion from stimulated BMDMs.

(J and K) Body weights (J) and bacterial load (K) after *S. aureus* infection. Data are pooled from at least three independent experiments; in (C)–(F), (J), and (K), data are represented as mean  $\pm$  SEM.

Statistical differences calculated by t test (C–E) or two-way ANOVA (F). See also Figure S1.



**Figure 2. BHB Regulates Neutrophil NLRP3 Inflammasome Activation**

(A and B) IL-1 $\beta$  secretion by stimulated mouse (A) and human (B) neutrophils (lines indicate group mean). (C) Secretion of S100A8 from stimulated human neutrophils. Different-colored symbols connected by dotted line represents a unique individual's response to all treatments. Columns indicate the mean for each treatment. (D and E) Western blot analysis of IL-1 $\beta$  secretion from LPS-primed adult and old murine neutrophils stimulated with (D) ATP  $\pm$  BHB or (E) ceramide  $\pm$  BHB. (F) Depiction of insulin-induced hypoglycemia in vivo. Arrow indicates blood collection time point. (G) Neutrophil IL-1 $\beta$  secretion from hypoglycemic subjects. Lines connect the responses of each individual for each treatment condition. Statistical differences were calculated by paired one-way ANOVA. See also Figure S2.

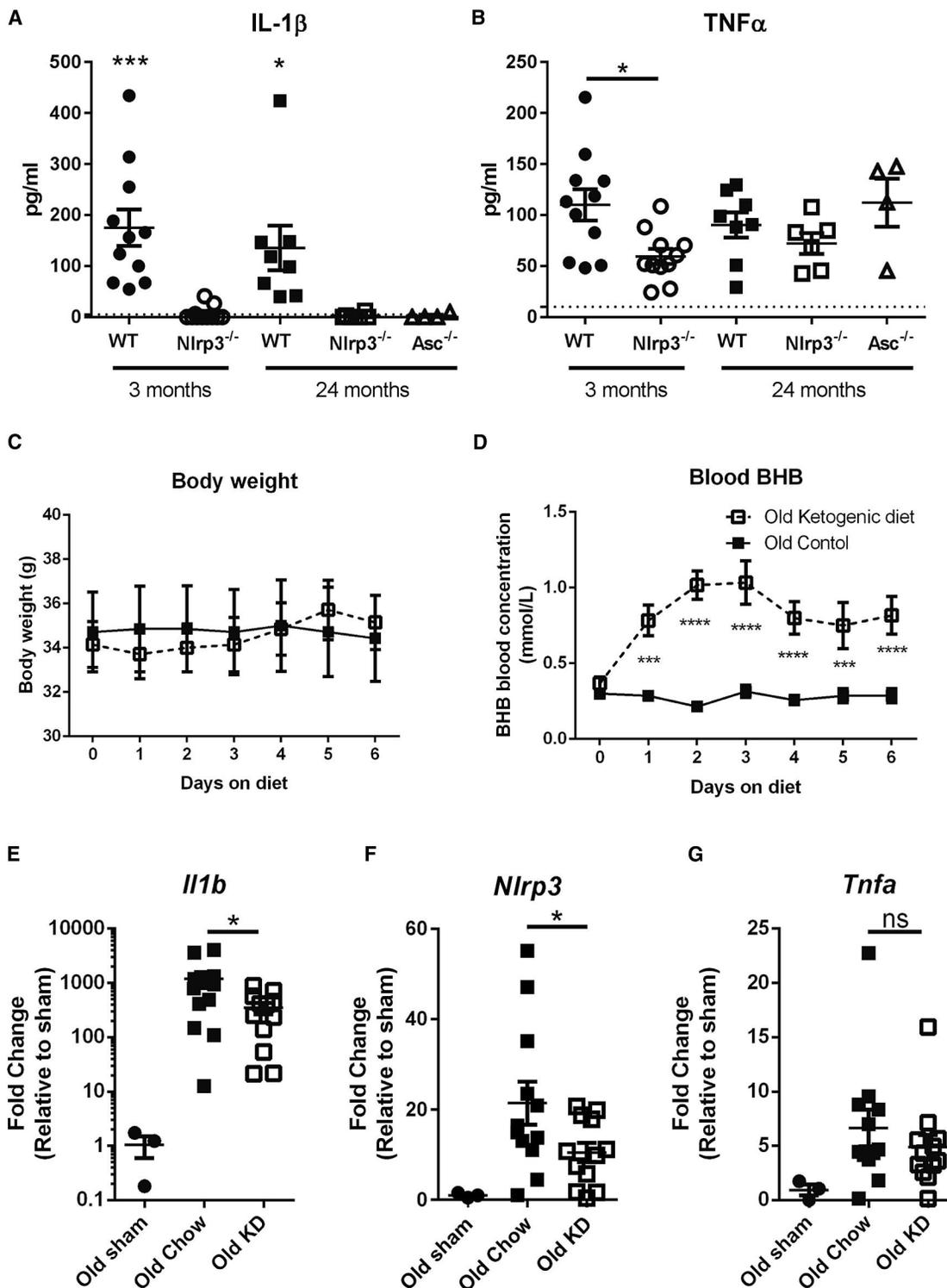
### BHB Acts on Neutrophils to Block IL-1 $\beta$ Secretion throughout the Lifespan

Gout flares are mediated by both macrophage and neutrophil activation. Because BHB inhibits NLRP3 inflammasome activation in macrophages (Youm et al., 2015) and in an in vitro model of gout (Figure 1B), we next tested whether BHB regulates IL-1 $\beta$  secretion from neutrophils. Indeed, BHB dose-dependently inhibited IL-1 $\beta$  secretion in isolated murine neutrophils (Figures 2A and S2). Aging is a major risk factor for gout, and if not properly managed, gouty flares increase in frequency and intensity in the elderly. Importantly, BHB potently inhibited the NLRP3 inflammasome-induced IL-1 $\beta$  secretion in neutrophils of young and elderly humans (Figure 2B), but not secretion of S100A8 protein implicated in the propagation of gouty flares (Figure 2C). BHB inhibited IL-1 $\beta$  secretion in response to classical NLRP3 inflammasome activation (Figure 2D) as well as the age-related lipotoxic DAMP ceramide (Figure 2E) in isolated adult and old murine neutrophils. Ketogenesis is typically induced during hypoglycemia or lack of glucose availability, which may alter neutrophil function. Therefore, we next determined whether neutrophils in humans exposed to hypoglycemia in vivo respond to

BHB (Figure 2F). Indeed, IL-1 $\beta$  secretion is still sensitive to BHB-mediated inhibition regardless of glucose availability in vivo (Figure 2G). These experiments reveal a regulatory role for BHB in neutrophil inflammasome activation regardless of the host's age.

### BHB Reduces Urate-Crystal-Induced Inflammation during Aging

Neutrophils are reported to accumulate several defects during aging, including impaired Toll-like receptor (TLR) signaling (Qian et al., 2014) and NETosis (Hazeldine et al., 2014). Neutrophils are also associated with immunopathology during infection in aged hosts (Bou Ghanem et al., 2015; Menter et al., 2014). Bone marrow neutrophils from aged mice exhibited no remarkable differences in abundance or expression of NLRP3 inflammasome proteins compared to young controls (Figure S2). To test neutrophil-intrinsic inflammasome activation defects during aging, isolated neutrophils from adult and old mice were stimulated with LPS+ATP. IL-1 $\beta$ , but not tumor necrosis factor  $\alpha$  (TNF- $\alpha$ ) secretion, was NLRP3- and ASC-dependent and no age-related differences were observed (Figures 3A and 3B). Due to the potent inhibitory effects of BHB on aged neutrophils



**Figure 3. Elevated BHB Protects against MSU-Induced Peritonitis in Aged Mice**

(A and B) IL-1 $\beta$  (A) and TNF- $\alpha$  (B) were measured in culture supernatants after LPS+ATP stimulation of isolated neutrophils from mice of indicated ages and genotypes. Statistical differences were calculated by age-specific t test (3 months) or one-way ANOVA (24 months). Data are pooled from two independent experiments.

(C and D) Body weights (C) and blood BHB concentrations (D) were measured daily in old mice fed KD, prior to MSU injection.

(E–G) Gene expression of (E) *Il1b*, (F) *Nlrp3*, and (G) *Tnfa* within peritoneal cells 4 hr post-MSU. Data are presented as fold change relative to sham and are pooled from three independent experiments.

All data are shown as mean  $\pm$  SEM. Statistical differences were calculated by t test between MSU-injected groups. See also Figure S3.

in vitro (Figures 2B, 2D, and 2E), we next sought to determine whether BHB could reduce neutrophilic inflammation in vivo during aging. To induce endogenous ketogenesis, old mice were fed KD for 1 week. During this time, no alterations to body weight (Figure 3C) were observed and blood BHB levels increased rapidly (Figure 3D). KD did not change peritoneal neutrophil infiltration (Figure S3) but prevented upregulation of *Nlrp3* and *Il1b* gene expression, but not the general inflammatory marker *Tnf $\alpha$*  in old mice (Figures 3E–3G) after MSU-induced peritonitis. This suggests that BHB levels can be elevated during aging to reduce acute urate-crystal-induced inflammatory responses.

### **$\beta$ -Hydroxybutyrate Inhibits NLRP3 Inflammasome Priming and Assembly**

Neutrophils contain both inflammasome-dependent and inflammasome-independent IL-1 $\beta$  cleavage processes (Cassel et al., 2014; Joosten et al., 2009; Karmakar et al., 2015; Mankan et al., 2012). In our experimental conditions, neutrophil IL-1 $\beta$  secretion, in response to several DAMPs, including extracellular ATP, the crystalline silica, or lipotoxic ceramide, was entirely caspase-1/11-dependent (Figure 4A). In cell-free assays, BHB had no effect on enzyme activity of neutrophil serine proteases elastase (Figure 4B) or cathepsin G (Figure 4C), both of which have been implicated in IL-1 $\beta$  cleavage (Cassel et al., 2014; Guma et al., 2009). We next investigated the mechanism by which BHB inhibits NLRP3 inflammasome activation in neutrophils. GPR109a can bind BHB on the cell surface to inhibit inflammation; however, we found that niacin, a high-affinity ligand for GPR109a, failed to impact neutrophil IL-1 $\beta$  cleavage, suggesting that in vitro GPR109a signaling is dispensable for neutrophil inflammasome activation (Figure 4D). Ketogenesis occurs during energy restriction, a physiological state in which autophagy is induced. However, BHB's inhibition of IL-1 $\beta$  secretion does not rely on autophagy, because treatment of neutrophils with the autophagy inhibitor 3-MA did not prevent BHB-mediated inhibition of IL-1 $\beta$  secretion (Figure 4E). Similarly, treating cells with the tricarboxylic acid (TCA) cycle entry inhibitor (aminoxy)acetic acid hemihydrochloride (AOA) did not prevent BHB's inhibition of IL-1 $\beta$  secretion, suggesting oxidation of BHB is not required for its effect (Figure 4E). Interestingly, the non-oxidizable chiral enantiomer S-BHB, which cannot enter the TCA cycle, also inhibited IL-1 $\beta$  secretion from neutrophils (Figure 4F), suggesting that in the presence of glucose in vitro, myeloid cells energetically spare BHB to block inflammasome activation.

Unaltered caspase-11 activation or gasdermin D expression in the presence of BHB (Figure S4A) suggested that BHB also does not impact neutrophil pyroptosis (He et al., 2015; Kayagaki et al., 2015; Shi et al., 2015) and BHB did not alter cell viability (Figure S4B). BHB treatment led to increased histone H3 acetylation (Figure 4G), probably due to its reported histone deacetylase (HDAC) inhibitor activity (Shimazu et al., 2013). Interestingly, BHB inhibited phosphorylation of NF- $\kappa$ B (Figure 4H), a necessary signaling event for inflammasome activation, confirming previous findings (Fu et al., 2014). Finally, BHB was tested in a mouse model of the severe human disease familial cold autoinflammatory syndrome (FCAS), which contains an L351P nucleo-

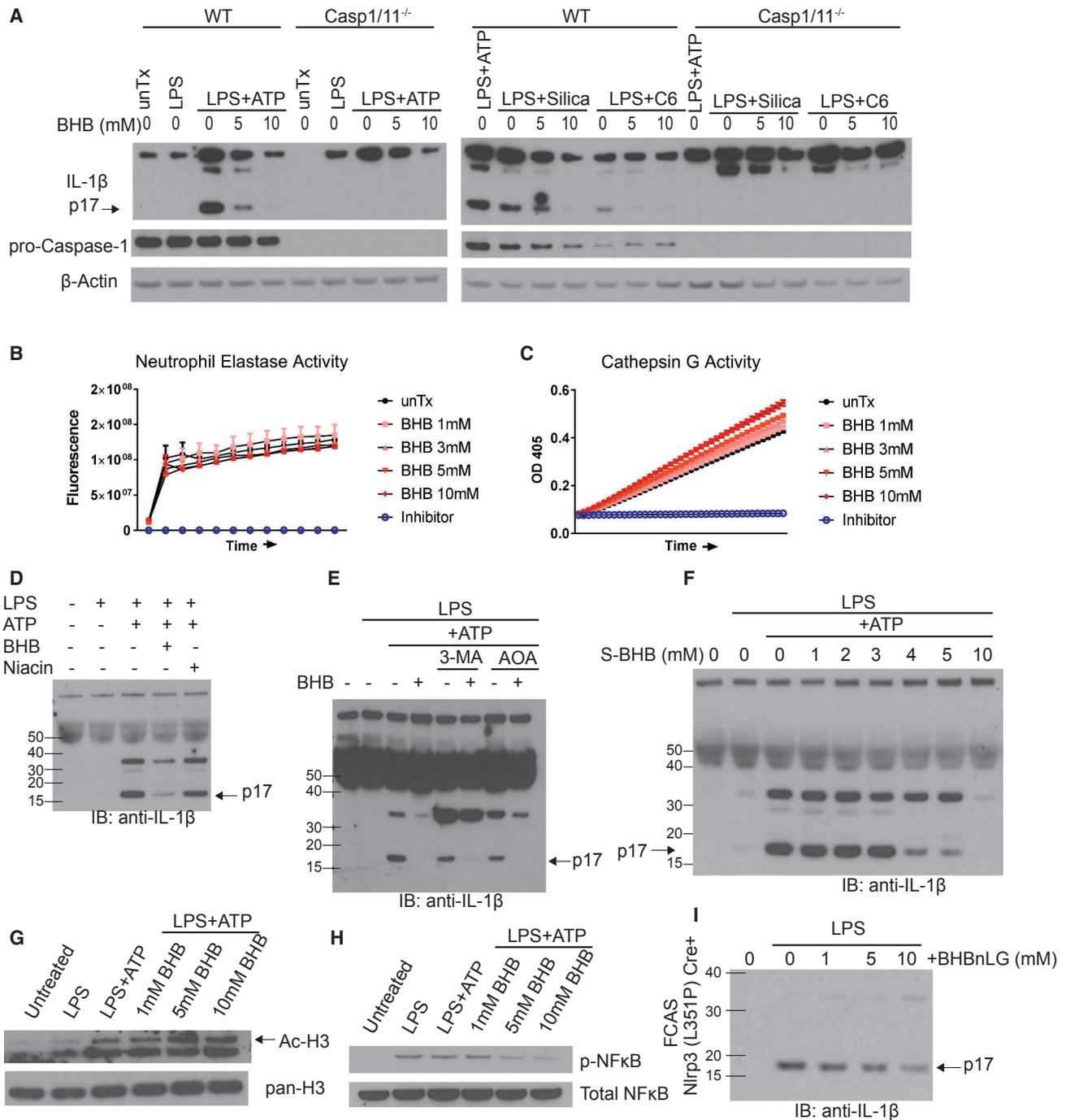
tide substitution in *Nlrp3* that facilitates constitutive inflammasome assembly (Brydges et al., 2009; Yu et al., 2006). IL-1 $\beta$  secretion from neutrophils in this FCAS model was dose-dependently inhibited by treatment with BHB-conjugated nanolipogels (Figure 4I). These data suggest that the inhibitory effect of BHB upon NLRP3 inflammasome activation is two-pronged: (1) it prevents TLR4-mediated priming, and (2) it blocks the physical assembly of the NLRP3 inflammasome complex.

## **DISCUSSION**

Gout is a chronic disease characterized by recurrent painful gouty flares. The incidence of gout has steadily increased, and individuals over age 65 account for the majority of gout-related hospitalizations (Lim et al., 2016). Although a primary treatment strategy for gout is to lower uric acid levels, many of these medications induce gouty flares, presumably due to disruption of tophi, which results in poor adherence by patients. We found that elevated blood BHB levels protected outbred rats from joint swelling and systemic inflammation after intra-articular injection of MSU crystals. Reduced joint swelling was due to a qualitative reduction in the inflammatory response, as pathology analysis revealed reduced tissue damage despite similar neutrophil and macrophage infiltration into the joints.

In vitro, BHB also inhibits neutrophil IL-1 $\beta$  secretion from adult and elderly individuals. Notably, elevated BHB levels during KD did not increase disease severity during *S. aureus* infection in mice and even reduced bacterial burdens in the lungs of infected mice, suggesting high translational potential of BHB against gouty flares. It was recently reported that glucose metabolism promotes mortality during bacterial infection and LPS sepsis (Wang et al., 2016), which may explain the beneficial effects of KD in *S. aureus*-infected mice, as BHB reduces glucose availability. Notably, other macronutrients in KD can induce hormonal alterations and plasma lipid profile changes in patients. Moreover, at the cellular level, limited glucose availability during KD can induce autophagy, which inhibits the inflammasome. Thus, additional studies are required to rule out the exact contribution of each of these mechanisms in mediating anti-inflammasome effects of KD in vivo in models of gout. Regardless, our findings that BHB can target neutrophil inflammasome have important clinical implications, as an estimated 8 million Americans have gout (Zhu et al., 2011) and the cumulative nature of the disease causes the risk and frequency of gouty flares to increase during aging.

S100A8/9 proteins are also reported to increase during age-related inflammatory diseases, including obesity (Nagareddy et al., 2014), cardiovascular disease (Ma et al., 2012), and Alzheimer's disease (Wang et al., 2014), and in the aged prostate (Yanamandra et al., 2009). This implies that these proteins may provide an endogenous, local signal 1 for NLRP3 inflammasome activation and may be universal biomarkers of inflammation (Vogl et al., 2007). During a gouty flare, neutrophils are recruited to inflamed joints by resident macrophages and secrete a variety of proinflammatory molecules, including S100A8/9 and IL-1 $\beta$  (Ryckman et al., 2004). BHB inhibited NLRP3 inflammasome activation in S100A9-primed macrophages (Figure 1B). Our data suggest that BHB can break this feed-forward cycle to



**Figure 4. BHB Inhibits Inflammasome Priming and Assembly**

(A) Western blot of culture supernatants from LPS-primed wild-type (WT) and caspase-1/11<sup>-/-</sup> neutrophils after ATP, silica, or ceramide stimulation as indicated. (B and C) Elastase (B) and cathepsin G (C) activity; n = 4, data are presented as mean  $\pm$  SEM.

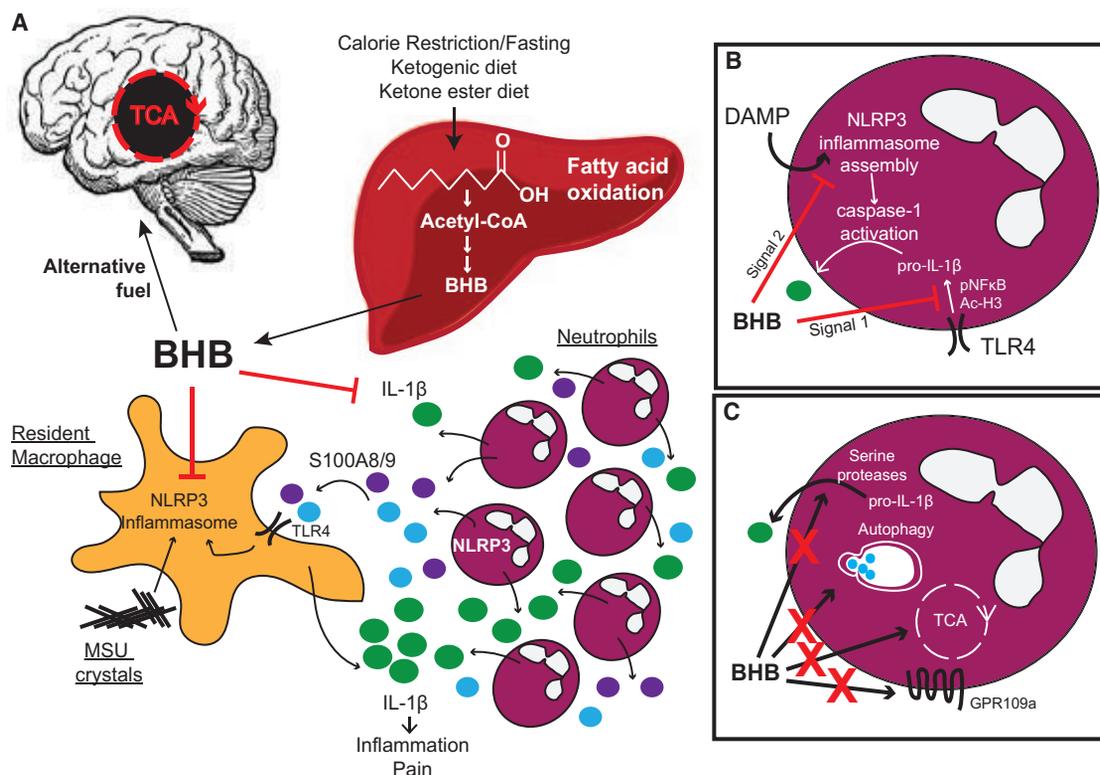
(D-F) Culture supernatants were analyzed for neutrophil IL-1 $\beta$  secretion by western blot after treatment with BHB  $\pm$  (D) niacin, (E) 3-MA or AOA, and (F) S-BHB as indicated.

(G) Total and acetylated H3 expression in cell lysates.

(H) NF- $\kappa$ B phosphorylation in neutrophil cell lysates.

(I) IL-1 $\beta$  secretion from neutrophils derived from a mouse model of FCAS bearing activating mutation of Nlrp3.

For all blots, each sample is pooled from at least n = 4 mice per experiment. Each blot is representative of at least two independent experiments. See also Figure S4.



**Figure 5. Hypothetical Model for the Dual Role of BHB as a Regulatory Metabolite**

(A) The metabolic role of BHB as an alternative energy source during limited energy availability, coordinated with the anti-inflammatory role of BHB on the NLRP3 inflammasome.

(B) BHB inhibits both the priming signal and the secondary assembling signal in neutrophils.

(C) Inhibition of IL-1 $\beta$  secretion by BHB does not involve inhibition of serine protease activity, autophagy, oxidation in the TCA cycle, or the surface receptor GPR109a, although additional experiments will be needed to further clarify the mechanism.

prevent swelling and inflammation (Figure 5), although defining the exact mechanism will require further experimentation. Neutrophils can secrete IL-1 $\beta$  by caspase-1-independent mechanisms. Although we were not able to detect IL-1 $\beta$  secretion from caspase-1/11 knockout mice, cell-free assays revealed no effect on enzymatic activity of serine proteases neutrophil elastase or cathepsin G. These data highlight the targeted effects of BHB upon the NLRP3 inflammasome, making it an ideal candidate for preventing NLRP3-driven inflammation.

Together with our prior findings that BHB inhibits NLRP3 inflammasome activation in macrophages (Youm et al., 2015) and our current data that BHB also regulates neutrophil inflammasome activation, we propose that strategies to increase BHB levels are likely to be therapeutically beneficial in gout patients. Furthermore, the NLRP3-targeted effects of BHB make it ideal for reducing or preventing many age-related inflammatory disease that have been shown to be driven by chronic NLRP3 inflammasome activation (Youm et al., 2013). Adherence to a high-fat low-carbohydrate KD is difficult and can promote dyslipidemia in gout patients. However, ketone esters have been delivered orally to humans, which increased circulating BHB levels and enhanced physical performance in athletes (Clarke et al., 2012; Cox et al., 2016), and should be explored for

inducing mild ketosis to prevent inflammation in individuals with gout.

## EXPERIMENTAL PROCEDURES

### Animals

All animals were housed under specific-pathogen-free conditions under a 12-hr light/dark cycle. All mice used were on the C57BL/6 genetic background. See Supplemental Experimental Procedures for details of mouse and rat strains. KD (Research Diets D12369B) was initiated 1 week prior to experimental manipulation. Blood BHB concentrations were measured with Precision Xtra  $\beta$ -ketone strips. All animal experiments were performed in compliance with the Yale University Institutional Animal Care and Use Committee.

### Human Subjects

Healthy adult (18–45 years) and old (>65 years) males and females with no current steroid use were recruited. Individuals were not fasting at time of peripheral blood collection, except for hypoglycemic studies. Insulin was used to induce hypoglycemia, see Supplemental Experimental Procedures for detailed procedure. Informed consent was obtained from all subjects and all studies were approved by the Institutional Review Committee of Yale University.

### Cell Isolation and Activation

Neutrophils were isolated from bone marrow (mice) or peripheral blood (humans) by negative magnetic selection (Stem Cell Technology; Figure S2).

BMDM were generated and all cells were stimulated as previously described (Youm et al., 2015), see Supplemental Experimental Procedures for detailed treatment conditions.

### In Vitro Activation Measurements

All biochemical analysis methods are described in detail in Supplemental Experimental Procedures. Cytokine secretion was measured by Multiplex (Figures 3A and 3B; IL-1 $\beta$  and TNF- $\alpha$  from Life Technologies) or ELISA (human IL-1 $\beta$  from eBioscience 88-7261-22; S100A8 from Thermo EHS100A8) according to manufacturers' protocols. Elastase (ab118971) and cathepsin G (ab204693) enzymatic activity was assessed using kits (Abcam) according to manufacturer's instructions.

### Peritonitis Model

Peritonitis was induced by intraperitoneal (i.p.) injection of MSU crystals (InvivoGen). Mice were injected with 2.5 mg MSU in PBS. 4 hr later, total peritoneal cells were collected by lavage. Cells were counted using a hemacytometer, and phenotype and gene expression were analyzed by flow cytometry and RT-PCR, respectively.

### Staphylococcus aureus Infection

Mice were infected intranasally with 10<sup>9</sup> colony-forming units (CFUs) or 10<sup>6</sup> CFUs of *S. aureus* (strain 14458). Bacterial burdens were determined by plating serial dilutions 24 hr post-infection. Bronchoalveolar lavage fluid (BALF) was collected by washing lungs three times with 1 mL sterile PBS.

### Gout Model

Gout was induced in rats by intra-articular injection of 1.25 mg MSU in the knee. Knee thickness was measured with digital calipers. IL-1 $\beta$  was measured in serum by ELISA (eBioscience BMS630). Knees were fixed and decalcified in Bouin's solution (Sigma). Tissue sectioning and H&E staining were performed by the Yale Mouse Research Pathology and Histology Core. For pathology analysis, all sections were taken from the mid-sagittal region of the femoro-tibial joint, encompassing cruciate ligaments and menisci. Images are oriented with the patellar ligament on the right. In high-power images, all images were taken from the anterior synovial tissue.

### Statistical Analysis

Statistical analyses were performed using GraphPad Prism software (GraphPad) as indicated in the text and figure legends.  $p < 0.05$  was considered statistically significant for all tests. All experiments were performed at least twice. All graphs shown are combined from all replicates of each experiment, and each data point represents an individual test subject. All data are expressed as mean  $\pm$  SEM unless otherwise specified. For all statistical differences, \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and \*\*\*\* $p < 0.0001$ . Western blot experiments were repeated at least three times, each time pooling cells from  $n = 3$ –5 mice.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.02.004>.

### AUTHOR CONTRIBUTIONS

E.L.G. performed experiments and data analysis and prepared the manuscript. J.L.A. performed gout experiments and clinical evaluations. R.D.M. performed infection experiments. C.W. and L.A.M.-R. provided S100A9 reagents and expertise. A.C.S. provided human samples. R.I.H. provided expertise and samples from hypoglycemic clamp studies. C.J.Z. evaluated gout pathology. A.I. designed and interpreted the infection experiments. V.D.D. conceived and supervised the project, interpreted the data, and prepared the manuscript.

### ACKNOWLEDGMENTS

We thank Kim Nguyen for expert technical assistance. V.D.D. is supported in part by US NIH grants AG043608, P01AG051459, AR070811, and AI105097 and The Glenn Foundation for Aging Research. E.L.G. is supported in part by postdoctoral fellowship from the American Foundation of Aging Research (AFAR) and American Heart Association (AHA). R.D.M. is supported by NIH T32 grants AI007019-38 and AI055403 and the Francis Trudeau Trainee Fellowship. A.C.S. is supported by NIH grants K24 AG042489, U19 AI089992, and P30 AG21342 (Yale Claude D. Pepper Older Americans Independence Center). R.I.H. is supported by NIH/NIDDK grants K08 DK082618 and R01 DK101984-02 and the hypoglycemic clamp studies are supported by the Yale Center for Clinical Investigation (YCCI) and Hospital Research Unit (HRU) personnel via DK045735, a Yale Diabetes Research Center grant, and Clinical Translational Science Award UL1-RR-024139 from the National Center for Advancing Translational Sciences, a component of the NIH, and the NIH Roadmap for Medical Research. A.I. is supported by the Howard Hughes Medical Institute.

Received: May 26, 2016

Revised: November 16, 2016

Accepted: January 30, 2017

Published: February 28, 2017

### REFERENCES

- Bou Ghanem, E.N., Clark, S., Roggensack, S.E., McIver, S.R., Alcaide, P., Haydon, P.G., and Leong, J.M. (2015). Extracellular adenosine protects against *Streptococcus pneumoniae* Lung infection by regulating pulmonary neutrophil recruitment. *PLoS Pathog.* *11*, e1005126.
- Brydges, S.D., Mueller, J.L., McGeough, M.D., Pena, C.A., Misaghi, A., Gandhi, C., Putnam, C.D., Boyle, D.L., Firestein, G.S., Horner, A.A., et al. (2009). Inflammation-mediated disease animal models reveal roles for innate but not adaptive immunity. *Immunity* *30*, 875–887.
- Cahill, G.F., Jr. (2006). Fuel metabolism in starvation. *Annu. Rev. Nutr.* *26*, 1–22.
- Cassel, S.L., Janczy, J.R., Bing, X., Wilson, S.P., Olivier, A.K., Otero, J.E., Iwakura, Y., Shayakhmetov, D.M., Bassuk, A.G., Abu-Amer, Y., et al. (2014). Inflammation-independent IL-1 $\beta$  mediates autoinflammatory disease in Pstpip2-deficient mice. *Proc. Natl. Acad. Sci. USA* *111*, 1072–1077.
- Clarke, K., Tchabanenko, K., Pawlosky, R., Carter, E., Todd King, M., Musa-Veloso, K., Ho, M., Roberts, A., Robertson, J., Vanitallie, T.B., and Veech, R.L. (2012). Kinetics, safety and tolerability of (R)-3-hydroxybutyl (R)-3-hydroxybutyrate in healthy adult subjects. *Regul. Toxicol. Pharmacol.* *63*, 401–408.
- Cox, P.J., Kirk, T., Ashmore, T., Willerton, K., Evans, R., Smith, A., Murray, A.J., Stubbs, B., West, J., McLure, S.W., et al. (2016). Nutritional ketosis alters fuel preference and thereby endurance performance in athletes. *Cell Metab.* *24*, 256–268.
- Dalbeth, N., Merriman, T.R., and Stamp, L.K. (2016). Gout. *Lancet* *388*, 2039–2052.
- Dessein, P.H., Shipton, E.A., Stanwix, A.E., Joffe, B.I., and Ramokgadi, J. (2000). Beneficial effects of weight loss associated with moderate calorie/carbohydrate restriction, and increased proportional intake of protein and unsaturated fat on serum urate and lipoprotein levels in gout: a pilot study. *Ann. Rheum. Dis.* *59*, 539–543.
- Duff, G.W., Atkins, E., and Malawista, S.E. (1983). The fever of gout: urate crystals activate endogenous pyrogen production from human and rabbit mononuclear phagocytes. *Trans. Assoc. Am. Physicians* *96*, 234–245.
- Edgeworth, J., Gorman, M., Bennett, R., Freemont, P., and Hogg, N. (1991). Identification of p8,14 as a highly abundant heterodimeric calcium binding protein complex of myeloid cells. *J. Biol. Chem.* *266*, 7706–7713.

- Edwards, C., Canfield, J., Copes, N., Rehan, M., Lipps, D., and Bradshaw, P.C. (2014). D-beta-hydroxybutyrate extends lifespan in *C. elegans*. *Aging (Albany NY)* 6, 621–644.
- Fu, S.P., Li, S.N., Wang, J.F., Li, Y., Xie, S.S., Xue, W.J., Liu, H.M., Huang, B.X., Lv, Q.K., Lei, L.C., et al. (2014). BHBA suppresses LPS-induced inflammation in BV-2 cells by inhibiting NF- $\kappa$ B activation. *Mediators Inflamm.* 2014, 983401.
- Garrod, A.B. (1876). *A Treatise on Gout and Rheumatic Gout (Rheumatoid Arthritis)* (Longman Green).
- Goldberg, E.L., and Dixit, V.D. (2015). Drivers of age-related inflammation and strategies for healthspan extension. *Immunol. Rev.* 265, 63–74.
- Guma, M., Ronacher, L., Liu-Bryan, R., Takai, S., Karin, M., and Corr, M. (2009). Caspase 1-independent activation of interleukin-1 $\beta$  in neutrophil-predominant inflammation. *Arthritis Rheum.* 60, 3642–3650.
- Hazeldine, J., Harris, P., Chapple, I.L., Grant, M., Greenwood, H., Livesey, A., Sapey, E., and Lord, J.M. (2014). Impaired neutrophil extracellular trap formation: a novel defect in the innate immune system of aged individuals. *Aging Cell* 13, 690–698.
- He, W.T., Wan, H., Hu, L., Chen, P., Wang, X., Huang, Z., Yang, Z.H., Zhong, C.Q., and Han, J. (2015). Gasdermin D is an executor of pyroptosis and required for interleukin-1 $\beta$  secretion. *Cell Res.* 25, 1285–1298.
- Holzinger, D., Nippe, N., Vogl, T., Marketon, K., Mysore, V., Weinhage, T., Dalbeth, N., Pool, B., Merriman, T., Baeten, D., et al. (2014). Myeloid-related proteins 8 and 14 contribute to monosodium urate monohydrate crystal-induced inflammation in gout. *Arthritis Rheumatol.* 66, 1327–1339.
- Joosten, L.A., Netea, M.G., Fantuzzi, G., Koenders, M.I., Helsen, M.M., Sparrer, H., Pham, C.T., van der Meer, J.W., Dinarello, C.A., and van den Berg, W.B. (2009). Inflammatory arthritis in caspase 1 gene-deficient mice: contribution of proteinase 3 to caspase 1-independent production of bioactive interleukin-1 $\beta$ . *Arthritis Rheum.* 60, 3651–3662.
- Joosten, L.A., Netea, M.G., Mylona, E., Koenders, M.I., Malireddi, R.K., Oosting, M., Stienstra, R., van de Veerdonk, F.L., Stalenhoef, A.F., Giamarellos-Bourboulis, E.J., et al. (2010). Engagement of fatty acids with Toll-like receptor 2 drives interleukin-1 $\beta$  production via the ASC/caspase 1 pathway in monosodium urate monohydrate crystal-induced gouty arthritis. *Arthritis Rheum.* 62, 3237–3248.
- Karmakar, M., Katsnelson, M., Malak, H.A., Greene, N.G., Howell, S.J., Hise, A.G., Camilli, A., Kadioglu, A., Dubyak, G.R., and Pearlman, E. (2015). Neutrophil IL-1 $\beta$  processing induced by pneumolysin is mediated by the NLRP3/ASC inflammasome and caspase-1 activation and is dependent on K<sup>+</sup> efflux. *J. Immunol.* 194, 1763–1775.
- Kayagaki, N., Stowe, I.B., Lee, B.L., O'Rourke, K., Anderson, K., Warming, S., Cuellar, T., Haley, B., Roose-Girma, M., Phung, Q.T., et al. (2015). Caspase-11 cleaves gasdermin D for non-canonical inflammasome signalling. *Nature* 526, 666–671.
- Li, C., Hsieh, M.C., and Chang, S.J. (2013). Metabolic syndrome, diabetes, and hyperuricemia. *Curr. Opin. Rheumatol.* 25, 210–216.
- Lim, S.Y., Lu, N., Oza, A., Fisher, M., Rai, S.K., Menendez, M.E., and Choi, H.K. (2016). Trends in Gout and Rheumatoid Arthritis Hospitalizations in the United States, 1993–2011. *JAMA* 315, 2345–2347.
- Ma, L.P., Haugen, E., Ikemoto, M., Fujita, M., Terasaki, F., and Fu, M. (2012). S100A8/A9 complex as a new biomarker in prediction of mortality in elderly patients with severe heart failure. *Int. J. Cardiol.* 155, 26–32.
- Mankan, A.K., Dau, T., Jenne, D., and Hornung, V. (2012). The NLRP3/ASC/Caspase-1 axis regulates IL-1 $\beta$  processing in neutrophils. *Eur. J. Immunol.* 42, 710–715.
- Martinson, F., Pétrilli, V., Mayor, A., Tardivel, A., and Tschopp, J. (2006). Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature* 440, 237–241.
- McCarty, D.J., and Hollander, J.L. (1961). Identification of urate crystals in gouty synovial fluid. *Ann. Intern. Med.* 54, 452–460.
- Menter, T., Giefing-Kroell, C., Grubeck-Loebenstein, B., and Tzankov, A. (2014). Characterization of the inflammatory infiltrate in Streptococcus pneumoniae pneumonia in young and elderly patients. *Pathobiology* 81, 160–167.
- Mitchell, S.J., Madrigal-Matute, J., Scheibye-Knudsen, M., Fang, E., Aon, M., González-Reyes, J.A., Cortassa, S., Kaushik, S., Gonzalez-Freire, M., Patel, B., et al. (2016). Effects of sex, strain, and energy intake on hallmarks of aging in mice. *Cell Metab.* 23, 1093–1112.
- Nagareddy, P.R., Kraakman, M., Masters, S.L., Stirzaker, R.A., Gorman, D.J., Grant, R.W., Dragoljevic, D., Hong, E.S., Abdel-Latif, A., Smyth, S.S., et al. (2014). Adipose tissue macrophages promote myelopoiesis and monocytosis in obesity. *Cell Metab.* 19, 821–835.
- Qian, F., Guo, X., Wang, X., Yuan, X., Chen, S., Malawista, S.E., Bockenstedt, L.K., Allore, H.G., and Montgomery, R.R. (2014). Reduced bioenergetics and toll-like receptor 1 function in human polymorphonuclear leukocytes in aging. *Aging (Albany NY)* 6, 131–139.
- Roddy, E., and Choi, H.K. (2014). Epidemiology of gout. *Rheum. Dis. Clin. North Am.* 40, 155–175.
- Ryckman, C., McColl, S.R., Vandal, K., de Médicis, R., Lussier, A., Poubelle, P.E., and Tessier, P.A. (2003). Role of S100A8 and S100A9 in neutrophil recruitment in response to monosodium urate monohydrate crystals in the air-pouch model of acute gouty arthritis. *Arthritis Rheum.* 48, 2310–2320.
- Ryckman, C., Gilbert, C., de Médicis, R., Lussier, A., Vandal, K., and Tessier, P.A. (2004). Monosodium urate monohydrate crystals induce the release of the proinflammatory protein S100A8/A9 from neutrophils. *J. Leukoc. Biol.* 76, 433–440.
- Seegmiller, J.E., and Howell, R.R. (1962). The old and new concepts of acute gouty arthritis. *Arthritis Rheum.* 5, 616–623.
- Shi, J., Zhao, Y., Wang, K., Shi, X., Wang, Y., Huang, H., Zhuang, Y., Cai, T., Wang, F., and Shao, F. (2015). Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. *Nature* 526, 660–665.
- Shimazu, T., Hirschey, M.D., Newman, J., He, W., Shirakawa, K., Le Moan, N., Grueter, C.A., Lim, H., Saunders, L.R., Stevens, R.D., et al. (2013). Suppression of oxidative stress by  $\beta$ -hydroxybutyrate, an endogenous histone deacetylase inhibitor. *Science* 339, 211–214.
- So, A., De Smedt, T., Revaz, S., and Tschopp, J. (2007). A pilot study of IL-1 inhibition by anakinra in acute gout. *Arthritis Res. Ther.* 9, R28.
- So, A., De Meulemeester, M., Pikhak, A., Yücel, A.E., Richard, D., Murphy, V., Arulmani, U., Sallstig, P., and Schlesinger, N. (2010). Canakinumab for the treatment of acute flares in difficult-to-treat gouty arthritis: Results of a multicenter, phase II, dose-ranging study. *Arthritis Rheum.* 62, 3064–3076.
- Vogl, T., Tenbrock, K., Ludwig, S., Leukert, N., Ehrhardt, C., van Zoelen, M.A., Nacken, W., Foell, D., van der Poll, T., Sorg, C., and Roth, J. (2007). Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock. *Nat. Med.* 13, 1042–1049.
- Wang, C., Klechikov, A.G., Gharibyan, A.L., Wärmländer, S.K., Jarvet, J., Zhao, L., Jia, X., Narayana, V.K., Shankar, S.K., Olofsson, A., et al. (2014). The role of pro-inflammatory S100A9 in Alzheimer's disease amyloid-neuroinflammatory cascade. *Acta Neuropathol.* 127, 507–522.
- Wang, A., Huen, S.C., Luan, H.H., Yu, S., Zhang, C., Gallezot, J.D., Booth, C.J., and Medzhitov, R. (2016). Opposing effects of fasting metabolism on tissue tolerance in bacterial and viral inflammation. *Cell* 166, 1512–1525.
- Yanamandra, K., Alexeyev, O., Zamotin, V., Srivastava, V., Shchukarev, A., Brorsson, A.C., Tartaglia, G.G., Vogl, T., Kaye, R., Wingsle, G., et al. (2009). Amyloid formation by the pro-inflammatory S100A8/A9 proteins in the ageing prostate. *PLoS ONE* 4, e5562.
- Youm, Y.H., Grant, R.W., McCabe, L.R., Albarado, D.C., Nguyen, K.Y., Ravussin, A., Pistell, P., Newman, S., Carter, R., Laque, A., et al. (2013). Canonical Nlrp3 inflammasome links systemic low-grade inflammation to functional decline in aging. *Cell Metab.* 18, 519–532.

Youm, Y.H., Nguyen, K.Y., Grant, R.W., Goldberg, E.L., Bodogai, M., Kim, D., D'Agostino, D., Planavsky, N., Lupfer, C., Kanneganti, T.D., et al. (2015). The ketone metabolite  $\beta$ -hydroxybutyrate blocks NLRP3 inflammasome-mediated inflammatory disease. *Nat. Med.* *21*, 263–269.

Yu, J.W., Wu, J., Zhang, Z., Datta, P., Ibrahimi, I., Taniguchi, S., Sagara, J., Fernandes-Alnemri, T., and Alnemri, E.S. (2006). Cryopyrin and pyrin activate

caspase-1, but not NF-kappaB, via ASC oligomerization. *Cell Death Differ.* *13*, 236–249.

Zhu, Y., Pandya, B.J., and Choi, H.K. (2011). Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007–2008. *Arthritis Rheum.* *63*, 3136–3141.

slas  
2018

INTERNATIONAL  
CONFERENCE & EXHIBITION

**FEB. 3-7** | [SLAS2018.ORG](http://SLAS2018.ORG)  
SAN DIEGO, CA  
SAN DIEGO CONVENTION CENTER

# WHERE DISCOVERY MEETS TECHNOLOGY

**SHORT COURSES:** FEBRUARY 3-4 | **CONFERENCE:** FEBRUARY 3-7 | **EXHIBITS:** FEBRUARY 5-7 | [SLAS2018.ORG](http://SLAS2018.ORG)

## Keynote Speakers



**OPENING KEYNOTE**  
Benjamin F. Cravatt



**CLOSING KEYNOTE**  
Marc Abrahams

## Educational Tracks

SLAS2018 offers participants a world-class scientific program that showcases the top life sciences discovery and technology education from industry leaders. Podium and poster presentations offer SLAS2018 participants compelling content, best practice and new perspectives on emerging scientific technologies from a broad range of industries and academic research perspectives.

The SLAS2018 scientific program features 10 educational tracks, including three tracks making their debut in San Diego:

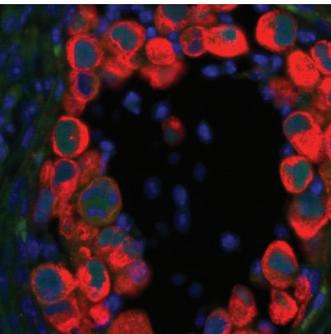
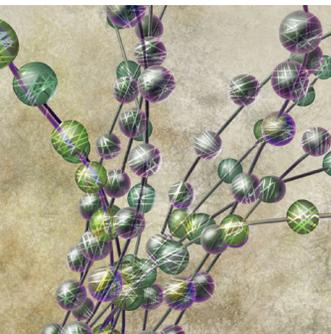
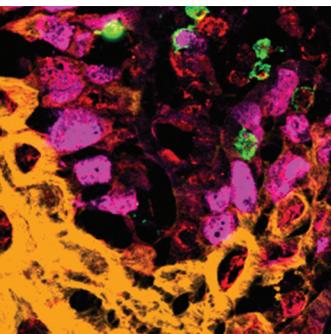
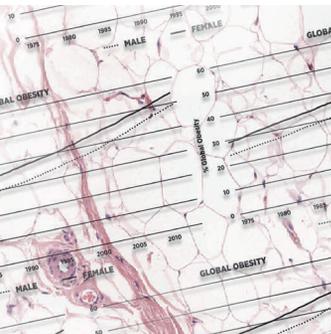
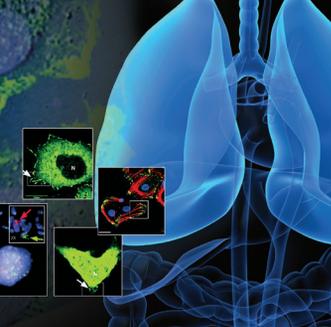
- Advances in Bioanalytics and Biomarkers
- Assay Development and Screening
- Automation and High-Throughput Technologies
- Biologics Discovery **(NEW!)**
- Cellular Technologies
- Chemical Biology **(NEW!)**
- Data Analysis and Informatics
- Drug Target Strategies
- High-Definition Biotechnology **(NEW!)**
- Micro- and Nanotechnologies

**REGISTER NOW**

Complete information about SLAS2018 can be found at [SLAS2018.ORG](http://SLAS2018.ORG).

LIFE SCIENCES

 **slas**



# 2018 SCIENTIFIC CONFERENCES

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

## **Fifth AACR-IASLC International Joint Conference: Lung Cancer Translational Science from the Bench to the Clinic**

*Conference Cochairs: Charles M. Rudin and Charles Swanton*

January 8-11, 2018 | San Diego, CA

## **Obesity and Cancer: Mechanisms Underlying Etiology and Outcomes**

*Conference Cochairs: Lewis C. Cantley, Michael N. Pollak, and Elizabeth A. Platz*

January 27-30, 2018 | Austin, TX

## **Immunobiology of Primary and Metastatic CNS Cancer: Multidisciplinary Science to Advance Cancer Immunotherapy**

*Conference Cochairs: Hideho Okada, Robyn S. Klein, Ignacio Melero, and Patricia S. Steeg*

February 12-15, 2018 | San Diego, CA

## **Third AACR-SNMMI Joint Conference on State-of-the-Art Molecular Imaging in Cancer Biology and Therapy**

*Conference Cochairs: Todd E. Peterson and David R. Piwnica-Worms*

February 14-17, 2018 | San Diego, CA

## **Targeting DNA Methylation and Chromatin for Cancer Therapy**

*Conference Cochairs: Stephen B. Baylin, Margaret A. Goodell, and Peter A. Jones*

March 1-4, 2018 | Atlanta, GA

## **AACR Annual Meeting 2018**

*Program Committee Chair: Elaine R. Mardis*  
April 14-18, 2018 | Chicago, IL

## **Cancer Dormancy and Residual Disease**

*Conference Cochairs: Julio A. Aguirre-Ghiso, Ann F. Chambers, Cyrus M. Ghajar, Christoph A. Klein, and Dorothy A. Sipkins*

June 19-22, 2018 | Montreal, QC, Canada

## **Advances in Malignant Lymphoma: Maximizing the Basic-Translational Interface for Clinical Application**

*Scientific Committee Chair: Ari M. Melnick*

June 22-26, 2018 | Boston, MA

## **Sixth JCA-AACR Special Joint Conference on the Latest Advances in Lung Cancer Research: From Basic Science to Therapeutics**

*Organizing Committee: Hiroyuki Mano, Seiji Yano, Hiroyoshi Nishikawa, Alice T. Shaw, Roy S. Herbst, and Charles M. Rudin*

July 10-12, 2018 | Kyoto, Japan

## **Intestinal Stem Cells and Colon Cancer: Biology to Therapy**

*Conference Cochairs: Anil K. Rustgi, Johanna Bendell, Hans Clevers, Christina Curtis, and Owen Sansom*

September 27-30, 2018 | Washington, DC

## **Metabolism and Cancer**

*Conference Cochairs: Ralph J. Deberardinis, Tak W. Mak, Joshua D. Rabinowitz, and M. Celeste Simon*

September 28-October 1, 2018 | New York, NY

## **Fourth CRI-CIMT-EATI-AACR International Cancer Immunotherapy Conference: Translating Science into Survival**

September 30-October 3, 2018 | New York, NY

## **Tumor Immunology and Immunotherapy**

*Conference Cochairs: James P. Allison, Lisa M. Coussens, Ira Mellman, and Drew M. Pardoll*

November 27 - 30, 2018 | Miami Beach, FL

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

**AACR** American Association  
for Cancer Research®

FINDING CURES TOGETHER®



# IT'S LIGHTING THE WAY TO A NEW LEVEL OF UNDERSTANDING

Your everyday and innovative cellular assays can reveal deep biological understanding, if you've got the sensitive imaging and intuitive data analysis to uncover it. The Operetta CLS™ high-content analysis system features a unique combination of microscopy technologies – automated water-immersion objectives, high-power 8x LED

illumination, true confocal optics, and an ultrasensitive sCMOS camera – to deliver all the speed, sensitivity, and resolution you need to reveal subcellular details. And with our simple, powerful Harmony™ 4.5 software, Operetta CLS lets you discover subtle phenotypic changes. Operetta CLS: When great technologies combine, what happens is illuminating.



Operetta CLS High-Content Analysis System

Check out the new Operetta CLS HCA System at [www.perkinelmer.com/LightTheWay](http://www.perkinelmer.com/LightTheWay)

  
**PerkinElmer**  
For the Better