

/// NEW STEMFLEX MEDIUM ///
DESIGNED FOR BETTER EVERYTHING



Enhanced flexibility and superior performance in today's applications

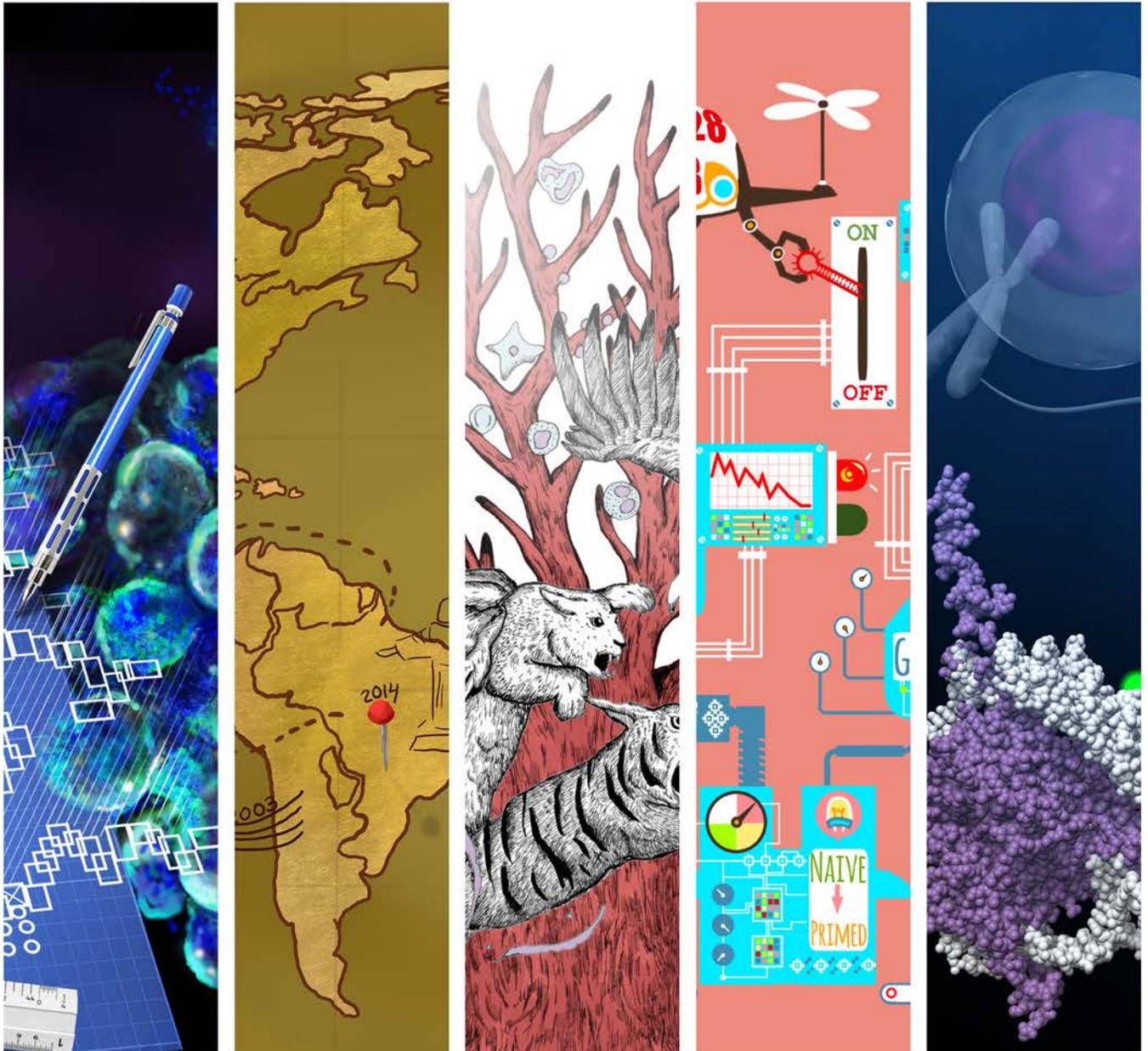
Introducing a media that works for the applications stem cell researchers are doing today, including gene editing and single-cell passaging. Gibco™ StemFlex™ Medium supports optimal, reproducible results while also letting you choose your matrix, passaging reagent, and feeding schedule.

Do what you do better at
thermofisher.com/stemflex

ThermoFisher
SCIENTIFIC

Cell Stem Cell

Best of 2016



Including an hiPSC Disease Modeling SnapShot

From the moment cells are placed in culture, the learning opportunity begins.



With the IncuCyte S3 Live-Cell Analysis System you can continuously monitor stem cell reprogramming and differentiation events - right inside your incubator.

The IncuCyte automatically acquires and analyzes images around the clock, for days, weeks, even months, enabling you to derive deeper and more physiologically relevant insights into stem cell reprogramming and differentiation processes from real-time kinetic data.

The system accommodates multiple users and vessel types simultaneously, offering a simple, yet powerful workflow to support an entire research team, **from incubation to publication.**

Learn about the benefits of real-time, automated analysis for Stem Cells and other research applications:

IncuCyte[®]
by ESSEN BIOSCIENCE

IncuCyte.com

THE FUTURE

AT YOUR

FINGERTIPS

Innovative Design Feature
ICONIC gas mixing technology

Innovative Design Feature
Removable Pop-Off™ front cover for easy set-up and cleaning

Quick and easy access with Ezee Sleeve™ Gloveless Access ports

Fast interlock purge cycle time



Dual chambers available for multiple atmospheres



Intuitive user controls and clean display

The InvivoO₂ workstation was designed and built to mimic the physiology of your subject matter, giving you the reassurance of precise results under controlled conditions. It allows you to study the most complex of cell interactions under precise physiological conditions, regulating and maintaining oxygen, carbon dioxide, temperature and humidity. Whether you're hoping to replicate the environment of blood vessels or lung tissue, the InvivoO₂ is the best tool for the job.

Get in touch today to hear how we can work with you. | U.S.: (800) 992-2537
www.bakerruskinn.com



Light re-engineered

An Employee-Owned Company

Multiline Laser Illuminator - High Power, Low Cost

LDI^{*}

- 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

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



Filter sets optimized for the LDI

www.89north.com
sales@89north.com
 +1.802.881.0302
 1.877.417.8313

www.chroma.com
sales@chroma.com
 +1.802.428.2500
 1.800.824.7662

Quantitative NIR Westerns are Just the Beginning



-  2 Lasers for NIR Detection
-  3 Visible Fluorescent Channels for Flexibility
-  Sensitive Chemiluminescent Detection



LEARN MORE ABOUT THE c600 AT
azurebiosystems.com/image | info@azurebiosystems.com



Meet Our Discovery Squad



Wes™



Milo™



Ella



Luminex® 100/200™

Luminex



MAGPIX®



FLEXMAP 3D®

Learn more | bio-techne.com/immunoassays

Foreword

Five years ago, Cell Press launched the “Best of” reprint collections across a number of our journals, including *Cell Stem Cell*. We are pleased to now bring you a new “Best of” that focuses on articles published over the course of 2016—*Cell Stem Cell*'s 10th year as a journal.

For this edition, we made our selection by looking at reviews and research articles with the most full-text HTML and PDF downloads since publication and then choosing a representative group from each of the two published volumes. This collection covers a broad range of topics and includes papers that feature some of the exciting experimental approaches—such as single-cell analysis, genome editing, reprogramming, and organoid technology—that have helped drive the field forward in recent years. It also reflects the way in which the stem cell field responded quickly to address the public health emergency posed by the Zika virus outbreak earlier last year. Additionally included is a graphical SnapShot that lays out key advances made in hiPSC disease modeling.

We recognize that no one measurement can be indicative of the “best” papers over a given period of time, especially when the articles are relatively new and their true significance may still need time to be established. Nevertheless, we hope that this combination of approaches to highlighting articles will give you a snapshot of different perspectives on the studies that we published during 2016 and over the course of our first 10 years.

To access this collection online, please visit cell.com/bestof, where you'll also be able to view the entire “Best of” Cell Press collection. To check out the latest findings that we've had the privilege to publish in the realm of stem cell research, visit cell.com/cell-stem-cell/home and follow us on Twitter [@cellstemcell](https://twitter.com/cellstemcell). Also, be sure to visit cell.com to find other high-quality papers published in the full collection of Cell Press journals.

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

Cell Stem Cell

For information for the Best of Series, please contact:

Jonathan Christison
Program Director, Best of Cell Press
e: jchristison@cell.com
p: 617-397-2893
t: @CellPressBiz



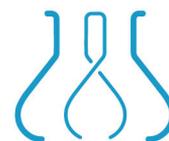
Give your cells a competitive edge

Sometimes it's the simplest of things that
make the difference for cell health.

Anyone who has ever cultured stem cells in a lab knows that it requires a tremendous investment in time, energy, and expense. You just can't leave anything to chance. That's why so many researchers choose Biological Industries' **NutriStem® hPSC XF** and **MSC NutriStem® Media**. From optimal growth and expansion to weekend-free culture, NutriStem® is designed to promote superior growth in a flexible culture format so you can achieve more reliable, consistent cell health.

- Defined, serum-free, and xeno-free
- Produced under GMP
- FDA Drug Master File (DMF) available
- Scale-up options available

Visit bioind.com to learn more.



BI
Biological Industries
Culture of Excellence

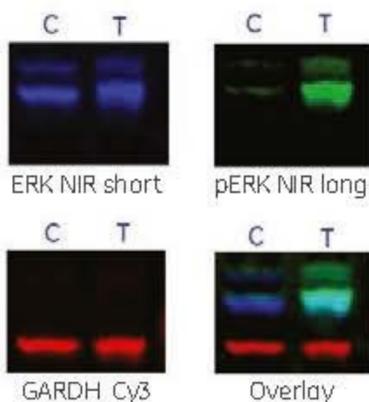




Quantify three proteins at a time with near-infrared technology

The Amersham™ Typhoon™ NIRplus offers a new level of flexibility without compromising accuracy. Designed to improve workflow and deliver significant time savings, NIRplus has a large scan area and 10 µm resolution for imaging traditional western blots and up to nine microtiter plates at one time. Additionally, the NIRplus features a three-laser configuration for simultaneous detection of three proteins in the same gel or blot.

Add value to your experiments with Amersham Typhoon NIRplus



Triplex protein detection
Distinguish among three different proteins using multifluorescence and high resolution of distinct bands.

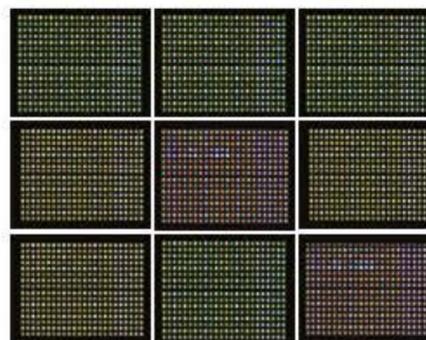


Image nine microtiter plates simultaneously
Three channel multiplex detection on 384-well microtiter plates. Amersham Typhoon NIRplus can scan up to nine microtiter plates per scan. Blue: Cy3, Green: NIR short, Red: NIR long



Nikon
100th
anniversary

BioStation. The Ultimate Live Cell Environment.



Protect Cells. Certify Cells. Expect Significant Results.

BioStation Cell Culture Observation Systems from Nikon offer perfect environmental control over cell cultures throughout the imaging process. This allows imaging experiments to be conducted in complete safety with no fear of contamination or other factors which damage live cells.

The **BioStation CT** is a cell culture imaging powerhouse, holding up to 30 culture vessels and providing complete control over temperature, humidity, and gas concentration.

The **BioStation IM-Q** is a compact cell incubation and monitoring device that allows users with minimal microscopy experience to conduct live cell imaging locally or by remote operation.

Which BioStation is right for your research?

Call us at 1-800-52-NIKON or visit
www.nikonbiostation.com

Bio Station

Imaging Solutions for Stem Cell Research



nikon-instruments-inc- nikoninstrumentsinc

Nikon Instruments Inc. ■ www.nikoninstruments.com ■ nikoninstruments@nikon.net

Cell Stem Cell

Best of 2016

SnapShot

Key Advances in hiPSC Disease Modeling

Valeria Orlova and Christine Mummery

Review

Engineering Stem Cell Organoids

Xiaolei Yin, Benjamin E. Mead, Helia Safaee, Robert Langer, Jeffrey M. Karp, and Oren Levy

Articles, Reports, and Resources

Zika Virus Infects Human Cortical Neural Progenitors and Attenuates Their Growth

Hengli Tang, Christy Hammack, Sarah C. Ogden, Zhexing Wen, Xuyu Qian, Yujing Li, Bing Yao, Jaehoon Shin, Feiran Zhang, Emily M. Lee, Kimberly M. Christian, Ruth A. Didier, Peng Jin, Hongjun Song, and Guo-li Ming

Complete Meiosis from Embryonic Stem Cell-Derived Germ Cells In Vitro

Quan Zhou, Mei Wang, Yan Yuan, Xuepeng Wang, Rui Fu, Haifeng Wan, Mingming Xie, Mingxi Liu, Xuejiang Guo, Ying Zheng, Guihai Feng, Qinghua Shi, Xiao-Yang Zhao, Jiahao Sha, and Qi Zhou

Mesenchymal Inflammation Drives Genotoxic Stress in Hematopoietic Stem Cells and Predicts Disease Evolution in Human Pre-leukemia

Noemi A. Zambetti, Zhen Ping, Si Chen, Keane J.G. Kenswil, Maria A. Mylona, Mathijs A. Sanders, Remco M. Hoogenboezem, Eric M.J. Bindels, Maria N. Adisty, Paulina M.H. Van Strien, Cindy S. van der Leije, Theresia M. Westers, Eline M.P. Cremers, Chiara Milanese, Pier G. Mastroberardino, Johannes P.T.M. van Leeuwen, Bram C.J. van der Eerden, Ivo P. Touw, Taco W. Kuijpers, Roland Kanaar, Arjan A. van de Loosdrecht, Thomas Vogl, and Marc H.G.P. Raaijmakers

Zika Virus Disrupts Neural Progenitor Development and Leads to Microcephaly in Mice

Cui Li, Dan Xu, Qing Ye, Shuai Hong, Yisheng Jiang, Xinyi Liu, Nana Zhang, Lei Shi, Cheng-Feng Qin, and Zhiheng Xu

Induction of Expandable Tissue-Specific Stem/Progenitor Cells through Transient Expression of YAP/TAZ

Tito Panciera, Luca Azzolin, Atsushi Fujimura, Daniele Di Biagio, Chiara Frasson, Silvia Bresolin, Sandra Soligo, Giuseppe Basso, Silvio Bicciato, Antonio Rosato, Michelangelo Cordenonsi, and Stefano Piccolo

Targeted Epigenetic Remodeling of Endogenous Loci by CRISPR/Cas9-Based Transcriptional Activators Directly Converts Fibroblasts to Neuronal Cells

Joshua B. Black, Andrew F. Adler, Hong-Gang Wang, Anthony M. D'Ippolito, Hunter A. Hutchinson, Timothy E. Reddy, Geoffrey S. Pitt, Kam W. Leong, and Charles A. Gersbach

(continued)

A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis

Masayuki Fujii, Mariko Shimokawa, Shoichi Date, Ai Takano, Mami Matano, Kosaku Nanki, Yuki Ohta, Kohta Toshimitsu, Yoshihiro Nakazato, Kenta Kawasaki, Toshio Uraoka, Toshiaki Watanabe, Takanori Kanai, and Toshiro Sato

CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs

Mohammad A. Mandegar, Nathaniel Huebsch, Ekaterina B. Frolov, Edward Shin, Annie Truong, Michael P. Olvera, Amanda H. Chan, Yuichiro Miyaoka, Kristin Holmes, C. Ian Spencer, Luke M. Judge, David E. Gordon, Tilde V. Eskildsen, Jacqueline E. Villalta, Max A. Horlbeck, Luke A. Gilbert, Nevan J. Krogan, Søren P. Sheikh, Jonathan S. Weissman, Lei S. Qi, Po-Lin So, and Bruce R. Conklin

De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data

Dominic Grün, Mauro J. Muraro, Jean-Charles Boisset, Kay Wiebrands, Anna Lyubimova, Gitanjali Dharmadhikari, Maaïke van den Born, Johan van Es, Erik Jansen, Hans Clevers, Eelco J.P. de Koning, and Alexander van Oudenaarden

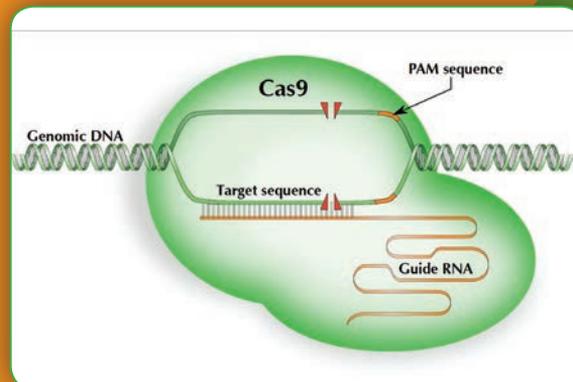


Make The Right Cut With OriGene's CRISPR/Cas9 Tools

CRISPR is revolutionizing genetics, particularly mammalian genetics.

With unprecedented efficiency, CRISPR enables precise genome disruption in almost all model organisms. RNA-guided cleavage paired with donor-guided repair allows easy introduction of any desired modification in a living cell.

OriGene offers a wide selection of ready-to-use CRISPR tools. Blue Heron, OriGene's gene synthesis division, provides quick, reliable, and confidential service for any CRISPR-related constructs.



- All-in-one gRNA/CAS9 vectors
- gRNA cloning service
- Donor cloning service
- Gene knockout kits
(pre-designed gRNAs and donor)



 **ORIGENE**

www.origene.com

ISSCR 2018

ANNUAL MEETING

**MELBOURNE
AUSTRALIA**

20-23 JUNE

ISSCR2018.ORG

**THE
GLOBAL
STEM CELL
EVENT**

**SAVE THE DATE:
20-23 JUNE 2018**

The **ISSCR Annual Meeting** brings the stem cell community together to exchange ideas, share insights and advance stem cell research and its application. Join us to engage with the latest science and technologies and envision new possibilities for the field.

Plan now for ISSCR 2018 in Melbourne, Australia.

ENVISION & ENGAGE

Co-sponsored by:



ISSCR



**INTERNATIONAL SOCIETY
FOR STEM CELL RESEARCH**



REVEAL THE POTENTIAL OF YOUR STEM CELLS

Predict Stem Cell Reprogramming and Differentiation Potential In Real-Time

Don't wait until **after** the experiment to find out your cells weren't ready to be reprogrammed or differentiated. Agilent Seahorse XF technology measures discrete changes in cellular bioenergetics in real-time, enabling researchers to predict the ability of somatic cells to reprogram, confirm pluripotency and monitor cell phenotype from pluripotent to a differentiated state. Achieve a new level of cell characterization, control and efficiency through label-free, real-time analysis.

Predict stem cell reprogramming and differentiation potential

Monitor as cells leave the pluripotent state to a differentiated phenotype

Confirm pluripotent stem cell phenotype and stability

Learn more about Agilent Seahorse XF products and Stem Cell Analysis at www.agilent.com/chem/StemCells

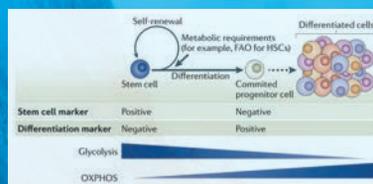


Fig 1 - Ito, K., et al. Metabolic requirements for the maintenance of self-renewing stem cells. Nat Rev Mol Cell Biol. 2014. 15: 243-56.

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

© Agilent Technologies, Inc. 2017

THE NYSCF CONFERENCE



OCTOBER 24 – 25, 2017
The Rockefeller University, New York City

NYSCF Conference Overview

The NYSCF Conference uniquely focuses on translational stem cell research to demonstrate the potential of stem cells and the power of scientists worldwide to advance novel cures for the major diseases facing the world today.

Designed for those with an interest in translational medicine, the conference is open to researchers, patient advocates, government and health officials, students, and industry representatives.

Featured Speakers

| | | | |
|------------------------|------------------------|-----------------------------|-----------------------|
| José Baselga, MD | Elaine Fuchs, PhD | Esteban Mazzoni, PhD | David Scadden, MD |
| Helen Blau, PhD | Valerie Horsley, PhD | Douglas A. Melton, PhD | Clive Svendsen, PhD |
| Pete Coffey, DPhil | Gordon Keller, PhD | Hiromitsu Nakauchi, MD, PhD | Takanori Takebe, MD |
| Valentina Fossati, PhD | Henry Klassen, MD, PhD | Malin Parmar, PhD | Irving Weissman, MD |
| Paul Frenette, MD | Tippi MacKenzie, MD | Emmanuelle Passegué, PhD | Thomas Zwaka, MD, PhD |

as of 05/08/2017

Registration

To register please visit www.nyscf.org/conference

Early registration is recommended as space is limited.

Conference registration includes all sessions.

Pricing

| | Early until 9/11 | Advanced until 10/2 | Regular from 10/3 |
|--|---------------------|------------------------|----------------------|
| Commercial / Industry | \$895 | \$895 | \$995 |
| *Healthcare Professional (non-academic) | \$400 | \$475 | \$550 |
| *Non-Profit / Government / Academic | \$325 | \$400 | \$475 |
| *Residents / Postdocs / Fellows | \$195 | \$225 | \$250 |
| *Students | \$45 | \$45 | \$55 |

*Proof of identification is required.

Call for Abstracts

Select researchers will have the chance to showcase their work at the conference poster session. Please submit abstracts electronically at nyscf.org/callforabstracts by 5:00 pm EDT on August 25, 2017 to be considered for selection.

Location

The Rockefeller University, Caspary Auditorium
1230 York Avenue (at 66th Street)
New York, NY 10065

Further NYSCF conference information, including hotel group pricing and cancellation policies, is available online.

Please visit: www.nyscf.org/conference

For questions, please contact events@nyscf.org or 212.365.7425



Snapshot: Key Advances in hiPSC Disease Modeling

Valeria Orlova and Christine Mummy

Department of Anatomy and Embryology, Leiden University Medical Centre, Einthovenweg 20, 2333ZC Leiden, The Netherlands

This SnapShot presents a timeline of key advances in directed differentiation and disease modeling using human pluripotent stem cells. The PMID for each paper is listed after the first author's name, and the papers are grouped by color codes referring to different systems. A related Review by Passier et al. (2016) in this issue of *Cell Stem Cell* provides further analysis of key advances in hiPSCs over the last 10 years.

Cell Stem Cell

2015

Lithium responses in bipolar disorder neurons (Merrens, 26524527)
 Atrial cardiomyocyte pharmacology assessment (Devata, 25700171)
 Contractile defect from MYBPC3 mutation (Ishiki, 26168474)
 Cardiomyocyte contractility depends on shape and substrate stiffness (Ribeiro, 26417073)
 Expansion and patterning of cardiovascular progenitors (Ishiki, 26192318)
 Hepatocyte and cardio-toxicity high content screening (Gimm, 26539751)
 Cardiac drug screening system (Muthur, 25748532)
 Single contracting cardiomyocytes (Feaster, 26423802)
 Bioprinted 3D mini-livers (Faulkner-Jones, 26486521)
 Cholangiocytes for disease modeling and drug validation (Sampazios, 26167629)
 Cholangiocytes (Ogawa, 26167630)
 3D cortical neurons and astrocytes (Papa, 26005811)
 Predicting neural toxicity (Sowatz, 26392547)
 Lung organoids (Dye, 25903487)
 Multi-lineage kidney organoids for modeling nephrogenesis (Ikasato, 26444236)
 CRISPR-mutant kidney organoids for disease modeling (Freedman, 26493500)
 Nephron organoids model kidney development and injury (Morizane, 26458176)
 Cell-cell interactions enhance hepatocyte maturation (Beuger, 25421237)
 Influenza in IRF7 deficiency (Ciancanelli, 25814066)
 Endothelial instruction of pancreas fate (Kao, 25601205)
 Engineered pulmonary vascular in decellularized lungs (Ren, 26380048)
 Hemogenic and arterial endothelium have distinct origins (Ibadi, 25915127)
 NOTCH1 haploinsufficiency modeling (Theobald, 25768004)
 Vasculature organ bud via mesenchymal cell condensation (Takebe, 25891906)
 Modeling Duchenne muscular dystrophy in muscle fibers (Chai, 26237517)

2014

cGMP expansion method for neural stem cells (Shaley, 24962813)
 Hyperexcitability of ALS motor neurons (Wang, 24703839)
 Automated analysis of engineered heart tissue (Stoehr, 24585781)
 Defined chemical conditions for cardiomyocyte generation (Burridge, 24930130)
 Disease modeling and drug screening for diabetic cardiomyopathy (Draebel, 25437537)
 Heart-on-chip disease modeling of Barth syndrome cardiomyopathy (Wang, 24813252)
 Epicardium (Witty, 25240927)
 Retinal pigment epithelium cell sheets for clinical application (Kamae, 24327394)
 Self-organizing kidney (Ikasato, 24335651)
 3D retina with photoreceptors (Zhong, 24915161)
 Gastric organoids (McCraken, 25363776)
 3D kidney via redefined nephron progenitors (Igauchi, 24322837)
 Lung and airway epithelium (Huang, 24291815)
 Functional β cells (Pagliuca, 25303535)
 Prediction of interindividual hepatocyte drug sensitivity (Takayama, 25385620)
 Cord blood endothelium (Prasanna, 25306246)
 Hutchinson-Gilford Progeria syndrome drug responses (Blondel, 24598761)

2013

Screen for candidate ALS therapeutics (Yang, 23602540)
 Astroglia regulate neuronal development (Yang, 23739711)
 Modeling of late-onset disease via induced aging (Miller, 24315443)
 Spinal cord astrocytes (Frydson, 23994478)
 Cardiomyocyte maturation (Nunes, 23793239)
 Drug screening for disease-specific cardiomyocyte toxicity (Liang, 23519760)
 Cerebral organoids and microcephaly (Lancaster, 23395685)
 Pre-rossette neural stem cell culture system (Ebert, 23474892)
 Drug screening and gene correction for liver disease (Choi, 23325555)
 Small molecule expansion and differentiation of hepatocytes (Shan, 23728495)
 Vascular endothelium (Adams, 24052946)
 Endothelium with limited variation (White, 23073939)
 Vasculature, functional liver from organ bud (Takebe, 23382371)
 Modeling and rescue of Williams-Beuren syndrome (Kinnear, 23283491)

2012

Drug screening for ALS (Egawa, 22855461)
 Probing sporadic and familial Alzheimer's disease (Ishai, 22278060)
 Modeling cardiac sodium channel disease (Davis, 22647976)
 Blood-brain barrier endothelium (Lippmann, 22729031)
 Vascular smooth muscle disease susceptibility (Cheung, 22282507)

2011

Schizophrenia (Blennand, 21490698)
 Treatment of rat model of Parkinson's disease (Ahee, 21576821)
 Transplantable astroglia (Kienck, 21602806)
 Optimizing cardiomyocyte differentiation (Kattman, 21285278)
 Cardiomyocyte surface markers (Eliot, 22020085; Dubois, 22020386; Uesaki, 21876760)
 Self-organizing 3D optic cup (Eiraku, 21475194)
 Anterior foregut endoderm (Green, 21358635)
 Treatment of a rat model of arterial disease (Furutani, 21830062)

2010

Cardiovascular progenitors (Moretti, 19850773; Bin, 20335662)
 Intestinal tissue (Spence, 21151107)
 Functional hepatic endoderm (Sullivan, 1987180)
 Modeling liver metabolic disorder (Rashid, 20739751)

2009

Parkinson's disease iPSCs (Solimer, 19269371)
 Spinal muscular atrophy iPSCs (Ebert, 19036894)
 Functional cardiomyocytes (Zhang, 19213953)
 Vascular and hematopoietic cells (Langeringe, 19796226; Iwaura, 19423866; Choi, 192559536)

2008

Amyotrophic Lateral Sclerosis (ALS) iPSCs (Dinos, 18669821)

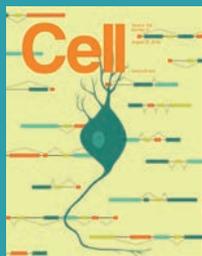
- Neuronal
- Cardiac
- Organoids
- Endoderm
- Mesoderm: vasculature | endothelium | muscle



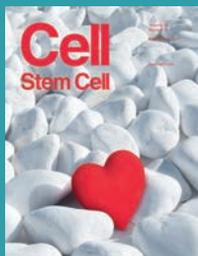
Introducing
STAR★METHODS

**Empowering methods,
to empower you.**

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



Cell



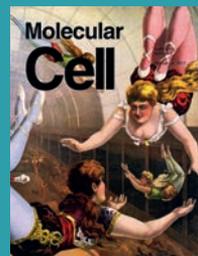
Cell Stem Cell



Cell Systems



Developmental Cell



Molecular Cell



Neuron

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

Visit www.cell.com/star-methods today!

CellPress

Engineering Stem Cell Organoids

Xiaolei Yin,^{1,2,3,4,5} Benjamin E. Mead,^{1,2,3,4,5} Helia Safaee,^{1,2,3,4} Robert Langer,^{4,5,6} Jeffrey M. Karp,^{1,2,3,4,*} and Oren Levy^{1,2,3,4,*}

¹Division of Biomedical Engineering, Department of Medicine, Center for Regenerative Therapeutics, Brigham and Women's Hospital, Cambridge, MA 02115, USA

²Harvard Medical School, Cambridge, MA 02115, USA

³Harvard Stem Cell Institute, Cambridge, MA 02138, USA

⁴Harvard - MIT Division of Health Sciences and Technology, Cambridge, MA 02139, USA

⁵David H. Koch Institute for Integrative Cancer Research

⁶Department of Chemical Engineering

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Correspondence: jeffkarp.bwh@gmail.com (J.M.K.), levyorn@gmail.com (O.L.)

<http://dx.doi.org/10.1016/j.stem.2015.12.005>

Organoid systems leverage the self-organizing properties of stem cells to create diverse multi-cellular tissue proxies. Most organoid models only represent single or partial components of a tissue, and it is often difficult to control the cell type, organization, and cell-cell/cell-matrix interactions within these systems. Herein, we discuss basic approaches to generate stem cell-based organoids, their advantages and limitations, and how bioengineering strategies can be used to steer the cell composition and their 3D organization within organoids to further enhance their utility in research and therapies.

Introduction

Model systems drive modern biological and biomedical research. These model systems aim to recapitulate body functions and processes from the molecular level to the cellular, tissue, organ, or whole organism level. The body can be viewed as a sum of a great number and wide variety of cellular and non-cellular materials formed in a highly organized manner (e.g., cell, tissue, and organ), as well as the entire interactome that includes internal (e.g., cell-cell, cell-matrix) or external (e.g., cell-environment) interactions. The hierarchical nature of all living beings suggests that multi-level recapitulation of the body could be achieved using model systems that consist of multiple cell types and their interactions (Figure 1).

Animal models most closely recapitulate in vivo human physiology, but they are limited by accessibility of imaging for observation, presence of confounding variables, limited throughput, limited usability, and differences between animal and human biology (Shanks et al., 2009). While simplistic models such as 2D monocultures of cell lines have their advantages, they often lack cell-cell and cell-matrix interactions that are required to maintain and define in situ phenotypes and thus fail to mimic cellular functions and signaling pathways present in tissues. Purified populations of primary cells also can lose their phenotype when cultured in 2D. 3D cell aggregate cultures of mesenchymal stem cells (MSCs) (Bartosh et al., 2010) or tumor cells (Vinci et al., 2012) exhibit improved function, though they lack relevant tissue organization present in vivo. Tissue explants or slices may transiently capture physiologically relevant cell organization and interactions, yet they tend to quickly lose their phenotype and are difficult to maintain for extended periods of time (Gähwiler et al., 1997). Other 3D culture systems include cell spheroids that often lack the presence of relevant stem or progenitor cell populations required to sustain the 3D culture and thus lack cells with the capacity for self-renewal and differentiation. While it is important to harness biological systems that can address specific scientific questions to achieve a balance between practica-

bility and faithfulness, most current model systems exhibit a large gap between the cellular level and the tissue/organ level.

In general, stem cells exhibit an intrinsic ability to assemble into complex structures. When placed within a hydrogel (often Matrigel) and in the presence of suitable exogenous factors, the stem cells can be coaxed into forming structures that contain organized clusters of cells. The recent availability of stem cell-derived organoid systems to provide 3D self-organized tissue models provides a compelling new class of biological model to serve as both tissue and organ proxies (Lancaster and Knoblich, 2014). Organoids recapitulate a large number of biological parameters including the spatial organization of heterogeneous tissue-specific cells, cell-cell interactions, cell-matrix interactions, and certain physiological functions generated by tissue-specific cells within the organoid. Organoids bridge a gap in existing model systems by providing a stable system amenable to extended cultivation and manipulation, while being more representative of in vivo physiology.

While a wide variety of organoids have been generated, most organoid models only represent single or partial components of a tissue, and it is often difficult to control the cell type, organization, and cell-cell or cell-matrix interactions within these systems. Bioengineers have long aspired to deconstruct biological systems and manipulate or reconstruct the system in a controlled manner. Bioengineering approaches have enabled us to steer cell behavior and cell organization, which are fundamental processes in organoid formation, and improved systems are on the horizon. In this Review, we will discuss the basic principles in the process of organoid formation, their advantages and limitations, and how bioengineering approaches can be used to increase their utility in research and therapies.

Organoids: Self-Organizing Systems of Stem Cells and Their Progeny

Organoids have been generated from both pluripotent stem cells (PSCs) and adult stem cells (ASCs) by mimicking the

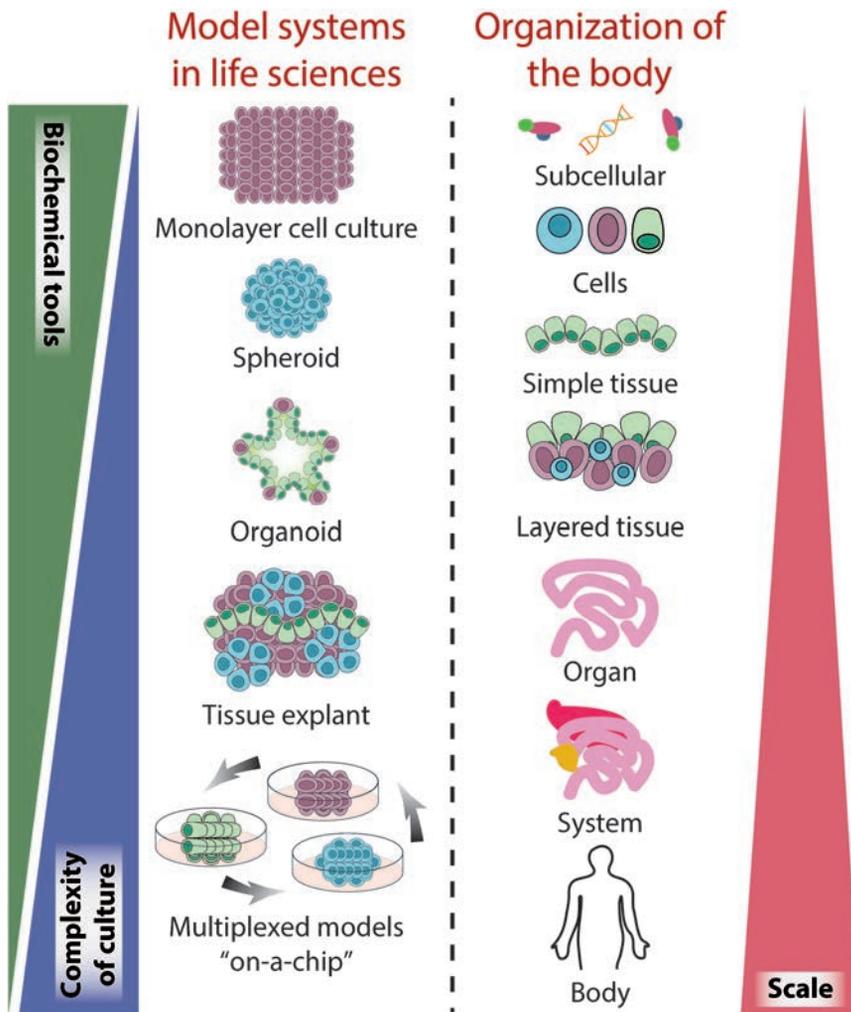


Figure 1. Model Systems in the Life Sciences

Organisms comprise a hierarchy of systems from the subcellular level to the whole body. In the life sciences, many models have been developed across this organismal hierarchy, to address specific questions across biology and medicine. Each model system possess unique attributes; in general, with increasing scale comes increasing system complexity and challenges in cell culture and the reduced availability of biochemical and quantitative tools, which can limit study insights. Organoid models provide a unique opportunity to incorporate moderate system complexity while still affording many tools for probing structure and function. When compared to tissue explants, organoid systems can mimic similar cell-cell and cell-matrix interactions while maintaining the ability for long-term cultures thanks to maintained signaling cues important for survival.

2010), liver (Huch et al., 2013), pancreas (Huch et al., 2013) and prostate (Karthaus et al., 2014). Furthermore, cancer organoids have been generated following similar approaches for tissues such as colon (Sato et al., 2011a), prostate (Gao et al., 2014), and pancreas (Boj et al., 2015).

In the process of organoid formation, a number of common factors are used to control the self-renewal and differentiation of stem cells or assist self-organization. Growth factors or small molecules are used to manipulate multiple signaling pathways important in cell survival, proliferation, and self-renewal, often in a tissue-specific manner. Paired with the biochemical cues, Matrigel is a common and important component of the system

biochemical and physical cues of tissue development and homeostasis (Lancaster and Knoblich, 2014). In a most simplified view, the development of the human body is a precisely controlled process of step-wise differentiation from the zygote and the subsequent self-organization of the cells generated in this process (Figure 2). This process can be partially reproduced when PSCs form a teratoma containing a variety of semi-organized tissues following uncontrolled differentiation and self-organization (Przyborski, 2005). Similarly, this process can be controlled in vitro with PSCs induced to differentiate down specific lineages. If provided the proper 3D scaffold and biochemical factors, differentiated cells from PSCs will self-organize to form tissue-specific organoids including the optic cup (Eiraku et al., 2011), brain (Lancaster et al., 2013), intestine (Spence et al., 2011), liver (Takebe et al., 2013), and kidney (Takasato et al., 2014). Additionally, the homeostasis of many tissues in vivo is maintained by tissue-specific ASCs through self-renewal and differentiation, followed by self-organization of the stem cells and their progeny. This process can also be reproduced in vitro under specific culture conditions to control self-renewal and differentiation, resulting in self-organized tissue organoids including intestine (Sato et al., 2009), stomach (Barker et al.,

that provides a scaffold and additional supplementation of signaling cues via basement membrane ligands to support cell attachment and survival as well as organoid formation (Xu et al., 2001). Often, organoid systems are governed by the stem cell microenvironment (or niche) they foster, which offers a point of control. The stem cell niche contains a wide range of elements including biochemical and biophysical signals, cell-cell interactions, and the extracellular matrix (ECM) interaction (Li and Xie, 2005). Current organoid systems mostly rely on intrinsic or extrinsic biochemical signals (e.g., growth factors) and cell-autonomous or cell-cell interactions to control the stem cell fate. Although these are all essential factors to control the differentiation and organization of the cells, organoid formation is highly dependent on cell-autonomous self-organization (Lancaster and Knoblich, 2014), which is not yet easily controlled.

During the establishment of organoid model systems, several bioengineering approaches that were developed in other fields including stem cell niche engineering and tissue engineering have become available to steer the behavior and organization of organoids. Here we will discuss these strategies and highlight examples of how bioengineering approaches can be used to

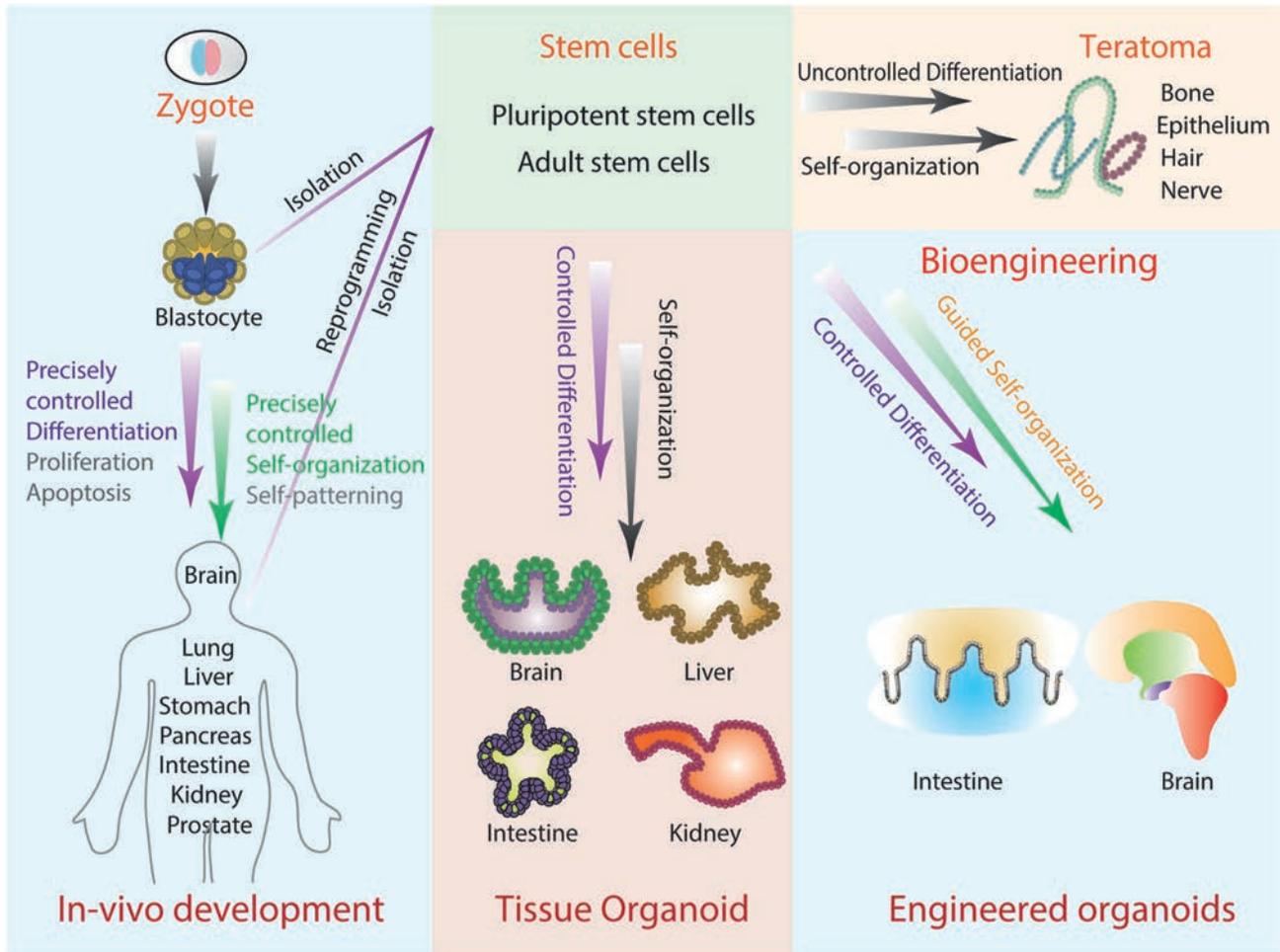


Figure 2. Organoid Development

The process of organoid formation is similar to organism development originating from a zygote and giving rise to a mature adult organism. This includes precisely controlled differentiation, proliferation, and apoptosis paired with multi-cellular self-organization and patterning, which leads to diverse mature tissues. Organoid systems are derived from ESCs (isolated from a blastocyte), iPSCs (reprogrammed from adult tissues), or ASCs (isolated from mature tissues). The driver stem cell population undergoes a similar process of culture-controlled differentiation and self-organization to give rise to tissue-specific organoids. In the case of uncontrolled differentiation (and especially following transplantation), PSCs will produce teratomas, self-organized multi-tissue tumors. Bioengineering strategies can be used to further control differentiation and organization of organoid systems to be further developed into models more representative of in vivo tissues.

increase control over organoid behavior, from tools at the sub-cellular level to those affecting system physiology.

Bioengineering Organoid Systems

Organoid systems offer one of the most promising platforms for harnessing stem cells, specifically because they are capable of recapitulating many important properties of a stem cell niche and its resulting tissue. However, like any model system, gaps between in vitro and in vivo remain, which may in part be addressed with directed bioengineering efforts. These efforts may enhance the utility of organoids in drug screening, regenerative therapy, or studies of physiological and pathological processes. Bioengineering approaches can also be applied to develop bottom-up synthetic organoid constructs, or multiplexed organoid components, enabling improved system control and the development of additional models for basic and translational stem cell niche and organoid research.

Organoid systems leverage the self-renewal and differentiation capability of stem cells and the intrinsic self-organization

ability to form organized structures. While stem cells are the “work horse” of the organoid system, the behavior of stem cells is controlled by the microenvironment. In organoids, niche components are derived by the cells (e.g., in the case of autocrine, paracrine, or juxtacrine signals) or exogenously added to the system (e.g., in the case of ECM substrates, small molecules, and growth factors). The interplay of these creates a dynamic environment in structure and function that is spatially and temporally coordinated and instructs the self-renewal/differentiation of the stem cells and self-assembly of the cells in organoids. To enhance control of organoids and further modulate the system for downstream applications, systematic engineering approaches are needed to manipulate each structural layer during the process of organoid formation. This may be achieved through modulating the cells generated by controlling stem cell self-renewal and differentiation, by directly modifying the stem cells, or by indirectly controlling them via manipulation of the microenvironment, as well as by modulating the organization of

cells in the system. Recent advances in biomaterials, micro/nanotechnology, and stem cell-driven tissue assembly have enabled extensive progress toward such systems. By combining the novel approaches from these fields, it will be possible to design microenvironments that resemble *in vivo* structure and function, giving rise to several dynamic and self-assembled organoid tissues.

Design Process

A major aim of engineering organoid systems is to improve system utility in downstream applications. Thus, effort is required to create better proxies for *in vivo* tissues and organs and improve organoid system modularity to accommodate high-throughput formats or even multi-tissue organoid compatibility in larger multiplexed systems (e.g., human-on-a-chip). However, in practice, tradeoffs must be made to optimize the design of the system for its intended application while acting within technological constraints.

The intended function of an *in vitro* organoid system dictates the design specifications. For instance, modeling developmental processes or homeostasis in tissues dictates that closely mimicking the *in vivo* tissue is top priority. In that case, recapitulating complex niche components and interactions will be essential. For expanding human cell cultures or tissues for transplantation purposes, expanding stem cells in a simplified yet efficient system will be more suitable. System design will have to account for parameters such as scalability, as well as simplifying the retrieval of cells and minimizing perturbations from dynamic interactions, to maintain the homogeneity of the system. For drug screening, the system should have high prediction power by recapitulating critical (if not all) aspects of the target *in vivo* tissue. In addition, systems should enable simple readout, a necessity for high-throughput formats. Thus, bioengineering organoid systems involves the careful selection of essential culture components (based on previous investigations concerning the specific tissue in question), as well as a thorough consideration of the objectives of the study. This way, an ideal balance among simplicity, complexity, faithfulness, and controllability can be achieved.

Engineering the Niche

Stem cell behavior *in vivo* is highly regulated through the extrinsic biochemical and biophysical signals from specialized microenvironments. These microenvironments consist of a complex array of signaling mechanisms from niche support cells, the ECM, and mechanical forces, as well as systemic and physiochemical conditions such as oxygen and pH levels. The components that make up the dynamic environment integrate sustained and rapid short-term signals to either maintain a stem cell state of quiescence, or instead, induce developmental pathways or regenerative responses. In addition to this, the niche can be remodeled and directed by its stem cell constituents. Engineering a biomimetic system that incorporates each of these signaling pathways and interactions will enable better control over the growth and differentiation of stem cells *in vitro* and enable their manipulation in accordance with the intended application.

Customized biomimetic scaffolds. ECM is one of the main components of the stem cell niche, providing structural support and mediating instructive signaling for cell polarization, retention, and mobilization (Peerani and Zandstra, 2010). ECM com-

ponents such as laminin, fibronectin, and collagen make up the physical framework of tissues and also influence cell behavior by engaging their integrin receptors (Vazin and Schaffer, 2010). Several approaches can be used to simulate the native ECM of a stem cell niche or mature tissue, including 3D scaffolds with microscale or nanoscale topography, producing customized biomaterials (Peerani and Zandstra, 2010). Alternatively, ECM scaffolds can also be produced from decellularized matrices. For example, upon reaching confluency *in vitro*, the underlying ECM matrix of bone marrow stromal cells can be decellularized and used to mimic endosteal or vascular niches along with the appropriate growth factors (Tan and Barker, 2013).

Matrigel is often a critical component of organoid culture. However, using Matrigel, which is generally not well defined, disregards the specific ECM cues required by different tissue types. Additionally, given its heterogeneous composition, it does not allow the morphogenetic processes, which are tightly governed *in vivo* by specific spatio-temporal cues, to be easily manipulated. As an alternative, essential signals from native ECMs can be incorporated into synthetic polymer matrices to produce designer ECMs, with specifically tailored compositions. For example, glycosaminoglycans such as hyaluronic acid (HA) are important ECM molecules that have shown to play an important part in modulating neural stem cell (NSC) and hematopoietic stem cell (HSC) behavior in their niche (Vazin and Schaffer, 2010), making them potentially suitable components of biomaterial-based systems for emulating NSC and HSC niche structure. Biomimetic scaffolds can be constructed from either synthetic polymers (such as polyacrylamine and polyethylene glycol [PEG]) or natural macromolecules (for instance, agarose or collagen) that can then be used to make them permissive to biological processes. For example, a biomimetic hydrogel scaffold constructed from pullulan (a polysaccharide polymer) complexed with collagen to mimic the epidermal niche ECM was able to enhance vascularization and healing when delivering MSCs to wound-sites (Rustad et al., 2012). Despite this progress, synthetic scaffolds remain primitive compared to substrates like Matrigel, and they lack the critical dynamic property of cell-driven remodeling.

Synthetic environments can be engineered to more closely replicate natural ECM by decorating bioinert matrices with signaling proteins via chemical/enzymatic crosslinking through adhesive or proteolytically cleavable sites. High-throughput screening of such platforms would enable analysis of multiple variables to determine which combination of signals is most suitable for modulating the stem cell activity and help design the synthetic ECM analogs to elicit a desired response. Cellular microarrays are capable of performing such high-throughput analysis. By tethering microenvironmental signals such as ECM components, soluble factors, and cell-cell interaction proteins to discrete locations, an array of artificial scaffolds can be generated (Gjorevski et al., 2014; Gobaa et al., 2011). Another method for providing relevant ECM support and signaling cues is microcontact printing. This technique directly deposits proteins, ECM, or cells onto a partially polymerized hydrogel substrate. This is done using a stamp often made up of poly(dimethyl siloxane) (PDMS) using soft lithography techniques (Perl et al., 2009). However, morphogenesis and organogenesis are inherently 3D

processes and thus their extrinsic regulation cannot be thoroughly understood under 2D analysis. Recently, this technology has been extended to a 3D microarray platform (Ranga et al., 2014), providing a high-throughput method of unveiling the influence of signaling proteins as well as matrix elasticity and degradability on stem cell regulation in a spatially relevant context. Achieving an ECM design that has physiologically relevant topography is also important. Nanolithography strategies such as electrospinning, electron-beam, nano-imprint lithography, and selective etching have been used to form nanofibrous substrates, nanopits, nanopillars, or nanochannels on various materials. Electron-beam lithography uses a computer-guided electron gun to scan the surface of substrates, producing patterns at nanoscale resolutions, while electrospinning is used to produce ultra-fine fibers that can form randomly oriented fibrous meshes suitable for a tissue engineering scaffold. Other methods include nano-imprint lithography (pressing a rigid mold into a layer of heated polymer) and selective etching (using a chemical etchant to roughen a surface). These strategies have been used to generate surface features that approach the natural shape and dimensions of the basement membrane fiber and pore sizes, also mimicking the porosity of natural ECM. These arrangements have been shown to support human embryonic stem cell self-renewal and control human MSC differentiation, based on the various configurations of these surfaces (Murphy et al., 2014).

Mechanical signaling applied by surrounding tissues also plays an important role in modulating cellular behavior in vivo. The absence of such forces in vitro can account for important morphological differences, such as the lack of villi formation in intestinal organoids (Gjorevski et al., 2014; Shyer et al., 2013). To dynamically tune the mechanical properties of the microenvironment, light-mediated patterning technology can be employed. A PEG hydrogel containing photolabile crosslinks can undergo local degradation when exposed to light, softening the gel. Alternatively, by including photo-initiators in the gel, shining light on specific regions will cause additional crosslinking and local stiffening (Guvendiren and Burdick, 2012). These light-dependent strategies can also be implemented for producing stiffness gradients at a microscale resolution. By photo-masking UV-crosslinked materials, the amount of light exposure can be used for reconstituting the natural stiffness variations within an artificial matrix (Vincent et al., 2013).

Adaptability to cell-induced modifications. An important step toward creating materials that can interface with cells is to engineer bioactive interfaces that offer both specificity and flexibility. To accomplish this, materials and surfaces must be rationally designed to impart specific biofunctionalities. For example, to design growth-promoting surfaces, generic cell adhesion motifs such as RGD (arginine-glycine-aspartic acid), or more cell-specific adhesion-promoting proteins and complementary receptors, can be used to engage cell integrins and promote adhesion and spreading. However, such materials must also be able to accommodate the changes brought on by the cellular activities that they are hosting.

Cells can actively modify a surface as they go through rapid and dynamic processes of cell adhesion and growth. As part of their natural growth process, cells begin to pull together and rearrange surface-bound biomolecules after adhering. On an en-

gineered surface where ligands are covalently bound, this process is impeded. To resolve this, adhesion motifs like RGD can be linked to material surfaces by long, flexible tethers, allowing the cell to pull them into clusters (Kuhlman et al., 2007). Artificial systems that incorporate carefully arranged proteins are also susceptible to enzymatic cleavage as cells try to remodel their surrounding ECM for migration or growth. It is possible to graft shorter polypeptides, which are less susceptible to proteolysis. However, an alternative approach would be to design intentionally degradable materials that take advantage of the cell-secreted proteases. For example, by crosslinking polymers such as PEG with synthetic peptides recognizable by cell proteases, the point of cleavage can be controlled. Scaffold degradation, mediated by the cell itself, can be used to promote cell migration in a pre-designed orientation (Raeber et al., 2005).

Spatio-temporal control. In addition to establishing a 3D culture scaffold that can accommodate synthetic analogs of cell-ECM interactions, it is critical to integrate other niche components for the in vitro culturing of stem cells to drive organoid formation. A major challenge in constructing an artificial scaffold in vitro is to precisely replicate the spatial presentation of signals to cells. In traditional 3D cultures, cells are flooded with biochemical signals without any spatio-temporal control. This gives rise to the major differences observed between organogenesis in vivo and in vitro. Recent advances yielded unique approaches to overcome this limitation by manipulating the ECM using light-mediated patterning. For example, by incorporating biomolecule-binding sites into a hydrogel and masking their active sites with a photo-degradable moiety, it was possible to control MSC migration within a PEG hydrogel (Gjorevski et al., 2014; Kloxin et al., 2009). The light releases the active transglutaminase factor XIII (FXIIIa) substrate, allowing ECM proteins and growth factors to be tethered within the matrix based on a light-induced pattern. This mechanism offers a powerful tool for controlling the precise pattern, location, and time when a signal can be presented to a cell.

Soluble growth factors can also be delivered using microbeads or degradable vehicles within the culturing scaffold. Nanoparticles loaded with growth factors can also be directly conjugated on the surface of the cell, slowly releasing molecules that will primarily be recaptured by the particle-carrying cells in an autocrine signaling loop (Stephan and Irvine, 2011). These methods allow some control over the release kinetics of signaling cues based on the design of the material. Bioresponsive biomaterials can also be constructed, where growth factors are released based on the release of matrix metalloproteinases from cells. PEG-based hydrogels have been fabricated with peptide sequences sensitive to proteolytic degradation, and they have been successfully used to deliver human fibroblast cells with bioactive molecules to help regenerate bone in vivo (Vazin and Schaffer, 2010). Microfluidic devices can deliver ligands as well, forming ligand gradients through the manipulation of the flow rate and the flow profile. Although this method may not readily allow macroscale architecture, it requires small sample sizes, which renders it a promising tool for high-throughput cell culture screening and analysis.

Bio-printing and bottom-up approaches to engineer organoids. To engineer organoids with controlled cell-cell and cell-matrix interactions, cells and materials can be directly deposited

onto surfaces to produce 3D co-cultures of two or more cell types with customized geometries that can be pre-defined by imaging data via CT or MRI scans of the desired tissues or organ. Bio-printing evolved from 2D inkjet printing; instead of depositing drops of black ink onto a substrate, the ink comprises a biomaterial with living cells that is precisely positioned in an additive layer-by-layer approach to create 3D biological structures that mimic the structure and function of native tissues and organs (Atala and Yoo, 2015; Murphy and Atala, 2014). Current printers can be used to create complex components of the ECM with multiple materials containing multiple cell types (Atala and Yoo, 2015; Murphy and Atala, 2014). Hydrogel-based bio-inks containing nutrients to support cell survival and function, for example, are used to place cells in 3D printers, resulting in 3D tissue-like masses that are attempts to resemble native tissues (Atala and Yoo, 2015). For example, a beating heart organoid was recently created that was able to respond to electrical and chemical cues by altering its beating patterns. Other examples of bio-printing-created organoids are human liver, muscle, and blood microvessel organoids (Atala and Yoo, 2015). Despite great progress, several challenges remain, including the ability to print in high resolution, to achieve relevant and controllable cell densities, and to achieve long-term cell functionality of the bio-inks. In addition to bio-printing, bottom-up approaches have been demonstrated to provide microscale spatial control of cell-cell interaction. By first assembling microscale cell-laden constructs individually, and then inducing controlled multi-construct organization, it is possible to assemble spatially controlled cell aggregates (Du et al., 2008). Bottom-up approaches present an interesting possibility for the construction of controlled stem cell niches or the construction of multi-tissue organoid systems.

Vascularization. Permitting sufficient nutrient and oxygen supply is an additional important consideration in developing functional *in vitro* tissue structures and organoids. *In vivo*, organs consist of hierarchically branched vascular networks and almost all cells are within a few hundred microns of a capillary to ensure a sufficient supply through diffusion. Integrating a vascular structure that allows adequate delivery of oxygen and nutrients is a necessary step in fully recapitulating larger diffusion-limited organoids. One strategy for achieving this consists of a cell-based approach, where endothelial cells are seeded within the system in order to form new blood vessels, in a process known as neoangiogenesis. The other strategy is scaffold-based, where synthetic scaffolds are used to create micro-engineered 3D structures that are tunable in geometrical, mechanical, and biological properties. Many attempts to create *in vitro* vascularization have combined these approaches. Microfluidic devices that allow uniform distribution of flow and mass transfer can be produced using soft lithographic and micro-molding processes, with polymers such as PDMS, poly-lactic (co-glycolic acid) (PLGA), and poly-glycerol sebacate (PGS). Bio-printing methods can then be used for seeding the channels with multiple types of vascular cells (Golden and Tien, 2007; King et al., 2004; Visconti et al., 2010). An additional strategy involves a modular assembly process, wherein cells are seeded in collagen constructs small enough to avoid diffusion limitations, which are then coated with endothelial cells and combined to create larger perfusion-capable structures (McGuigan and Sefton, 2006). It should be

noted that for organoid systems, these strategies will need to be modified so they would be able to integrate within the 3D macroscale tissue structures and allow perfusion and the specialized physiological functions of that tissue.

Scaffolds can also be functionalized with a combination of proangiogenic biomolecules, such as vascular endothelial growth factors (VEGFs), platelet-derived growth factors (PDGFGs), and basic fibroblast growth factors (BFGFs) for rapid formation of mature vascular networks (Richardson et al., 2001; Zisch et al., 2003). These angiogenic growth factors can not only trigger neo-vascularization, but they can also direct endothelial progenitor cell migration via gradients and promote cell assembly. Using a time-dependent release from biodegradable porous scaffolds or microparticles, these immobilized proteins can be delivered in a spatio-temporally controlled and sustained manner (Karal-Yilmaz et al., 2011; Layman et al., 2012).

Cell retrieval. Multiple strategies have emerged to release cells from surfaces with minimal impact on cell phenotype. For example, thermoresponsive polymers such as poly(*N*-isopropylacrylamide) and co-polymers such as di(ethylene glycol) methacrylate and a 9-mer oligo(ethylene glycol) methacrylate can change from a collapsed state to a hydrated, extended state (Lutz et al., 2006; Wischerhoff et al., 2008). In the extended state, the material becomes resistant to proteins, forming a layer that allows only a weak attraction between the cell and the surface, gently releasing the cells without the use of harsh chemical treatments or proteolytic cleavage.

Indeed, there is a significant need for new approaches that enable efficient retrieval of cells. For example, to produce sufficient quantities of human PSCs (hPSCs) for stem cell-based therapies, hPSCs must be rapidly and robustly expanded in culture and then collected. Similar to the example above, other synthetic polymer hydrogel systems are being developed to achieve this manufacturing goal. Poly(*N*-isopropylacrylamide)-co-poly(ethylene glycol) (PNIPAAm-PEG) is another hydrogel that possesses thermoresponsive properties. This hydrogel enables simple encapsulation and retrieval of hPSCs by going from liquid to a solid gel with temperature switches between 4°C and 37°C. This specific hydrogel matrix also demonstrates significantly increased cell expansion from 2D adherent formats, as well as a more homogenous population that retains its pluripotent phenotype (McDevitt, 2013), making it a scalable and compatible platform for manufacturing practices.

Monitoring niche components *in vitro*. Certain components of the stem cell niche, such as cell-cell interactions, oxygen distribution, local pH levels, and nutrient transport, are difficult to investigate *in vivo*, although there has been considerable recent progress. Sensors and devices that can be incorporated into the *in vitro* niche, with minimal perturbations to the natural processes, can enable precise monitoring of culturing conditions and help us better understand such parameters. Some of these capabilities have already been demonstrated *in vivo*. For example, using two-photon phosphorescence lifetime microscopy, local oxygen levels were measured in the bone marrow of live mice with micrometer spatial resolution (Spencer et al., 2014). The influence of cell-contact-mediated signaling in the stem cell niche can be further analyzed with the use of microscale devices. For example, microfabricated structures can be used to precisely control the spatial

positioning of cells and study their interactions (Hui and Bhatia, 2007).

Additionally, 3D microfluidic devices, which can be readily monitored with several imaging modalities, are able to closely imitate key physiological and structural features of small functional units of organs and provide a platform for studying such biological systems in more detail. For example, microfluidic features like microchannels can permit fluid flow at rates that are observed *in vivo*. Microfluidic devices such as organ-on-a-chip technology can also mimic peristaltic contractions and the essential function of blood vessels for delivering oxygen and nutrients, while removing waste (Bhatia and Ingber, 2014; Huh et al., 2010). Miniaturized models of functional biological units have already been fabricated on a chip, including models for lung, liver, kidney, intestine, heart, fat, bone marrow, cornea, skin, and the blood-brain barrier (see Bhatia and Ingber, 2014 for review). Organ-on-a-chip systems are amendable to high-resolution, real-time imaging as well as analysis of biochemical, genetic, and metabolic processes under conditions that closely resemble *in vivo* conditions.

Although microfluidic-based organ-on-a-chip systems can integrate key components together and still allow precise control and measurement, certain features of conventional 3D organoid culture systems may still be more advantageous. For example, traditional 3D organoid culture systems can generate more tissue mass, allowing scientists to perform analytical experiments that usually require large samples. 3D cultures also allow the growth of macroscale architecture and highly complex and spatially heterogeneous tissues that cannot be supported at the microscale. Additionally, using microfluidic chip technology poses some experimental nuances. Fabrication of a chip requires micro-engineering capabilities, and the process is susceptible to bacterial contamination and bubble formation, which will interfere with cell health and chip function and fabrication. Nevertheless, despite these setbacks, microfluidic chips still offer an unprecedented flexibility in independently controlling and monitoring features such as cell and tissue position, fluid flow, and mechanical cues, helping to dissect their contribution to tissue and organ function. Most importantly, this technology reconciles a major drawback of macroscale 3D culturing. As functionality and complexity of 3D culture systems increases, certain tissues become more inaccessible. Hence, it becomes harder to perform high-resolution imaging and to track cell activity. Organ-on-a-chip allows cells to be easily integrated with fluorescence confocal microscopy, microfluidometry, trans-epithelial electrical resistance measurements, multiple electrode arrays, and other analytical systems (Bhatia and Ingber, 2014; Huh et al., 2010).

Bioprocessing and scale-up. One of the main challenges in translating tissue engineering to the clinic is the existing bioprocessing gap including scalability and standardization issues. Scaling up the organoid platform to a production scale faces a similar hurdle, and future designs of the niche-mimetic organoid system should accommodate some degree of scale-up. Aggregate-based stirred suspension bioreactors have previously demonstrated the ability to control stem cell expansion and differentiation (Fluri et al., 2012; King and Miller, 2007; Ungrin et al., 2008). However, to further extend our control over stem cell behavior, bioreactors that can better re-create the stem cell

niche using a two-phase system can be implemented (Kirouac and Zandstra, 2008). This consists of providing both bulk signals, such as those offered as physiological conditions of pH, oxygen tension, glucose levels, and temperature, and cell-level signals, such as those from cell-cell and cell-ECM interactions. Bulk signals can be provided using the exchange of culture media and controlled via sensors and process control loops, while microenvironment signals can be controlled by scaffold design and the same techniques of protein patterning used in lab-scale cultures. For such 3D systems, which require both continuous flow and substrate interaction, scaffold porosity and permeability will be important properties to consider. Other system characteristics, such as the ease of cell retrieval, and stringent process monitoring methods for product quality control, will also need to be incorporated. Ultimately, before organoids can transition to the clinic and be manufactured on a large scale, their dynamic response to system parameters must first be understood. This includes the derivation of mathematical models that can accurately predict system behavior and the identification of adjustable input parameters and robust cell markers for confirming output quality.

Model Organoid Systems: Applying Bioengineering Approaches

Intestinal Organoids

The intestinal epithelium is an actively renewing tissue fueled by Lgr5 intestinal stem cells (ISCs) located at the bottom of the intestinal crypts (Barker et al., 2007). The self-renewal and differentiation of ISCs is cooperatively controlled by signals from the underlying mesenchyme (e.g., BMP and Wnt) as well as cells in the epithelium, specifically Paneth cells (e.g., Notch and Wnt signals) (Sato et al., 2011b). The identification of Lgr5 ISCs and the knowledge of the signals controlling ISC behavior has collectively led to the establishment of intestinal organoids, where isolated ISCs are cultured in Matrigel with conditions that permit the self-renewal and differentiation of the stem cells, followed by self-organization of the generated cells (Sato et al., 2009).

Intestinal organoids recapitulate many aspects of the intestine *in vivo* (Figure 3). Within these organoids are all major cell types of the intestinal epithelium including enterocytes, entero-endocrine cells, Paneth cells, and goblet cells. Intestinal organoids also encompass crypt-villus structures, essential cell-cell interactions including the Paneth cell-stem cell axis, and functions including absorptive and secretory activities. However, intestinal organoids do not fully recapitulate the *in vivo* epithelium, evidenced by the lack of BMP signaling gradients—the BMP inhibitor Noggin diffuses throughout the organoid from the culture media (Sato et al., 2009). Moreover, in colon and human intestinal organoid culture, Wnt proteins and other factors (e.g., Tgf- β inhibitor, p38 inhibitor, Nicotinamide, etc.) are added, which essentially prevents the differentiation of the stem cells in culture, resulting in reduced diversity of cell types (Jung et al., 2011; Sato et al., 2011a). The incorporation of spatially controlled growth factor gradients or patterned ECM represents an attractive strategy to more closely mimic physiological conditions.

In addition, modeling intestinal diseases often requires the presence of additional tissues of the intestine (i.e., immune cells and mesenchymal cells) and their interactions, which may have

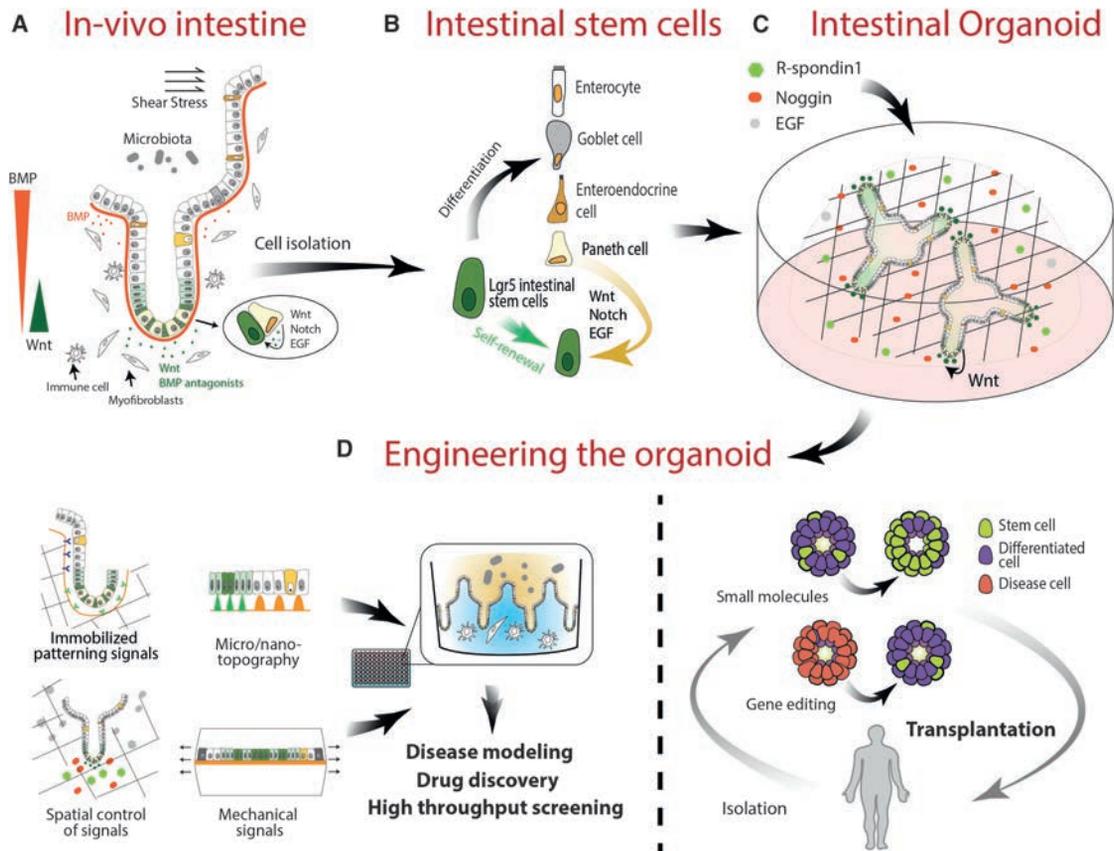


Figure 3. Intestinal Organoids

(A) The intestinal epithelium encompasses a dynamic environment in which multiple cues drive rapid and sustained tissue turnover, which is emulated in part in organoid culture. The spatial arrangement of signaling cues and neighboring cells in vivo facilitates sustained epithelial regeneration from Lgr5⁺ stem cells.

(B) Lgr5⁺ stem cells are directly responsible for the generation of terminally differentiated epithelial cells including the Paneth cell, which directly provides cues for sustaining the stem cell niche.

(C) In organoid cultures, growth factors provided by the stem cell microenvironment and surrounding mesenchyme are supplemented with exogenous factors to sustain Lgr5⁺ stem cells and the organoid.

(D) By applying multiple engineering strategies, it is possible to further emulate the in vivo intestinal epithelium to build improved systems for disease modeling, drug discovery, and screening. Additionally, it is possible to use directed gene editing or small molecule treatment to drive organoids down particular paths of enrichment, thereby enabling the potential for organoid transplantation therapy.

important disease implications (Lindemans et al., 2015). Direct access to the luminal compartment is also needed for studies on drug absorption, or microbe-epithelium interactions (Wilson et al., 2015). Such studies would also require a continuous and intact epithelial layer, which is absent in the organoid system. By combining spatial-temporal control of signals presented to the cells, and guided cell organization with structured scaffolds, these important targets could be incorporated. Directed organization of the cells will also have the potential to introduce additional cell types from lineages other than the epithelium, including immune and mesenchymal cells, enabling improved disease modeling.

Also of interest is the direct use of intestinal organoids as therapy. Intestinal organoids have been transplanted into damaged colon for tissue repair (Yui et al., 2012), but only with limited engraftment success. For this purpose, highly efficient expansion of a pure population of ISCs in a biochemically defined system (free of Matrigel) would be the goal, along with delivery techniques to improve targeting of ulcers, graft survival, and

engraftment. Recently we have shown that small molecules can be used to significantly increase the expansion efficiency of ISCs (Yin et al., 2014).

Brain Organoids

The human brain embodies biological system complexity, and its development involves a high degree of coordination between the NSCs and the dynamic niche in which they exist. Through providing different levels of morphogens (i.e., BMP, Wnt, Shh, RA, and FGF), PSCs can be induced to differentiate into many different neural subtypes, such as cortical pyramidal neurons (Espuny-Camacho et al., 2013), midbrain dopaminergic neurons (Chambers et al., 2009; Lee et al., 2000; Perrier et al., 2004; Yan et al., 2005), and spinal cord motor neurons (Dimos et al., 2002; Li et al., 2005; Soundararajan et al., 2006; Wichterle et al., 2002) (see Petros et al., 2011 for review). Furthermore, the SFEBq (serum-free floating culture of embryoid body-like aggregates with quick re-aggregation) protocol has been used to generate the more complex architectures, such as sub-brain regions like the cerebral cortex (Danjo et al., 2011; Eiraku et al., 2008;

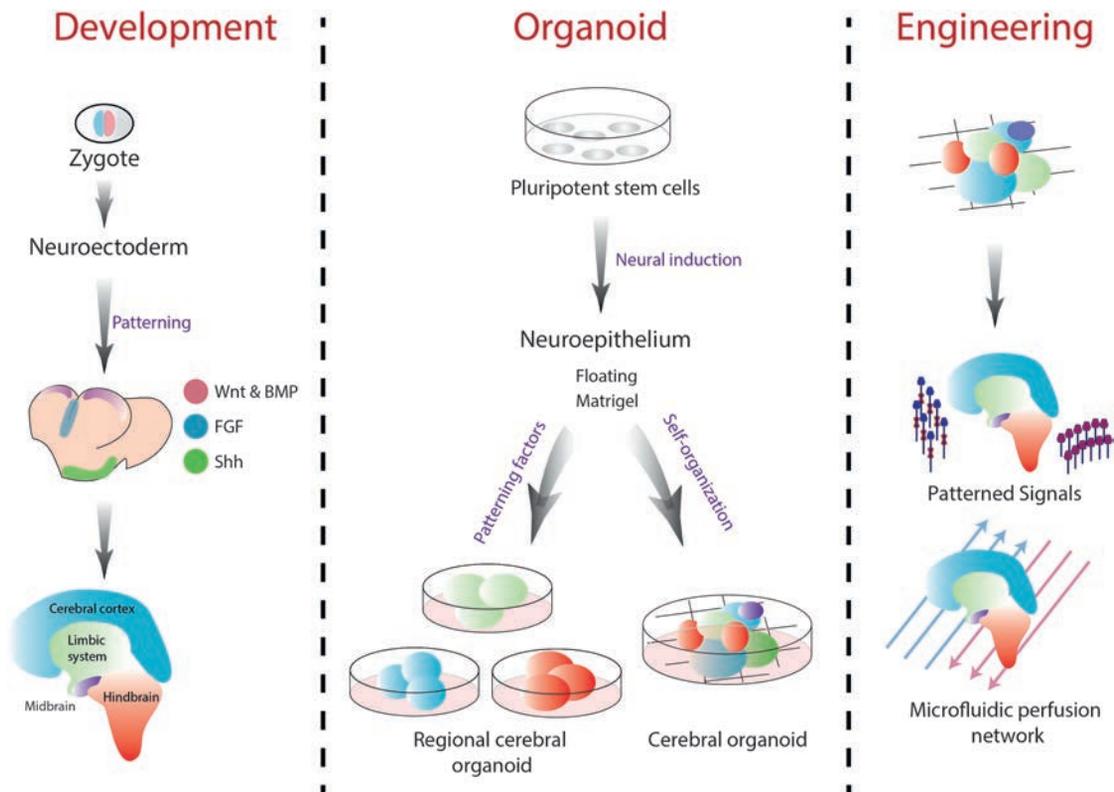


Figure 4. Brain Organoids

Brain organoids follow a modified path of in vivo development, where ESCs develop into the brain structure following specific spatio-temporal cues, ultimately leading to the mature brain structure. In vitro this process begins with either ESCs or iPSCs, and following structural and biochemical cues, multiple organoid lineages can be produced. This includes the self-organized multi-regional organoids or the patterned induction of organoids mimicking specific regions of the brain. A promising set of tools including customized and responsive biomaterials, patterned signaling, and microfluidic networks can be applied to further refine the spatial development of brain organoids.

Kadoshima et al., 2013; Mariani et al., 2012) and the pituitary (Suga et al., 2011). Alternatively, Lancaster et al. reported a culture system to generate heterogeneous neural organoids that contained multiple, but interdependent, brain regions within individual organoids (Lancaster et al., 2013). Here, the generated neuroectodermal tissues were maintained in 3D Matrigel for further expansion, without the addition of neural inducing or patterning factors. When transferred to a spinning bioreactor as Matrigel droplets, the cerebral organoids also showed enhanced nutrient absorption and grew as large as 4 mm in 2 months, and they generated distinct brain regions such as the dorsal cortex, ventral telencephalon, choroid plexus, hippocampus, and retina.

Although the cerebral organoid system has achieved some ability to model human brain development, several limitations still exist. Specifically, owing to the absence of surrounding tissue (and tissue cross-communication) and body axis, the current models are not organized to form the brain shape and structure as they exist in vivo. It is also worth noting that although culturing without any patterning factors has improved the developed cerebral organoids, patterning factors are necessary to ensure some level of controlled tissue organization. Bioengineering approaches such as spatial-temporal control of differentiation or cell patterning signals using customized scaffolds with immobilized signals, or signal gradient formation using control-released

particles or microfluidics, will have the potential to guide the differentiation and patterning of brain regions in the organoids. Nutrient and oxygen delivery limits the size of cerebral organoids and may lead to undesired differentiation in regions of poor supply, despite the addition of agitation. This also contributes to stochastic growth patterns and limited maturation of key cell types in the brain organoids (Chambers et al., 2013). A potential solution exists in implementing co-cultures that can vascularize brain organoids or implementing microfluidic perfusion networks (Figure 4).

Future Directions: Bioengineering Strategies to Advance Organoid-Based Therapies
Targeted Genome Editing: Enabling Tighter Control of the Stem Cell Niche

A major limitation to constructing more sophisticated organoids in vitro is that current control of driver stem cell fate and behavior is primarily achieved through exogenous signals. Targeted genome editing, especially via CRISPR or TALEN technology, is an extremely powerful tool to accurately manipulate endogenous genes in clinically relevant cells and organisms. This approach can be used to enhance niche function, in which we can not only steer cell fate through exogenous niche components, but also reprogram the internal decision-making structure of driver and support cells. Reprogramming cells in organoid

systems would enable user-defined “training” in how these structures self-assemble and self-regulate, driving improved studies of niche environments and their impact on organoid formation and function.

For instance, future studies may enable organoids derived from patients with single-gene hereditary diseases, such as cystic fibrosis (CF) or beta thalassemia, to undergo CRISPR genome editing to correct the mutation and then undergo transplantation as functional organ-like units back into patients to advance tissue repair and functionality. A proof-of-concept use of genome editing in the organoid field was recently demonstrated by Schwank et al. (2013). First, efficient genome editing of human stem cells in primary intestinal organoids via CRISPR/Cas9 was demonstrated. Then, using CRISPR-Cas9-mediated homology-directed repair (HDR), intestinal organoids from two CF patients were genome edited to correct the mutation (deletion of phenylalanine at position 508) of the CF transmembrane conductor receptor (CFTR), the primary cause of the disease. It was demonstrated that the genome-edited intestinal organoid systems expressed the corrected CFTR allele, which also produced fully functional proteins (involved in chloride ion channels) in these organoids. This study demonstrates genome editing of organoids as a potential gene therapy strategy, with limited risks to off-target tissue mutagenesis. Another example is microvillus inclusion disease (MVID), in which patients display microvillus inclusion and loss of brush-border microvilli, resulting in life-threatening persisting diarrhea. Mutations in myosin Vb and Syntaxin were demonstrated to cause classic and variant MVID, respectively. Establishing patient-derived intestinal organoids to be genome edited to correct those mutations and then re-transplanted into the patients may be useful in improving intestinal function. We envision that in the near future, stem cells from patients with hereditary diseases will be used to establish organoids, which will then undergo direct genome editing and be transplanted as autologous therapy (Yui et al., 2012), correcting the tissue-specific functional defects. Such a strategy may provide a cure to currently incurable hereditary diseases such as CF.

Alternatively, tools for precise gene editing in organoids can be utilized for elucidation of signaling pathways responsible for disease development. By using CRISPR gene editing, Matano et al. introduced multiple mutations in niche signaling pathways into human intestinal epithelial organoids, which were cultured *in vitro* and then transplanted into mice so their role in tumor progression and micrometastases could be studied (Matano et al., 2015). Overall, the rapidly evolving genome-editing approaches possess immense promise in creating next-generation organoid culture systems, advancing the study of organogenesis, pathogenesis, and drug-screening-based and organoid-based therapies. In this case, the organoid system paired with controlled editing of niche direction presented an ideal platform for the study of the stem cell niche in tumor generation and propagation.

Genetic Circuits: Programming Organoids

A major need in organoid research is to control the response of the niche cells to changing stimuli—this applies for *in vitro* organoid systems as well as for organoids under *in vivo* settings (for example, upon transplantation). Synthetic biology is emerging as a promising field with great potential for developing the next gen-

eration of therapeutics and diagnostics (Purcell and Lu, 2014). Relying on basic molecular biology components, artificial gene circuits have produced programmable and responsive systems within living cells. With continuous improvements of these biological components, as well as the construction of higher-order devices such as switches, memory elements, cascades, time-delayed circuits, oscillators, and logic gates, artificial gene circuits have achieved sophisticated cellular computational capabilities in both single-cell and multi-cellular systems (Cheng and Lu, 2012). Transcriptional regulation, cellular memory storage, and integration of logic gates are all tools that enable the complex computational abilities of genetic circuits, and they can be incorporated into organoid systems for an additional layer of control over their behavior during culture and after transplantation. For example, organoids could be programmed to reach tissue-like homeostasis, or execute diverse functions from the multiple terminal cell populations contained within.

Many synthetic circuits have been designed using digital logic gates, relying on transcriptional control using activators, repressors, and other novel mechanisms (Lohmueller et al., 2012). Furthermore, synthetic biologists are now able to connect many single logic gates into multiple configurations to achieve more sophisticated genetic programs. One successful approach to interconnecting logic gates is to use diffusible signaling molecules. AHL (a diffusible quorum-sensing molecule) was used to connect logic gates between *Escherichia coli* cells, producing a complex network of light-sensitive “edge-detectors” (Tabor et al., 2009), an achievement that can potentially one day be incorporated in organoid systems to direct cellular organization. While unwanted crosstalk among the synthetic devices, signaling delays between layered circuits, and other limitations may limit the exclusive use of digital gates (Purcell and Lu, 2014), integrating both analog and digital processing may eventually achieve efficient cellular computation by reducing the size of genetic circuits and thus alleviating the cellular burden.

Although performed thus far mostly in bacteria, such as *E. coli*, the field of synthetic biology is continuously advancing, and soon mammalian cells, including niche-establishing PSCs, will be successfully engineered using similar synthetic biology approaches. We envision that genetic circuit-based coordination between multiple cell types has the potential to revolutionize organoid engineering. Using genetic circuits to engineer multiple types of responsive cells, an organoid structure may be generated rapidly upon administration of specific cues. *In vitro* niche designs can also be simplified by using reprogrammed cells that produce their own signals. As organoids grow in size, it will be possible to initiate an autocrine signaling pathway for cells that are inaccessible to exogenous signals. Applying logic gates in these genetic systems will help to fine-tune such responses. Furthermore, organoids may be programmed to respond to specific signals upon their transplantation (i.e., differentiation of cells in the organoids based on a decline of a specific hormone or the presence of other cues during the course of disease). These organoids could gauge information about their local environment and have a pre-programmed built-in response based on their intended function (e.g., therapeutic), further advancing organoid-based therapies.

In a system such as an organoid, multiple niche interactions take place in a highly dynamic and transient manner. Orchestrating the delivery of such a complex array of signaling

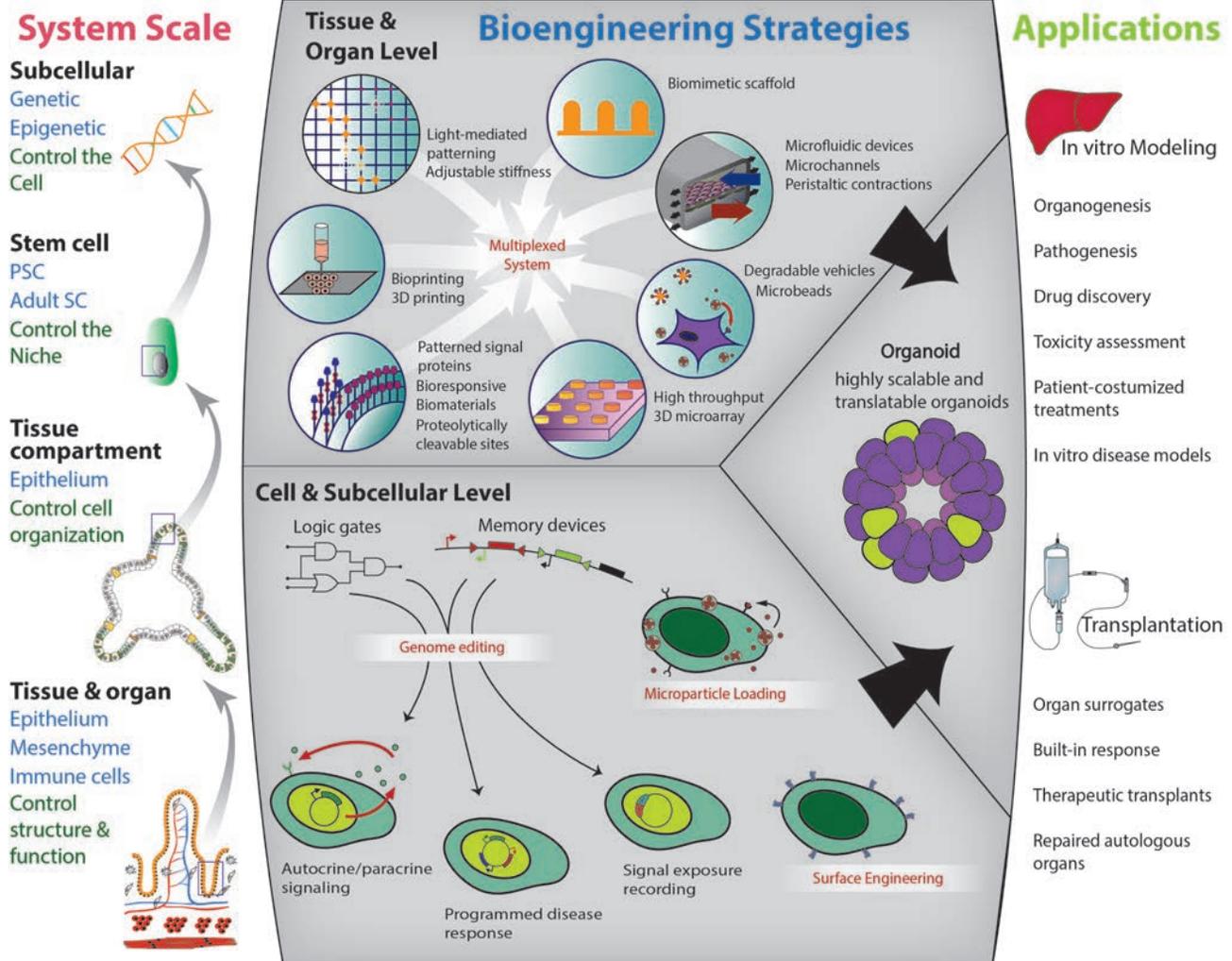


Figure 5. Bioengineering Approaches to Advance Organoid-Based Research and Therapy

With the right combination and sequence of input signals that the niche relays to the cell, it is possible to obtain the desired output (i.e., the in vitro disease model or tissue-specific organoid) and there are multiple bioengineering tools that can be harnessed to modify these signals and monitor relevant responses. Based on the same principle, following elucidation of organoid biology, there is potential to harness new knowledge to create synthetic niches. The niche can be engineered by combining multiple bioengineering techniques that mimic specific niche components (e.g., biomimetic scaffolds, tunable stiffness, appropriate topography, and spatio-temporally controlled signaling cues). The stem cells used to seed the organoid culture can also be engineered. In addition to exogenous signaling mechanisms, cell activity can be controlled through genome editing and surface modifications including drug delivery nano/microparticles. Using these methodologies, we can gain better control over the organoid to maximize functionality and sustainability in culture and ideally more closely mimic in vivo biology.

cues in vitro poses a significant engineering challenge. Applying cellular memory devices within niche cells can help to alleviate this problem. Cellular memory is an essential function that enables the storage of otherwise transient responses. A key development in the field of synthetic biology has been the creation of sophisticated memory devices, which are able to mimic cellular memory via genomically integrated circuits. Certain synthetic memory networks use positive feedback loops to achieve stable states and switch between these states based on repressive or activating inputs (Purcell and Lu, 2014). Other approaches involve the recombinase-catalyzed reconfiguration of DNA (recently demonstrated in *E. coli*), which can serve as both a method of implementing logic and a way to record information by embedding a memory of the received inputs into the DNA—a stable storage medium that persists even after cell death (Siuti

et al., 2013). This platform can also be used for creating biosensors that record the history of cellular exposure to either individual or a sequence of environmental signals (Siuti et al., 2013). Such tools can be applied to niche engineering, as a way to record and analyze niche input or to program cellular response. For example, by storing certain signaling cues in the cellular memory, it is possible to simplify the delivery of exogenous signals and apply them in sequence rather than simultaneously. Additionally, because organoid systems arise from clonal stem cell populations, the details of circuit construction are simplified to targeting the “driver” stem cells, and pairing circuitry with the transcriptional programs that naturally drive differentiation within organoids to obtain multiple circuits acting in synchrony. Also, for large-scale culturing processes that require a continuous supply of growth factors, memory devices can be used to

activate the desired response with singular/discreet exposure, cutting manufacturing costs. Future transplantation of memory-equipped organoids can also be extremely useful to increase organoid survival and maintain an extended organoid response in the body post-transplantation. For example, organoid systems can be “trained” to bypass certain physiological conditions upon transplantation, ignoring an anticipated sequence of signals based on a pre-programmed response.

Conclusions

The use of organoid platforms has led to advancements in vitro organogenesis and disease modeling, and subsequently, it has created exciting possibilities for the development of innovative new therapies. Important characteristics of multiple tissues and organ sub-regions, such as the formation of distinct brain regions of the dorsal cortex, ventral telencephalon, choroid plexus, or hippocampus and crypt-villus structures similar to those in the intestinal epithelial lining have been successfully recapitulated in 3D organoid models. In addition to the development of new biological models and tools for studying and manipulating tissue regeneration, organoid models can be used to model disease states, and they can also potentially be used to develop more predictive drug screening platforms and patient-specific treatments.

The utility of existing spheroid and tissue explant cultures can be attributed to their ability to mimic the complex niche interactions present in situ. Organoids provide a more advanced in vitro tool that enables more physiologically relevant experiments to be performed that cannot be conducted in animals or people. With the currently available extensive arsenal of bioengineering methods, it is possible to extend the utility of organoids with improved control over external cues and with an unprecedented opportunity to monitor and manipulate cellular behavior. Influencing how cells internally process exogenous signals offers a new layer of control, allowing the fine-tuning of organoids with genome editing and genetic circuits (Figure 5).

However, this is not without its challenges. While it is clearly important to not only continue to expand this library of bioengineering methods and determine how best to partner relevant technologies with unmet needs in the creation and manipulation of organoids, moving from in vitro cell monolayer model systems, which can be readily standardized, to multi-phenotype models that are combined with diverse bioengineering tools will create a standardization nightmare. This will undoubtedly make it difficult for groups to compare results between systems, as has been observed for products derived from iPSCs. It is also critical to consider challenges associated with using bioengineering strategies without interfering with the system's natural ability to individually tune each parameter, and to consider challenges to simultaneously and simply control organoids from micro-to-macroscopic levels and to maintain them in culture. Regardless, organoid systems have already found utility in many basic biological and therapeutic experiments to advance new knowledge and to advance us closer to therapies for diseases that previously appeared untouchable.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant HL095722 to J.M.K., NIH grant DE013023 to R.S.L. A Kenneth Rainin Foundation Break-

through Award to J.M.K., a Movember Prostate Cancer Foundation Challenge Award to J.M.K., The Charles A. King Trust Fellowship to O.L., The William Randolph Hearst BWH DOM Young Investigator Award to O.L.; and an NSF fellowship to B.E.M.

REFERENCES

- Atala, A., and Yoo, J.J. (2015). *Essentials of 3D Biofabrication and Translation* (Elsevier Science).
- Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., and Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* *449*, 1003–1007.
- Barker, N., Huch, M., Kujala, P., van de Wetering, M., Snippert, H.J., van Es, J.H., Sato, T., Stange, D.E., Begthel, H., van den Born, M., et al. (2010). *Lgr5*(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. *Cell Stem Cell* *6*, 25–36.
- Bartosh, T.J., Ylöstalo, J.H., Mohammadipoor, A., Bazhanov, N., Coble, K., Claypool, K., Lee, R.H., Choi, H., and Prockop, D.J. (2010). Aggregation of human mesenchymal stromal cells (MSCs) into 3D spheroids enhances their antiinflammatory properties. *Proc. Natl. Acad. Sci. USA* *107*, 13724–13729.
- Bhatia, S.N., and Ingber, D.E. (2014). Microfluidic organs-on-chips. *Nat. Biotechnol.* *32*, 760–772.
- Boj, S.F., Hwang, C.I., Baker, L.A., Chio, I., Engle, D.D., Corbo, V., Jager, M., Ponz-Sarvisé, M., Tiriác, H., Spector, M.S., et al. (2015). Organoid models of human and mouse ductal pancreatic cancer. *Cell* *160*, 324–338.
- Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* *27*, 275–280.
- Chambers, S.M., Tchieu, J., and Studer, L. (2013). Build-a-brain. *Cell Stem Cell* *13*, 377–378.
- Cheng, A.A., and Lu, T.K. (2012). Synthetic biology: an emerging engineering discipline. *Annu. Rev. Biomed. Eng.* *14*, 155–178.
- Danjo, T., Eiraku, M., Muguruma, K., Watanabe, K., Kawada, M., Yanagawa, Y., Rubenstein, J.L., and Sasai, Y. (2011). Subregional specification of embryonic stem cell-derived ventral telencephalic tissues by timed and combinatorial treatment with extrinsic signals. *J. Neurosci.* *31*, 1919–1933.
- Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Golland, R., et al. (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* *321*, 1218–1221.
- Du, Y., Lo, E., Ali, S., and Khademhosseini, A. (2008). Directed assembly of cell-laden microgels for fabrication of 3D tissue constructs. *Proc. Natl. Acad. Sci. USA* *105*, 9522–9527.
- Eiraku, M., Watanabe, K., Matsuo-Takasaki, M., Kawada, M., Yonemura, S., Matsumura, M., Wataya, T., Nishiyama, A., Muguruma, K., and Sasai, Y. (2008). Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell* *3*, 519–532.
- Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* *472*, 51–56.
- Espuny-Camacho, I., Michelsen, K.A., Gall, D., Linaro, D., Hasche, A., Bonnefont, J., Bali, C., Orduz, D., Bilheu, A., Herpoel, A., et al. (2013). Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. *Neuron* *77*, 440–456.
- Fluri, D.A., Tonge, P.D., Song, H., Baptista, R.P., Shakiba, N., Shukla, S., Clarke, G., Nagy, A., and Zandstra, P.W. (2012). Derivation, expansion and differentiation of induced pluripotent stem cells in continuous suspension cultures. *Nat. Methods* *9*, 509–516.
- Gähwiler, B.H., Capogna, M., Debanne, D., McKinney, R.A., and Thompson, S.M. (1997). Organotypic slice cultures: a technique has come of age. *Trends Neurosci.* *20*, 471–477.

- Gao, D., Vela, I., Sboner, A., Iaquinta, P.J., Karthaus, W.R., Gopalan, A., Dowling, C., Wanjala, J.N., Undvall, E.A., Arora, V.K., et al. (2014). Organoid cultures derived from patients with advanced prostate cancer. *Cell* **159**, 176–187.
- Gjorevski, N., Ranga, A., and Lutolf, M.P. (2014). Bioengineering approaches to guide stem cell-based organogenesis. *Development* **141**, 1794–1804.
- Gobaa, S., Hoehnel, S., Roccio, M., Negro, A., Kobel, S., and Lutolf, M.P. (2011). Artificial niche microarrays for probing single stem cell fate in high throughput. *Nat. Methods* **8**, 949–955.
- Golden, A.P., and Tien, J. (2007). Fabrication of microfluidic hydrogels using molded gelatin as a sacrificial element. *Lab Chip* **7**, 720–725.
- Guvendiren, M., and Burdick, J.A. (2012). Stiffening hydrogels to probe short- and long-term cellular responses to dynamic mechanics. *Nature Commun.* **3**, 792.
- 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.
- Huh, D., Matthews, B.D., Mammoto, A., Montoya-Zavala, M., Hsin, H.Y., and Ingber, D.E. (2010). Reconstituting organ-level lung functions on a chip. *Science* **328**, 1662–1668.
- Hui, E.E., and Bhatia, S.N. (2007). Micromechanical control of cell-cell interactions. *Proc. Natl. Acad. Sci. USA* **104**, 5722–5726.
- Jung, P., Sato, T., Merlos-Suárez, A., Barriga, F.M., Iglesias, M., Rossell, D., Auer, H., Gallardo, M., Blasco, M.A., Sancho, E., et al. (2011). Isolation and in vitro expansion of human colonic stem cells. *Nat. Med.* **17**, 1225–1227.
- Kadoshima, T., Sakaguchi, H., Nakano, T., Soen, M., Ando, S., Eiraku, M., and Sasai, Y. (2013). Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. *Proc. Natl. Acad. Sci. USA* **110**, 20284–20289.
- Karal-Yilmaz, O., Serhatli, M., Baysal, K., and Baysal, B.M. (2011). Preparation and in vitro characterization of vascular endothelial growth factor (VEGF)-loaded poly(D,L-lactic-co-glycolic acid) microspheres using a double emulsion/solvent evaporation technique. *J. Microencapsul.* **28**, 46–54.
- Karthaus, W.R., Iaquinta, P.J., Drost, J., Gracanin, A., van Boxtel, R., Wongvipat, J., Dowling, C.M., Gao, D., Begthel, H., Sachs, N., et al. (2014). Identification of multipotent luminal progenitor cells in human prostate organoid cultures. *Cell* **159**, 163–175.
- King, J.A., and Miller, W.M. (2007). Bioreactor development for stem cell expansion and controlled differentiation. *Curr. Opin. Chem. Biol.* **11**, 394–398.
- King, K.R., Wang, C.C.J., Kaazempur-Mofrad, M.R., Vacanti, J.P., and Borenstein, J.T. (2004). Biodegradable microfluidics. *Adv. Mater.* **16**, 2007.
- Kirouac, D.C., and Zandstra, P.W. (2008). The systematic production of cells for cell therapies. *Cell Stem Cell* **3**, 369–381.
- Kloxin, A.M., Kasko, A.M., Salinas, C.N., and Anseth, K.S. (2009). Photodegradable hydrogels for dynamic tuning of physical and chemical properties. *Science* **324**, 59–63.
- Kuhlman, W., Taniguchi, I., Griffith, L.G., and Mayes, A.M. (2007). Interplay between PEO tether length and ligand spacing governs cell spreading on RGD-modified PMMA-g-PEO comb copolymers. *Biomacromolecules* **8**, 3206–3213.
- Lancaster, M.A., and Knoblich, J.A. (2014). Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* **345**, 1247125.
- Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurler, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379.
- Layman, H., Li, X., Nagar, E., Vial, X., Pham, S.M., and Andreopoulos, F.M. (2012). Enhanced angiogenic efficacy through controlled and sustained delivery of FGF-2 and G-CSF from fibrin hydrogels containing ionic-albumin microspheres. *J. Biomater. Sci. Polym. Ed.* **23**, 185–206.
- Lee, S.H., Lumelsky, N., Studer, L., Auerbach, J.M., and McKay, R.D. (2000). Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. *Nat. Biotechnol.* **18**, 675–679.
- Li, L., and Xie, T. (2005). Stem cell niche: structure and function. *Annu. Rev. Cell Dev. Biol.* **21**, 605–631.
- Li, X.J., Du, Z.W., Zarnowska, E.D., Pankratz, M., Hansen, L.O., Pearce, R.A., and Zhang, S.C. (2005). Specification of motoneurons from human embryonic stem cells. *Nat. Biotechnol.* **23**, 215–221.
- Lindemans, C.A., Calafiore, M., Mertelsmann, A.M., O'Connor, M.H., Dudakov, J.A., Jenq, R.R., Velardi, E., Young, L.F., Smith, O.M., Lawrence, G., et al. (2015). Interleukin-22 promotes intestinal-stem-cell-mediated epithelial regeneration. *Nature*, in press. Published online December 9, 2015. <http://dx.doi.org/10.1038/nature16460>.
- Lohmueller, J.J., Armel, T.Z., and Silver, P.A. (2012). A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic Acids Res.* **40**, 5180–5187.
- Lutz, J.-F., Akdemir, O., and Hoth, A. (2006). Point by point comparison of two thermosensitive polymers exhibiting a similar LCST: is the age of poly(NIPAM) over? *J. Am. Chem. Soc.* **128**, 13046–13047.
- Mariani, J., Simonini, M.V., Palejev, D., Tomasini, L., Coppola, G., Szekely, A.M., Horvath, T.L., and Vaccarino, F.M. (2012). Modeling human cortical development in vitro using induced pluripotent stem cells. *Proc. Natl. Acad. Sci. USA* **109**, 12770–12775.
- Matano, M., Date, S., Shimokawa, M., Takano, A., Fujii, M., Ohta, Y., Watanabe, T., Kanai, T., and Sato, T. (2015). Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* **21**, 256–262.
- McDevitt, T.C. (2013). Scalable culture of human pluripotent stem cells in 3D. *Proc. Natl. Acad. Sci. USA* **110**, 20852–20853.
- McGuigan, A.P., and Sefton, M.V. (2006). Vascularized organoid engineered by modular assembly enables blood perfusion. *Proc. Natl. Acad. Sci. USA* **103**, 11461–11466.
- Murphy, S.V., and Atala, A. (2014). 3D bioprinting of tissues and organs. *Nat. Biotechnol.* **32**, 773–785.
- Murphy, W.L., McDevitt, T.C., and Engler, A.J. (2014). Materials as stem cell regulators. *Nat. Mater.* **13**, 547–557.
- Peerani, R., and Zandstra, P.W. (2010). Enabling stem cell therapies through synthetic stem cell-niche engineering. *J. Clin. Invest.* **120**, 60–70.
- Perl, A., Reinhoudt, D.N., and Huskens, J. (2009). Microcontact Printing: Limitations and Achievements. *Adv. Mater.* **21**, 2257–2268.
- Perrier, A.L., Tabar, V., Barberi, T., Rubio, M.E., Bruses, J., Topf, N., Harrison, N.L., and Studer, L. (2004). Derivation of midbrain dopamine neurons from human embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **101**, 12543–12548.
- Petros, T.J., Tyson, J.A., and Anderson, S.A. (2011). Pluripotent stem cells for the study of CNS development. *Front. Mol. Neurosci.* **4**, 30.
- Przyborski, S.A. (2005). Differentiation of human embryonic stem cells after transplantation in immune-deficient mice. *Stem Cells* **23**, 1242–1250.
- Purcell, O., and Lu, T.K. (2014). Synthetic analog and digital circuits for cellular computation and memory. *Curr. Opin. Biotechnol.* **29**, 146–155.
- Raeber, G.P., Lutolf, M.P., and Hubbell, J.A. (2005). Molecularly engineered PEG hydrogels: a novel model system for proteolytically mediated cell migration. *Biophys. J.* **89**, 1374–1388.
- Ranga, A., Gobaa, S., Okawa, Y., Mosiewicz, K., Negro, A., and Lutolf, M.P. (2014). 3D niche microarrays for systems-level analyses of cell fate. *Nat. Commun.* **5**, 4324.
- Richardson, T.P., Peters, M.C., Ennett, A.B., and Mooney, D.J. (2001). Polymeric system for dual growth factor delivery. *Nat. Biotechnol.* **19**, 1029–1034.
- Rustad, K.C., Wong, V.W., Sorkin, M., Glotzbach, J.P., Major, M.R., Rajadas, J., Longaker, M.T., and Gurtner, G.C. (2012). Enhancement of mesenchymal stem cell angiogenic capacity and stemness by a biomimetic hydrogel scaffold. *Biomaterials* **33**, 80–90.

- Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., and Clevers, H. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* *459*, 262–265.
- Sato, T., Stange, D.E., Ferrante, M., Vries, R.G., Van Es, J.H., Van den Brink, S., Van Houdt, W.J., Pronk, A., Van Gorp, J., Siersema, P.D., and Clevers, H. (2011a). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* *141*, 1762–1772.
- Sato, T., van Es, J.H., Snippert, H.J., Stange, D.E., Vries, R.G., van den Born, M., Barker, N., Shroyer, N.F., van de Wetering, M., and Clevers, H. (2011b). Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* *469*, 415–418.
- Schwank, G., Koo, B.K., Sasselli, V., Dekkers, J.F., Heo, I., Demircan, T., Sasaki, N., Boymans, S., Cuppen, E., van der Ent, C.K., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* *13*, 653–658.
- Shanks, N., Greek, R., and Greek, J. (2009). Are animal models predictive for humans? *Philos. Ethics Humanit. Med.* *4*, 2.
- Shyer, A.E., Tallinen, T., Nerurkar, N.L., Wei, Z., Gil, E.S., Kaplan, D.L., Tabin, C.J., and Mahadevan, L. (2013). Villification: how the gut gets its villi. *Science* *342*, 212–218.
- Siuti, P., Yazbek, J., and Lu, T.K. (2013). Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* *31*, 448–452.
- Soundararajan, P., Miles, G.B., Rubin, L.L., Brownstone, R.M., and Rafuse, V.F. (2006). Motoneurons derived from embryonic stem cells express transcription factors and develop phenotypes characteristic of medial motor column neurons. *J. Neurosci.* *26*, 3256–3268.
- Spence, J.R., Mayhew, C.N., Rankin, S.A., Kuhar, M.F., Vallance, J.E., Tolle, K., Hoskins, E.E., Kalinichenko, V.V., Wells, S.I., Zorn, A.M., et al. (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* *470*, 105–109.
- Spencer, J.A., Ferraro, F., Roussakis, E., Klein, A., Wu, J., Runnels, J.M., Zaher, W., Mortensen, L.J., Alt, C., Turcotte, R., et al. (2014). Direct measurement of local oxygen concentration in the bone marrow of live animals. *Nature* *508*, 269–273.
- Stephan, M.T., and Irvine, D.J. (2011). Enhancing Cell therapies from the Outside In: Cell Surface Engineering Using Synthetic Nanomaterials. *Nano Today* *6*, 309–325.
- Suga, H., Kadoshima, T., Minaguchi, M., Ohgushi, M., Soen, M., Nakano, T., Takata, N., Wataya, T., Muguruma, K., Miyoshi, H., et al. (2011). Self-formation of functional adenohypophysis in three-dimensional culture. *Nature* *480*, 57–62.
- Tabor, J.J., Salis, H.M., Simpson, Z.B., Chevalier, A.A., Levskaya, A., Marcotte, E.M., Voigt, C.A., and Ellington, A.D. (2009). A synthetic genetic edge detection program. *Cell* *137*, 1272–1281.
- Takasato, M., Er, P.X., Becroft, M., Vanslambrouck, J.M., Stanley, E.G., Elfant, A.G., and Little, M.H. (2014). Directing human embryonic stem cell differentiation towards a renal lineage generates a self-organizing kidney. *Nat. Cell Biol.* *16*, 118–126.
- Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.R., Ueno, Y., Zheng, Y.W., Koike, N., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* *499*, 481–484.
- Tan, S., and Barker, N. (2013). Engineering the niche for stem cells. *Growth Factors* *31*, 175–184.
- Ungrin, M.D., Joshi, C., Nica, A., Bauwens, C., and Zandstra, P.W. (2008). Reproducible, ultra high-throughput formation of multicellular organization from single cell suspension-derived human embryonic stem cell aggregates. *PLoS ONE* *3*, e1565.
- Vazin, T., and Schaffer, D.V. (2010). Engineering strategies to emulate the stem cell niche. *Trends Biotechnol.* *28*, 117–124.
- Vincent, L.G., Choi, Y.S., Alonso-Latorre, B., del Álamo, J.C., and Engler, A.J. (2013). Mesenchymal stem cell durotaxis depends on substrate stiffness gradient strength. *Biotechnol. J.* *8*, 472–484.
- Vinci, M., Gowan, S., Boxall, F., Patterson, L., Zimmermann, M., Court, W., Lomas, C., Mendiola, M., Hardisson, D., and Eccles, S.A. (2012). Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. *BMC Biol.* *10*, 29.
- Visconti, R.P., Kasyanov, V., Gentile, C., Zhang, J., Markwald, R.R., and Mironov, V. (2010). Towards organ printing: engineering an intra-organ branched vascular tree. *Expert Opin. Biol. Ther.* *10*, 409–420.
- Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* *110*, 385–397.
- Wilson, S.S., Tocchi, A., Holly, M.K., Parks, W.C., and Smith, J.G. (2015). A small intestinal organoid model of non-invasive enteric pathogen-epithelial cell interactions. *Mucosal Immunol.* *8*, 352–361.
- Wischerhoff, E., Uhlig, K., Lankenau, A., Börner, H.G., Laschewsky, A., Dusch, C., and Lutz, J.-F. (2008). Controlled cell adhesion on PEG-based switchable surfaces. *Angew. Chem. Int. Ed. Engl.* *47*, 5666–5668.
- Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. *Nat. Biotechnol.* *19*, 971–974.
- Yan, Y., Yang, D., Zarnowska, E.D., Du, Z., Werbel, B., Valliere, C., Pearce, R.A., Thomson, J.A., and Zhang, S.C. (2005). Directed differentiation of dopaminergic neuronal subtypes from human embryonic stem cells. *Stem Cells* *23*, 781–790.
- Yin, X., Farin, H.F., van Es, J.H., Clevers, H., Langer, R., and Karp, J.M. (2014). Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny. *Nat. Methods* *11*, 106–112.
- Yui, S., Nakamura, T., Sato, T., Nemoto, Y., Mizutani, T., Zheng, X., Ichinose, S., Nagaiishi, T., Okamoto, R., Tsuchiya, K., et al. (2012). Functional engraftment of colon epithelium expanded in vitro from a single adult Lgr5+ stem cell. *Nat. Med.* *18*, 618–623.
- Zisch, A.H., Lutolf, M.P., and Hubbell, J.A. (2003). Biopolymeric delivery matrices for angiogenic growth factors. *Cardiovasc. Pathol.* *12*, 295–310.

Zika Virus Infects Human Cortical Neural Progenitors and Attenuates Their Growth

Hengli Tang,^{1,11,*} Christy Hammack,^{1,11} Sarah C. Ogden,^{1,11} Zhexing Wen,^{2,3,11} Xuyu Qian,^{2,4,11} Yujing Li,⁹ Bing Yao,⁹ Jaehoon Shin,^{2,5} Feiran Zhang,⁹ Emily M. Lee,¹ Kimberly M. Christian,^{2,3} Ruth A. Didier,¹⁰ Peng Jin,⁹ Hongjun Song,^{2,3,5,6,7,*} and Guo-li Ming^{2,3,5,6,7,8,*}

¹Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

²Institute for Cell Engineering

³Department of Neurology

⁴Biomedical Engineering Graduate Program

⁵Cellular and Molecular Medicine Graduate Program

⁶Solomon Snyder Department of Neuroscience

⁷Kavli Neuroscience Discovery Institute

⁸Department of Psychiatry and Behavioral Sciences

Johns Hopkins University School of Medicine, Baltimore, MD 21025, USA

⁹Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA

¹⁰College of Medicine, Florida State University, Tallahassee, FL 32306, USA

¹¹Co-first author

*Correspondence: tang@bio.fsu.edu (H.T.), shongju1@jhmi.edu (H.S.), gming1@jhmi.edu (G.-l.M.)

<http://dx.doi.org/10.1016/j.stem.2016.02.016>

SUMMARY

The suspected link between infection by Zika virus (ZIKV), a re-emerging flavivirus, and microcephaly is an urgent global health concern. The direct target cells of ZIKV in the developing human fetus are not clear. Here we show that a strain of the ZIKV, MR766, serially passaged in monkey and mosquito cells efficiently infects human neural progenitor cells (hNPCs) derived from induced pluripotent stem cells. Infected hNPCs further release infectious ZIKV particles. Importantly, ZIKV infection increases cell death and dysregulates cell-cycle progression, resulting in attenuated hNPC growth. Global gene expression analysis of infected hNPCs reveals transcriptional dysregulation, notably of cell-cycle-related pathways. Our results identify hNPCs as a direct ZIKV target. In addition, we establish a tractable experimental model system to investigate the impact and mechanism of ZIKV on human brain development and provide a platform to screen therapeutic compounds.

Zika virus (ZIKV), a mosquito-borne flavivirus, is now reported to be circulating in 26 countries and territories in Latin America and the Caribbean (Petersen et al., 2016). While infected individuals can often be asymptomatic or have only mild symptoms, of mounting concern are reports linking ZIKV infection to fetal and newborn microcephaly and serious neurological complications, such as Guillain-Barré syndrome (Petersen et al., 2016). The World Health Organization declared a Public Health Emergency of International Concern on February 1 of 2016 (Heymann et al., 2016). ZIKV infects human skin cells, consistent with its

major transmission route (Hamel et al., 2015). ZIKV was detected in the amniotic fluid of two pregnant women whose fetuses had been diagnosed with microcephaly (Calvet et al., 2016), suggesting that ZIKV can cross the placental barrier. ZIKV was also found in microcephalic fetal brain tissue (Mlakar et al., 2016). Because so little is known about direct cell targets and mechanisms of ZIKV, and because access to fetal human brain tissue is limited, there is an urgent need to develop a new strategy to determine whether there is a causal relationship between ZIKV infection and microcephaly. Here we used human induced pluripotent stem cells (hiPSCs) as an in vitro model to investigate whether ZIKV directly infects human neural cells and the nature of its impact.

We obtained a ZIKV stock from the infected rhesus *Macaca* cell line LLC-MK2. We passaged the virus in the mosquito C6/C36 cell line and titered collected ZIKV on Vero cells, an interferon-deficient monkey cell line commonly used to titer viruses. Sequences of multiple RT-PCR fragments generated from this stock (Figure S1A) matched the sequence of MR766, the original ZIKV strain that likely passed from an infected rhesus monkey to mosquitoes (Dick et al., 1952). We first tested several human cell lines and found varying levels of susceptibility to ZIKV infection (Table S1). Notably, the human embryonic kidney cell line HEK293T showed low permissiveness for ZIKV infection (Figure S1C).

To identify direct target cells of ZIKV in the human neural lineage, we used a highly efficient protocol to differentiate hiPSCs into forebrain-specific human neural progenitor cells (hNPCs), which can be further differentiated into cortical neurons (Wen et al., 2014). The titer of ZIKV in the infected humans is currently unknown. We performed infections at a low multiplicity of infection (MOI < 0.1) and the medium containing virus inoculum was removed after a 2 hr incubation period. Infection rates were then quantified 56 hr later with RT-PCR using MR766-specific primers (Figure S1A) and with immunocytochemistry using an anti-ZIKV envelope antibody (Figures 1A and 1B). The hNPCs

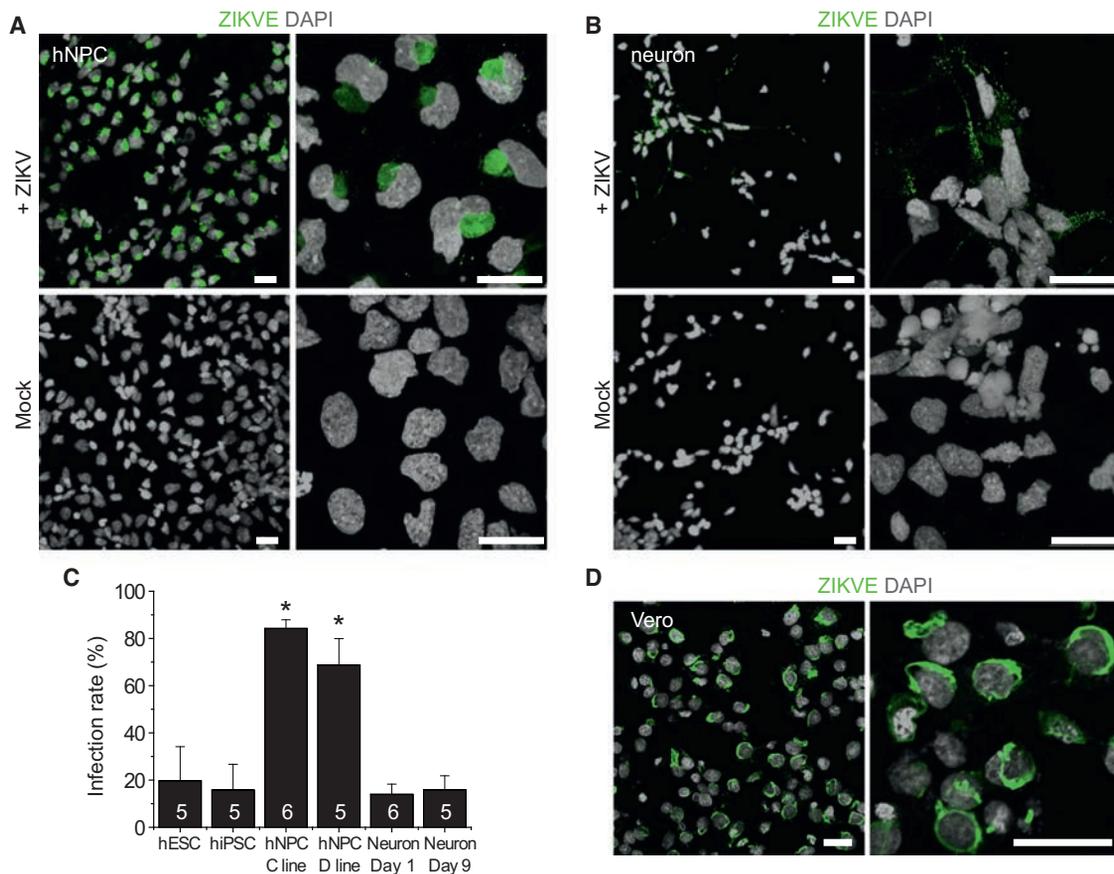


Figure 1. ZIKV Infects hiPSC-Derived Neural Progenitor Cells with High Efficiency

(A and B) Sample confocal images of forebrain-specific hNPCs (A) and immature neurons (B) 56 hr after infection with ZIKV supernatant, immunostained for ZIKV envelop protein (ZIKVE; green) and DAPI (gray). Cells were differentiated from the C1-2 hiPSC line. Scale bars, 20 μ m.

(C) Quantification of infection efficiency for different cell types, including hESCs, hiPSCs, hNPCs derived from two different hiPSCs, and immature neurons 1 or 9 days after differentiation from hNPCs. Both hESCs and hiPSCs were analyzed 72 hr after infection, whereas all other cells were analyzed 56 hr after infection. Numbers associated with bar graphs indicate numbers of independent experiments. Values represent mean \pm SD (* p < 0.01; Student's *t* test).

(D) Production of infectious ZIKV particles by infected hNPCs. Supernatant from hNPC cultures 72 hr after ZIKV infection was collected and added to Vero cells for 2 hr. The Vero cells were further cultured for 48 hr. Shown are sample images of ZIKVE immunostaining (green) and DAPI (gray). Scale bars, 20 μ m. See also Figure S1 and Table S1.

were readily infected by ZIKV *in vitro*, with the infection spreading to 65%–90% of the cells within 3 days of inoculation (Figures 1A and 1C). Quantitative analysis showed similar results for hNPCs derived from hiPSC lines of two different subjects (Figure 1C). As a control, we also exposed human embryonic stem cells (hESCs), hiPSCs, and immature cortical neurons to ZIKV under the same conditions. hESCs and hiPSCs could also be infected by ZIKV, but the infection was limited to a few cells at the colony edge with reduced expression of the pluripotent marker NANOG (Figures 1C and S1D; Table S1). Immature neurons differentiated from hNPCs also exhibited lower levels of infection under our conditions (Figures 1B and 1C). Together, these results establish that hNPCs, a constitutive population of the developing embryonic brain, are a direct cell target of ZIKV.

ZIKV envelope immunostaining exhibited the characteristic intracellular “virus factory” pattern of flaviviruses (Romero-Brey and Bartenschlager, 2014) (Figure 1A). We therefore tested infectivity using supernatant from infected hNPCs and observed robust infection of Vero cells (Figure 1D), indicating that produc-

tive infection of hNPCs leads to efficient secretion of infectious ZIKV particles.

We next determined the potential impact of ZIKV infection on hNPCs. We found a 29.9% \pm 6.6% reduction in the total number of viable cells 66–72 hr after ZIKV infection, as compared to the mock infection (n = 3). Interestingly, ZIKV infection led to significantly higher caspase-3 activation in hNPCs 3 days after infection, as compared to the mock infection, suggesting increased cell death (Figures 2A and 2B). Furthermore, analysis of DNA content by flow cytometry suggested cell-cycle perturbation of infected hNPCs (Figures 2C and S2A). Therefore, ZIKV infection of hNPCs leads to attenuated growth of this cell population that is due, at least partly, to both increased cell death and cell-cycle dysregulation.

To investigate the impact of ZIKV infection on hNPCs at the molecular level, we employed global transcriptome analyses (RNA-seq). Our genome-wide analyses identified a large number of differentially expressed genes upon viral infection (Figure S2B and Table S2). Gene Ontology analyses revealed a particular

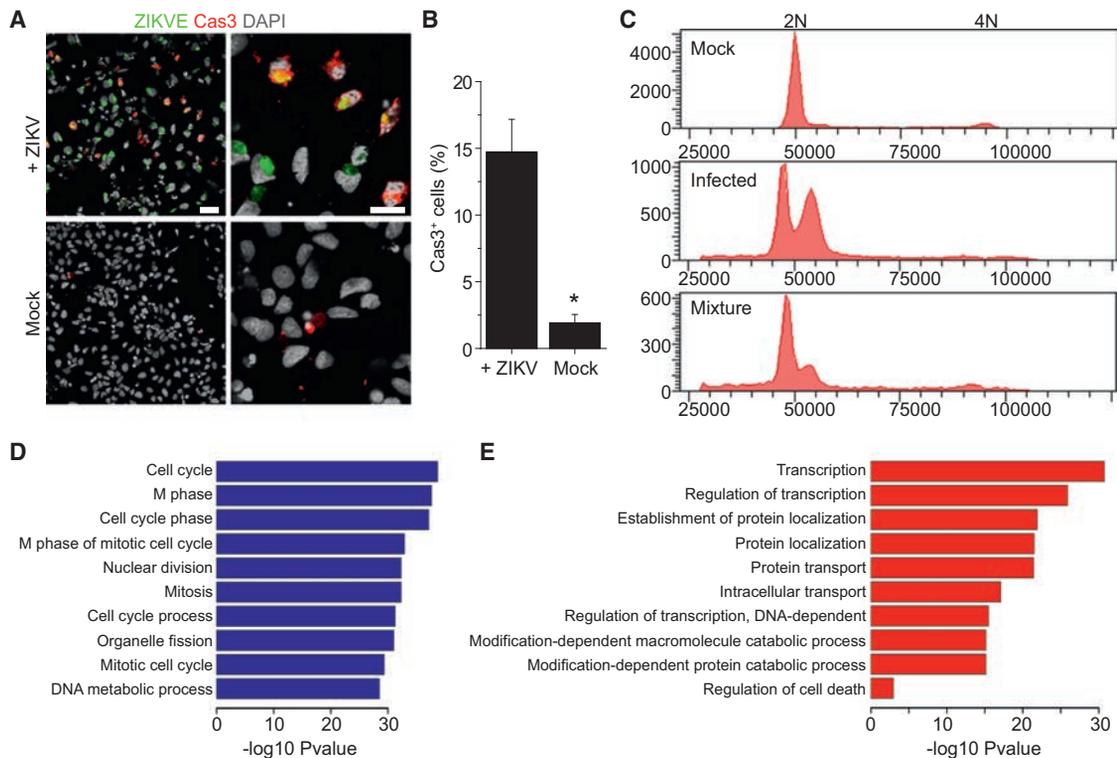


Figure 2. ZIKV-Infected hNPCs Exhibit Increased Cell Death and Dysregulated Cell-Cycle Progression and Gene Expression

(A and B) Increased cell death of ZIKV-infected hNPCs. Shown in (A) are sample images of immunostaining of hNPCs for ZIKVE (green) and cleaved-caspase-3 (Cas3; red) and DAPI (gray) 72 hr after ZIKV infection. Scale bars, 20 μ m. Shown in (B) is the quantification. Values represent mean \pm SEM (n = 6; *p < 0.01; Student's t test).

(C) Cell-cycle perturbation of hNPCs infected by ZIKV. Shown are sample flow cytometry analyses of distributions of hNPCs (from the C1-2 line) at different phases of the cell cycle 72 hr after ZIKV or mock infection. For the mixture sample, mock and infected hNPCs were mixed at a ratio of 1:1 following propidium iodide staining of each sample.

(D and E) RNA-seq analysis of hNPCs (C1-2 line) 56 hr after ZIKV or mock infection. Genes with significant differences in expression between infected and uninfected hNPCs were subjected to GO analyses. The top 10 most significant terms are shown for downregulated (D) and upregulated (E) genes, respectively. The $-\log_{10}$ p values are indicated by bar plots. An additional term of regulation of programmed cell death is also shown for upregulated genes (E). See also Figure S2 and Table S2.

enrichment of downregulated genes in cell-cycle-related pathways (Figure 2D), which is consistent with our flow cytometry findings (Figure 2C). Upregulated genes were primarily enriched in transcription, protein transport, and catabolic processes (Figure 2E). Consistent with increased caspase-3 activation observed by immunocytochemistry (Figures 2A and 2B), RNA-seq analysis revealed upregulation of genes, including caspase-3, involved in the regulation of the apoptotic pathway (Figure 2E). These global transcriptome datasets not only support our cell biology findings but also provide a valuable resource for the field.

It is not known whether specific strains of ZIKV circulating in geographically diverse parts of the world differ in their ability to impact neural development, and the strain we used had been discovered prior to the current reports of a potential epidemiological link between ZIKV and microcephaly. Nevertheless, our results clearly demonstrate that ZIKV can directly infect hNPCs in vitro with high efficiency and that infection of hNPCs leads to attenuated population growth through virally induced caspase-3-mediated apoptosis and cell-cycle dysregulation. Infected hNPCs also release infectious viral particles, which pre-

sents a significant clinical challenge for developing effective therapeutics to arrest or block the impact of infection. Future studies using the hiPSC/hNPC model can determine whether various ZIKV strains impact hNPCs differently and, conversely, whether a single ZIKV strain differentially affects hNPCs from hiPSCs of various human populations.

Flaviviruses tend to have broad cellular tropisms and multiple factors contribute to pathogenic outcomes, including specific cellular response and tissue accessibility. Dengue virus infects cells of several lineages and hematopoietic cells play an essential role in the associated pathogenesis (Pham et al., 2012). West Nile virus infects epithelial cells of multiple tissues and can be neuroinvasive (Suthar et al., 2013). We note that ZIKV also infects other human cell types, including skin cells and fibroblasts (Hamel et al., 2015), and it remains unknown how ZIKV may gain access to the fetal brain (Mlakar et al., 2016). The capacity of ZIKV to infect hNPCs and attenuate their growth underscores the urgent need for more research into the role of these cells in putative ZIKV-related neuropathology. The finding that ZIKV also infects immature neurons raises critical questions about pathological effects on neurons and other neural cell types in

the brain, as well as potential long-term consequences. Intriguingly, an early animal study showed ZIKV infection of neurons and astrocytes in mice and observed enlarged astrocytes (Bell et al., 1971). Our study also raises the question of whether ZIKV infects neural stem cells in adult humans (Bond et al., 2015).

In summary, our results fill a major gap in our knowledge about ZIKV biology and serve as an entry point to establish a mechanistic link between ZIKV and microcephaly. Our study also provides a tractable experimental system for modeling the impact of ZIKV on neural development and for investigating underlying cellular and molecular mechanisms. Of equal importance, our hNPC model and robust cellular phenotype comprise a readily scalable platform for high-throughput screens to prevent ZIKV infection of hNPCs and to ameliorate its pathological effects during neural development.

ACCESSION NUMBERS

The accession number for RNA-seq data reported in this paper is GEO: GSE78711.

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes two figures, two tables, and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.02.016>.

AUTHOR CONTRIBUTIONS

H.T., H.S., and G.-I.M. conceived of the research, designed the study, and wrote the manuscript. C.H., S.C.O., Z.W., and X.Q. performed experiments, analyzed data, and contributed equally to this study. Y.L., B.Y., J.S., F.Z., and P.J. performed RNA-seq analysis, and E.M.L., K.M.C., and R.A.D. contributed to additional data collection. All authors commented on the manuscript.

ACKNOWLEDGMENTS

We thank Yichen Cheng, Taylor Lee, and Jianshe Lang of the Tang laboratory, Lihong Liu and Yuan Cai of the Ming and Song laboratories, Luoxiu Huang of the Jin laboratory for technical assistance, Zhiheng Xu and additional laboratory members for suggestions, and Timothy Megraw for assistance with confocal imaging. H.T. thanks the College of Arts and Sciences and the Department of Biological Science at Florida State University for seed funding. This work was partially supported by NIH (AI119530/AI111250 to H.T., NS047344 to H.S., NS048271/NS095348 to G.-I.M., and NS051630/NS079625/MH102690 to P.J.), MSCRF (to H.S.), and start-up funding (to H.S.).

Received: February 24, 2016

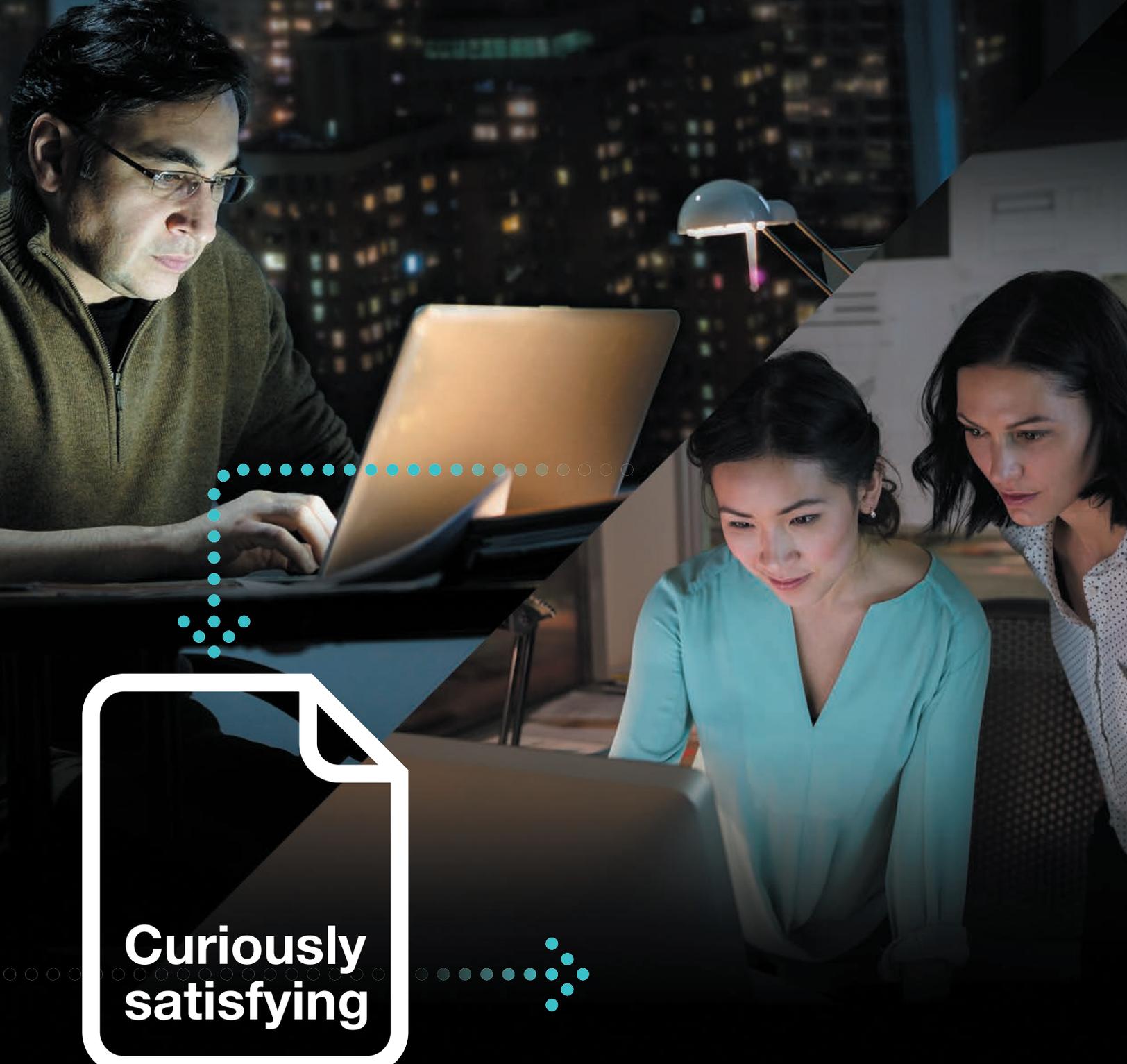
Revised: February 28, 2016

Accepted: February 29, 2016

Published: March 4, 2016

REFERENCES

- Bell, T.M., Field, E.J., and Narang, H.K. (1971). Zika virus infection of the central nervous system of mice. *Arch. Gesamte Virusforsch.* **35**, 183–193.
- Bond, A.M., Ming, G.L., and Song, H. (2015). Adult mammalian neural stem cells and neurogenesis: Five decades later. *Cell Stem Cell* **17**, 385–395.
- Calvet, G., Aguiar, R.S., Melo, A.S., Sampaio, S.A., de Filippis, I., Fabri, A., Araujo, E.S., de Sequeira, P.C., de Mendonça, M.C., de Oliveira, L., et al. (2016). Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. *Lancet Infect. Dis.* S1473-3099(16)00095-5. [http://dx.doi.org/10.1016/S1473-3099\(16\)00095-5](http://dx.doi.org/10.1016/S1473-3099(16)00095-5).
- Dick, G.W., Kitchen, S.F., and Haddock, A.J. (1952). Zika virus. I. Isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* **46**, 509–520.
- Hamel, R., Dejarnac, O., Wichit, S., Ekcharyawat, P., Neyret, A., Luplertlop, N., Perera-Lecoin, M., Surasombattana, P., Talignani, L., Thomas, F., et al. (2015). Biology of Zika virus infection in human skin cells. *J. Virol.* **89**, 8880–8896.
- Heymann, D.L., Hodgson, A., Sall, A.A., Freedman, D.O., Staples, J.E., Althabe, F., Baruah, K., Mahmud, G., Kandun, N., Vasconcelos, P.F., et al. (2016). Zika virus and microcephaly: why is this situation a PHEIC? *Lancet* **387**, 719–721.
- Mrakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., Kolenc, M., Resman Rus, K., Vesnaver Vipotnik, T., Fabjan Vodusek, V., et al. (2016). Zika virus associated with microcephaly. *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1600651>.
- Petersen, E., Wilson, M.E., Touch, S., McCloskey, B., Mwaba, P., Bates, M., Dar, O., Mattes, F., Kidd, M., Ippolito, G., et al. (2016). Rapid spread of Zika virus in the Americas - Implications for public health preparedness for mass gatherings at the 2016 Brazil Olympic Games. *Int. J. Infect. Dis.* **44**, 11–15.
- Pham, A.M., Langlois, R.A., and TenOever, B.R. (2012). Replication in cells of hematopoietic origin is necessary for Dengue virus dissemination. *PLoS Pathog.* **8**, e1002465.
- Romero-Brey, I., and Bartenschlager, R. (2014). Membranous replication factories induced by plus-strand RNA viruses. *Viruses* **6**, 2826–2857.
- Suthar, M.S., Diamond, M.S., and Gale, M., Jr. (2013). West Nile virus infection and immunity. *Nat. Rev. Microbiol.* **11**, 115–128.
- Wen, Z., Nguyen, H.N., Guo, Z., Lalli, M.A., Wang, X., Su, Y., Kim, N.S., Yoon, K.J., Shin, J., Zhang, C., et al. (2014). Synaptic dysregulation in a human iPS cell model of mental disorders. *Nature* **515**, 414–418.



**Curiously
satisfying**

**Introducing Sneak Peek, a preview
of Cell Press papers under review**

Exciting science needs to be shared, and fast. That's why we're introducing Cell Press Sneak Peek, a preview of the papers under review at our primary research journals. Sneak Peek makes papers discoverable earlier in the publication process—so authors can surface their research quickly and readers can build on their work.

**Go on, satisfy your curiosity!
Visit www.cell.com/sneakpeek**



Sneak Peek
A PREVIEW OF PAPERS UNDER REVIEW

Complete Meiosis from Embryonic Stem Cell-Derived Germ Cells In Vitro

Quan Zhou,^{1,2,8} Mei Wang,^{2,3,8} Yan Yuan,^{1,2,8} Xuepeng Wang,^{2,4} Rui Fu,² Haifeng Wan,² Mingming Xie,^{2,5} Mingxi Liu,¹ Xuejiang Guo,¹ Ying Zheng,⁶ Guihai Feng,² Qinghua Shi,⁷ Xiao-Yang Zhao,^{2,9,*} Jiahao Sha,^{1,*} and Qi Zhou^{2,*}

¹State Key Laboratory of Reproductive Medicine, Department of Histology and Embryology, Nanjing Medical University, Nanjing 210029, China

²State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

³College of the Life Sciences, Hunan Normal University, Changsha, Hunan 410081, China

⁴University of the Chinese Academy of Sciences, Beijing 100049, China

⁵College of Life Science, Anhui University of China, Hefei 230601, China

⁶Department of Histology and Embryology, Medical College of Yangzhou University, Yangzhou 225001, Jiangsu, China

⁷Molecular and Cell Genetics Laboratory, Chinese Academy of Sciences Key Laboratory of Innate Immunity and Chronic Disease, Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei 230027, China

⁸Co-first author

⁹Present address: Department of Developmental Biology, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, China

*Correspondence: qzhou@ioz.ac.cn (Q.Z.), shajh@njmu.edu.cn (J.S.), zhaoxiaoyang@smu.edu.cn (X.-Y.Z.)

<http://dx.doi.org/10.1016/j.stem.2016.01.017>

SUMMARY

In vitro generation of functional gametes is a promising approach for treating infertility, although faithful replication of meiosis has proven to be a substantial obstacle to deriving haploid gamete cells in culture. Here we report complete in vitro meiosis from embryonic stem cell (ESC)-derived primordial germ cells (PGCLCs). Co-culture of PGCLCs with neonatal testicular somatic cells and sequential exposure to morphogens and sex hormones reproduced key hallmarks of meiosis, including erasure of genetic imprinting, chromosomal synapsis and recombination, and correct nuclear DNA and chromosomal content in the resulting haploid cells. Intracytoplasmic injection of the resulting spermatid-like cells into oocytes produced viable and fertile offspring, showing that this robust stepwise approach can functionally recapitulate male gametogenesis in vitro. These findings provide a platform for investigating meiotic mechanisms and the potential generation of human haploid spermatids in vitro.

INTRODUCTION

In sexually reproducing organisms, transmission of genetic and epigenetic information between generations relies on gametes, reproductive cells of the germline that unite at fertilization to form a new organism. In mammals, germline specification occurs early during embryogenesis, when primordial germ cells (PGCs) first appear in the extra-embryonic compartment. These germ cell progenitors undergo a complex developmental program involving migration into the developing embryo, coloniza-

tion of the gonadal ridges, proliferation, and eventual progression through meiosis to form haploid sex-specific germ cells. Up to 15% of couples are infertile, and many of them have gametogenesis failure. Therefore, reproducing germ cell development in vitro has remained a central goal in reproductive biology and medicine (Sun et al., 2014). The effective production of functional gametes in culture would not only provide a system to investigate the genetic, epigenetic, and environmental factors that shape germ cell development but may also lead to clinical approaches addressing infertility resulting from defects in gametogenesis.

The recapitulation of meiosis, a process unique to germ cells, has remained a major obstacle toward the production of functional gametes in vitro. To avoid misconceptions, a consensus panel of reproductive biologists has therefore formulated a panel of “gold standard” criteria for in-vitro-derived gametes that are based on features that reflect key events of meiosis (Handel et al., 2014). To prove that meiosis occurred in vitro, all of the following must be shown: correct nuclear DNA content at specific meiotic stages (for male cells, pre-meiotic, meiotic S phase, first reductional, and second meiotic division stages), normal chromosome number and organization, appropriate nuclear and chromosomal localization of proteins involved in homologous synapsis and recombination, and capacity of the in vitro-produced germ cells to produce viable euploid offspring.

The production of mature male germ cells is the result of a complex developmental program that begins with the specification of PGCs during prenatal stages and relies on spermatogenesis, a complex multi-step expansion and maturation process that is initiated in the postnatal testis during puberty and encompasses mitotic proliferation of spermatogonia, meiotic division into haploid germ cells, and spermiogenic differentiation. In response to cytokines resembling those released by early extraembryonic tissues, mouse and human embryonic stem cells (ESCs) or induced pluripotent stem cells

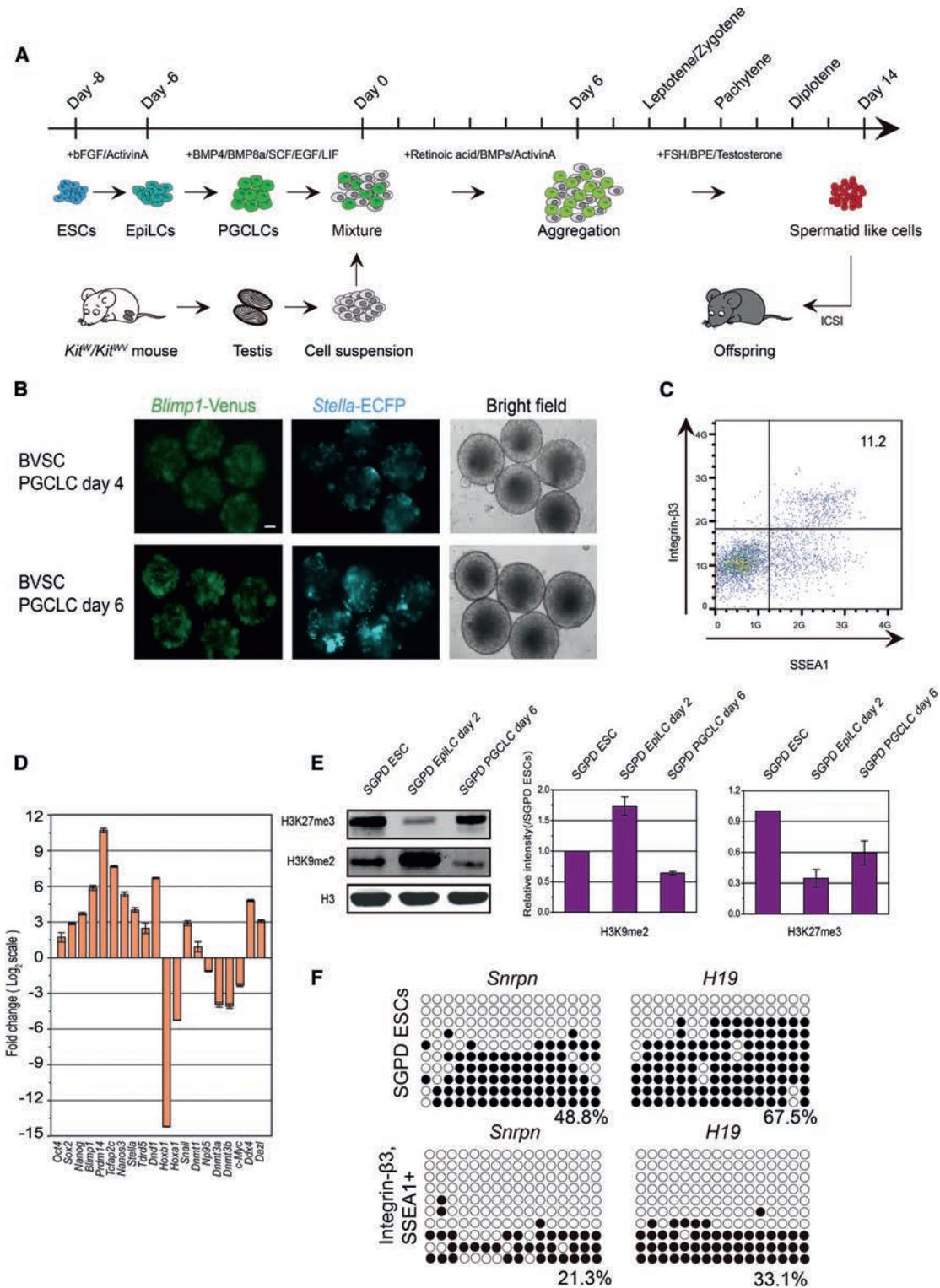


Figure 1. SGPD and BVSC ESCs Differentiated into PGCLCs

(A) Time line and culture conditions of in vitro spermatogenesis.

(B) Expression of *Blimp1*-mVenus and *Stella*-ECFP identifies BVSC ESC-derived PGCLCs on days 4 and 6 of PGCLC culture (days –2 and 0 of overall culture). Scale bar, 50 μ m.

(C) FACS of SSEA1 and integrin β 3 double-positive cells in day 6 aggregates (day –2).

(legend continued on next page)

(iPSCs) can adopt an epiblast-like developmental program in vitro and transiently form epiblast-like cells (EpiLCs) that are competent to undergo specification into PGC-like cells (PGCLCs) (Hayashi et al., 2011; Irie et al., 2015). Upon transplantation of murine PGCLCs into the testis environment, these further developed into haploid spermatozoa (Hayashi et al., 2011). In humans, the risk for tumorigenesis prohibits in vivo transplantation approaches so that the capacity of human PGCLCs to form spermatozoa remains unexplored. Several studies have reported the generation of haploid germ cells in vitro from pluripotent stem cells (Eguizabal et al., 2011; Geijsen et al., 2004), but none of these studies provided proof for all key hallmarks of meiosis, and the functionality of the haploid germ cells were not fully evaluated. A 3D culture system aimed at reconstructing mouse seminiferous tubules in vitro supported sperm formation from neonatal testicular cells. However, the function of the spermatids obtained with this system has not been evaluated (Yokonishi et al., 2013). Therefore, the full recapitulation of spermatogenesis in vitro to produce functional haploid male gametes has not yet been achieved but is highly anticipated.

Here we report complete in vitro meiosis from murine ESC-derived PGCLCs, resulting in the formation of male spermatid-like cells (SLCs) capable of producing viable fertile offspring via intracytoplasmic sperm injection (ICSI). We demonstrate that, upon co-culture with neonatal testicular somatic cells and sequential exposure to morphogens and sex hormones, ESC-derived PGCLCs recapitulate male gametogenesis in vitro, reproducing hallmarks of erasure of imprints, synapsis, and recombination.

RESULTS

Specification of PGCLCs from ESCs In Vitro

To visualize key stages during germ cell development from ESCs in vitro, we derived and used mouse ESC lines transgenic for fluorescent reporter proteins under the control of regulatory elements of germ cell markers. For identification of cells resembling PGCs, we used an ESC line expressing membrane-targeted Venus (mVenus) under the control of *Prdm1* (*Blimp1*) regulatory elements and enhanced CFP (ECFP) under the control of *Dppa3* (*Stella/Pgc7*), marking lineage-restricted PGC precursors of the proximal epiblast and *Prdm1*- and *Dppa3*-positive migrating PGCs, respectively (*Blimp1*-mVenus and *Stella*-ECFP [BVSC] double-transgenic ESC line). We also used an ESC line expressing EGFP controlled by the cell-specific stimulated by retinoic acid gene 8 (*Stra8*) promoter, reflecting early-stage spermatogonia through preleptotene-stage spermatocytes (Li et al., 2007). This line also expressed DsRed under the control of the protamine 1 (*Prrm1*) promoter, identifying post-meiotic spermatids (*Stra8*-EGFP and *Prrm1*-DsRed [SGPD] dou-

ble-transgenic ESC line; Figure S1A). Both ESC lines had a normal karyotype and produced completely ESC-derived live offspring by tetraploid complementation (Figures S1B and S1C), confirming pluripotency.

From ground-state SGPD and BVSC ESCs that were maintained under serum- and feeder free conditions in the presence of GSK3 inhibitor and MEK inhibitor (2i) and leukemia inhibitory factor (LIF), we first induced differentiation into EpiLCs using culture conditions adapted from a protocol published previously (Hayashi et al., 2011). In adherent culture, SGPD and BVSC ESCs underwent EpiLC differentiation when exposed to activin A and basic fibroblast growth factor (bFGF) (Figure 1A). This fate change was associated with a decrease in the expression of the pluripotency markers NANOG and SOX2, whereas levels of the pluripotency and germ cell marker OCT4 remained high (Figure S2A). Expression analysis by real-time RT-PCR confirmed downregulation of transcripts of multiple pluripotency and inner cell mass (ICM) marker genes, including *Prdm14*, *Zfp42*, *Tbx3*, *Tcl1*, *Esrrb*, *Klf2*, *Klf4*, and *Klf5*, whereas transcripts of the epiblast marker genes *Fgf5* and *Wnt3* and of *Dnmt3b* became upregulated (Figure S2B). Transcripts of endoderm marker genes (*Gata4*, *Gata6*, *Sox17*, and *Blimp1*) were expressed at low levels during differentiation.

Subsequent floating culture of EpiLC in an N2B27-based differentiation medium containing a cytokine mix of BMP-4, BMP-8a, epidermal growth factor (EGF), SCF, and LIF (Hayashi et al., 2011) resulted in robust activation of *Blimp1*-mVenus and *Stella*-ECFP after 4–6 days, suggesting induction of PGCLCs (Figure 1B). Fluorescence-activated cell sorting (FACS) for the PGC markers integrin $\beta 3$ and SSEA1 yielded 11.2% double-positive cells (Figure 1C). We next evaluated whether changes in gene expression, histone modification, and allele-specific methylation patterns associated with PGC commitment were recapitulated in our culture system. Comparison of expression profiles of SGPD ESC-derived EpiLCs and day 6 PGCLC aggregates revealed upregulation of transcripts of pluripotency genes, including *Oct4*, *Sox2*, and *Nanog*, in PGCLCs (Figure 1D). Similarly, transcript levels of PGC-related genes, including *Blimp1*, *Prdm14*, *Tcfap2c*, *Nanos3*, *Stella*, *Tdrd5*, *Dnd1*, *Dnmt1*, *Ddx4*, and *Dazl*, increased, whereas those of somatic cell-related genes such *hoxa1* and *hoxb1* and other genes, including *Dnmt3a/3b*, *Np95*, and *c-Myc*, became downregulated. This gene expression profile resembles that of in vivo PGCs (Kurimoto et al., 2008; Saitou et al., 2002), confirming in vitro PGCLC specification (Hayashi et al., 2011). Analysis of epigenetic profiles by western blot demonstrated similar dynamics of histone modification as observed during PGC formation in vivo (Seki et al., 2005) and PGCLC induction in vitro, with a transient increase and reduction of histone H3 lysine 9 dimethylation (H3K9me2) and histone H3 lysine 27 trimethylation (H3K27me3), respectively, during formation of EpiLCs, followed

(D) Gene expression changes in SGPD PGCLCs between day 0 (day –2 of overall culture) and day 6 of PGCLC culture (day 0). Average values \pm SD were plotted on the log₂ scale.

(E) Western blot analysis (left) of H3K9me2 and H3K27me3 in SGPD ESCs, SGPD day 2 EpiLCs (day –6 overall), and SGPD day 6 (day 0 overall) PGCLCs (SSEA1 and integrin $\beta 3$ double-positive cells). Quantification of H3K9me2 (center) and H3K27me3 normalized to H3 levels (right); mean \pm SD.

(F) Bisulfite sequencing of DMRs that regulate mono-allelic expression of the imprinted genes *Snrpn* and *H19* from the paternal and maternal allele, respectively. White and black circles indicate unmethylated and methylated CpGs, respectively.

See also Figures S1 and S2.

by the inverse pattern in PGCLCs (Figure 1E; Hayashi et al., 2011; Seki et al., 2005). Bisulfite sequencing of differentially methylated regions (DMRs) that regulate mono-allelic expression of the imprinted genes *Snrpn* and *H19* from the paternal and maternal allele, respectively, revealed a reduction of methylation that may indicate that erasure of imprinting was initiated in SGPD PGCLCs (Figure 1F). Collectively, both SGPD and BVSC ESCs were induced effectively to form PGCLCs.

Initiation of Meiosis in PGCLCs In Vitro in Co-culture

Completion of spermatogenesis from in vitro-produced PGCLCs has only been achieved after transplantation into mouse testes using the germ cell-deficient *KIT^W/KIT^{W-V}* mouse model to ensure the donor origin of haploid cells (Hayashi et al., 2011). To reconstitute an in vitro environment compatible with meiotic progression, we pursued co-culture of PGCLCs with early postnatal testicular cells from *KIT^W/KIT^{W-V}* mice. In male mice, fetal-stage gonadal somatic cells express high levels of CYP26B1, which metabolizes endogenous retinoic acid (RA), thereby prohibiting entry of male germ cells into meiosis (Bowles et al., 2006; MacLean et al., 2007). Initiation of meiosis remains suppressed in the male gonad until after birth (Zhou et al., 2008). Consistent with this, we found that early postnatal somatic testicular cells express low levels of CYP26B1, comparable with those of the female fetal gonad at the stage of meiotic induction of PGCs at embryonic day (E) 13.5 (Anderson et al., 2008; McLaren, 2003; Figure S3A). Therefore, early postnatal testicular cells of *KIT^W/KIT^{W-V}* mice may represent a permissive environment for the initiation of meiosis (Zhou et al., 2008) despite developmental asynchrony. We used a mixed cell culture system of SGPD PGCLCs and *KIT^W/KIT^{W-V}* testicular cells at a ratio of 1:1.

An RA signal is required during meiotic induction (Bowles et al., 2006; Menke et al., 2003; Zhou et al., 2008), and pathways responsive to activin A and BMP-2/4/7 (BMPs) regulate postnatal germ cell development, including self-renewal (Hu et al., 2004; Puglisi et al., 2004). To assess the effect of these morphogens on the initiation of meiosis in vitro, we exposed co-cultures of SSEA1 and integrin β 3 double-positive PGCLCs and *KIT^W/KIT^{W-V}* testicular cells to different combinations of activin A, BMPs, and RA (Figure 1A). Only in cultures supplemented with all three morphogens (activin A, BMPs, and RA) did *Stra8*-EGFP-expressing cells become apparent within 3 days of exposure, suggesting the initiation of meiosis in SGPD PGCLCs in vitro (Figure 2A). Gene expression analysis revealed that increased transcript levels of the germ cell markers *Nanos3* and *Ddx4* were detectable in all cultures exposed to activin A and BMPs. However, only co-cultures additionally supplemented with RA exhibited a marked upregulation of the meiosis markers *Stra8* and *Dmc1* after day 6 (Figure 2B). The expression level of *Rec8*, another target of RA but independent of *Stra8* (Koubova et al., 2014), was higher when treated with RA than that without RA treatment after day 6 (Figure 2B). Cells not exposed to activin A or BMPs exhibited poor proliferation capacity (Figure 2C). In co-cultures exposed to all three morphogens, we observed that testis somatic cells migrated actively toward PGCLCs, forming aggregation colonies with *Stra8*-EGFP-expressing cells within 6 days (Movie

S1). *Prrm1*-DsRed expression was not detected at this time. 336 of the 400 (84%) colonies showed *Stra8*-EGFP-positive signals. The *Stra8*-EGFP-positive colonies were positive for markers of germ cells (DDX4), testis somatic cells (GATA4), and Sertoli cells (SOX9), whereas the PGC markers SSEA1 and OCT4 were undetectable (Figure 2D). SGPD PGCLCs had therefore differentiated from a PGC/spermatogonial stem cell (SSC) state and integrated into colonies comprised of multiple testicular cell types, including *KIT^W/KIT^{W-V}* testicular somatic cells. In BVSC double-positive PGCLCs, downregulation of the *Blimp1*-mVenus and *Stella*-ECFP signal occurred within 3 days of co-culture in the presence of all three morphogens, supporting this hypothesis (Figure S3B). This synchronous process resembles the behavior of PGCs in the E13.5 female genital ridge, which simultaneously enter meiosis in the following few days. Our co-culture conditions therefore reconstitute a microenvironment that triggers the initiation of meiosis in germ cells.

Hormonal Stimulation Induces the Formation of Haploid SLCs In Vitro

Sex hormones, including follicle-stimulating hormone (FSH) and testosterone (T) regulate the progression of meiosis in mice (O'Shaughnessy, 2014). Starting on day 7 of co-culture, we therefore withdrew morphogens (Activin A, BMPs, and RA [ABR]) and exposed SGPD-derived cultures to combinations of FSH, bovine pituitary extract (BPE), and T. In the presence of FSH/BPE/T, postmeiotic *Prrm1*-DsRed-expressing cells became first detectable on day 10 (Figure 3A), and strong reporter expression on day 14 correlated with upregulation of haploid spermatid markers such as *Tp1*, *Prrm1*, *acrosin*, and *haprin*. In contrast, cultures supplemented with fewer hormone factors did not contain *Prrm1*-DsRed-expressing cells. A slight upregulation of *Sycp3* was detectable in cultures exposed to BPE and T (Figure 3B). In the presence of FSH/BPE/T, 14% of cells were identified to have 1C DNA content on day 14 of culture, indicating the formation of haploid male SLCs. Less than 3% of SLCs were detectable in cultures not exposed to FSH, and SLCs were absent from cultures without BPE or T (Figure 3C), indicating that all three hormones were required for progression of meiosis. No 1C cells could be induced from the somatic cells on day 14, even in the presence of FSH/BPE/T (Figure 3C). In addition to FSH, BPE contains a variety of growth factors and hormones. Therefore, it is possible that factors other than FSH contribute to the observed completion of meiosis in vitro with FSH, T, and BPE stimulation.

In FSH/BPE/T-supplemented cultures, haploid cells were first detectable around day 10 and increased over the next 4 days to 14%–20%. Approximately half of the haploid cells expressed *Prrm1*-DsRed on day 14 (Figure 4A). Cells had a normal karyotype during metaphase of meiosis I (Figure 4B).

Chromosome Synapsis and Recombination in Meiosis In Vitro

To observe the progression of meiosis during culture, we assessed chromosomal synapsis and recombination/crossover. These processes require the initiation and resolution of DNA double-strand breaks (DSBs). Nuclear spreads of differentiating cells on day 8 contained multiple SPO11 and RAD51

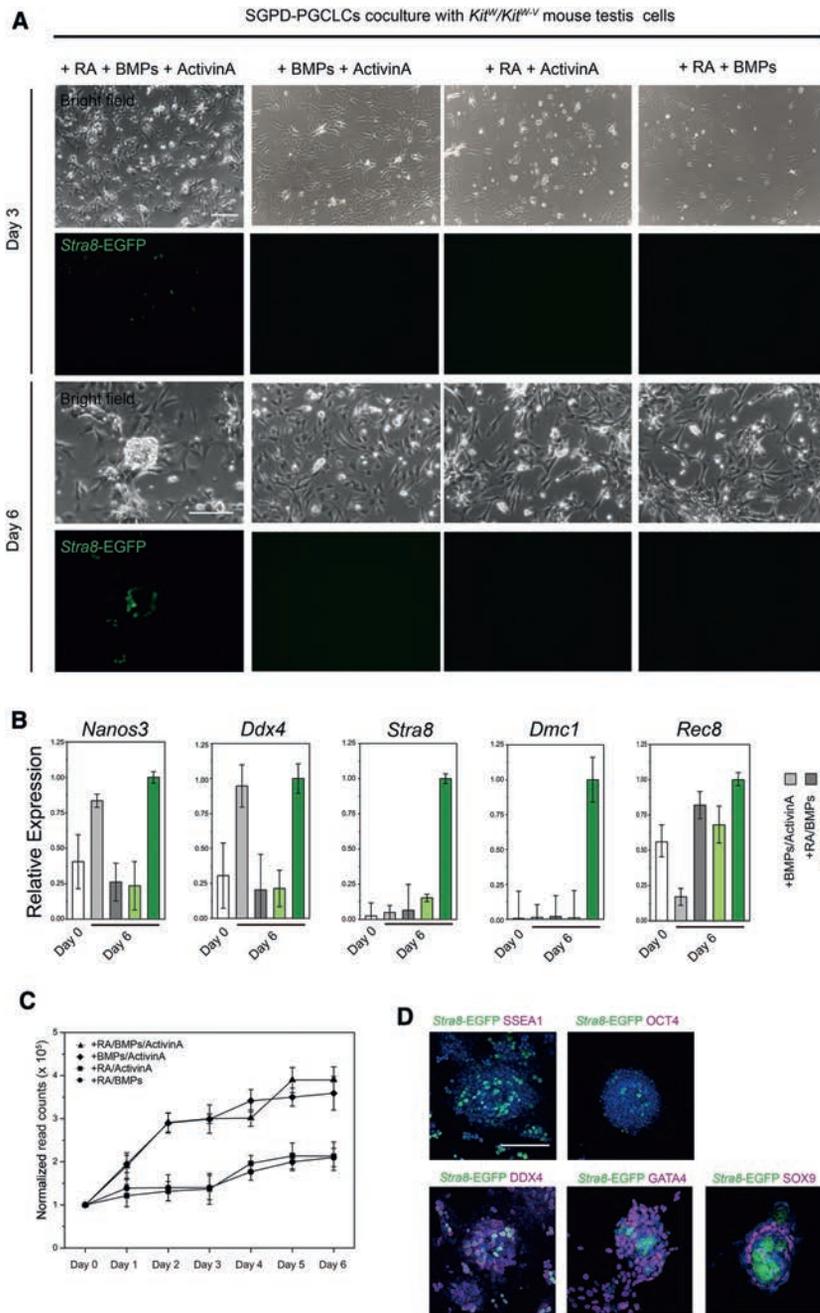


Figure 2. Induction of Meiosis in PGCLCs In Vitro

(A) SGPD-derived PGCLCs initiate meiosis in co-culture with *KIT^{fl}/KIT^{W-V}* testicular cells when exposed to activin A, BMPs, and RA. *Stra8*-EGFP-positive cells and colonies were detected after 3 and 6 days, respectively. Scale bar, 100 μ m. (B) qRT-PCR of co-cultures exposed to different morphogen combinations. Relative gene expression levels were normalized to *Rps2* and reflect mean \pm SD from three independent biological replicates. (C) Mean (\pm SD) cell counts of cultures. 10⁵ cells/well were plated on day 0. (D) Immunostaining of day 6 colonies with antibodies directed against germ cell and somatic cell markers combined with Hoechst staining (blue) and *Stra8* promoter-driven EGFP expression. Scale bar, 50 μ m. See also Figure S3 and Movie S1.

verse elements of the synaptonemal complex (Lammers et al., 1994; Meuwissen et al., 1992), and also contained lateral (SYCP3) elements (Eijpe et al., 2000), suggesting progression of synapsis. Quantitative analysis revealed that, after 8 days of differentiation in vitro, more than 90% of the primary spermatocytes were at the leptotene or zygotene stage of meiosis. On day 10, 64% were in the pachytene stage, indicating the successful repair of DNA DSBs and completion of synapsis, and more than 50% of the spermatocytes had entered the diplotene stage by day 12 (Figure 4D). These results demonstrated that meiosis in vitro encompassed synapsis and recombination and was synchronized in the majority of germ cells. Consistent with this, we found that upregulation of transcripts of meiotic factors occurred in a programmed manner, similar to one cycle of meiotic division in vivo (Figure 4E). Transcripts of the meiotic markers *Dmc1* and *Stra8* were first detectable on day 4, increased to the highest levels between days 7 and 10, and decreased by day 14. *Sycp3* transcript levels increased gradually from day 0 to day 10 and then dropped sharply on day 14, consistent with progression to the diplotene stage, and upregulation of haploid spermatid marker transcripts such as *Prm1*, *haprin*, and *acrosin* was most prominent on day 14. FACS of *Prm1*-DsRed-expressing SLCs on day 14 yielded approximately 2×10^4 SLCs/culture well. Because meiotic progression of a single PGCLC would result in four SLCs, the estimated conversion rate from PGCLCs (5×10^4 PGCLCs plated per well) to SLCs was about 10%. Dividing *Prm1*-DsRed-positive cells were present in co-cultures on day 12, indicating the formation of haploid cells (Movie S2).

foci, reflecting the generation of DSBs (Aravind et al., 1998; Keeney et al., 1997) and their resolution by homologous recombination repair (Neale and Keeney, 2006), respectively (Figure 4C; Figures S3C–S3F). Furthermore, we found that the distribution of phosphorylated H2A histone family member X (γ H2AX) recapitulated that of meiotic progression in vivo. Broad distribution throughout the nucleus in day 8 cells (Figure 4C; Figure S3C) reflected an association with DSBs, and disappearance from the autosome region on day 10 and accumulation on the unsynapsed sex chromosomes (Mahadevaiah et al., 2001) suggested completion of synapsis similar to pachytene stage spermatocytes in vivo (Figure 4C). The nuclei of in vitro spermatocytes were positive for SYCP1, a component of trans-

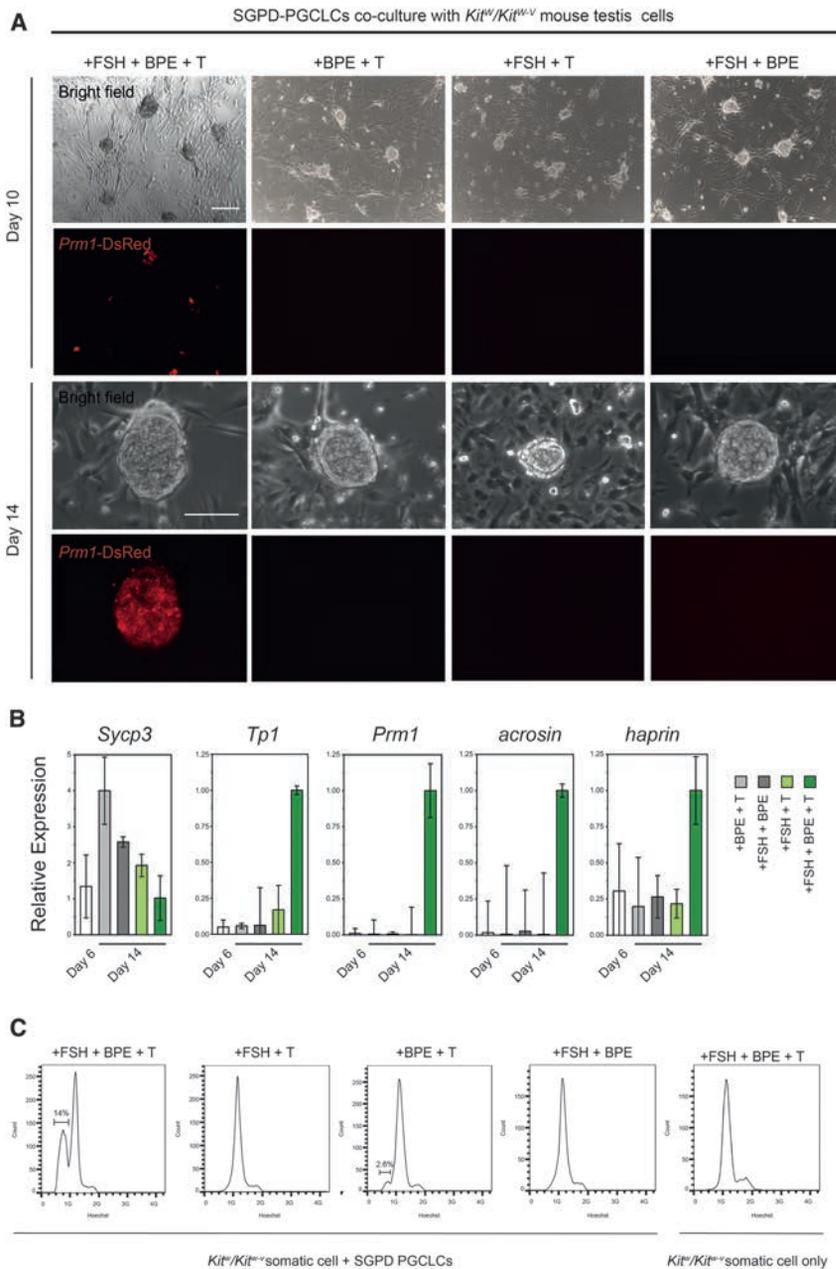


Figure 3. Hormonal Stimulation Induces the Formation of Haploid SLCs in Culture

(A) SGPD-derived co-cultures were supplemented with hormones from days 7–14. *Prm1*-DsRed-expressing cells were detected on day 10 in the presence of FSH/BPE/T. Scale bars, 100 μ m.

(B) qRT-PCR analysis of cultures supplemented with hormones as indicated. Relative gene expression levels were normalized to *Rps2* and reflect mean \pm SD from three independent biological replicates.

(C) FACS analysis of DNA content revealing the presence of 14% and 2.6% haploid SLCs only in cultures exposed to FSH/BPE/T and BPE/T (days 7–14), respectively.

See also Movie S2.

Global transcription profile clustering analyses revealed the similarity of in vitro SLCs to in vivo spermatids (Figure 5D). These analyses also confirmed that global transcription profiles of SSEA1/integrin β 3 double-positive PGCLCs on day 6 clustered closely with those of SSEA1/integrin β 3 double-positive PGCs from E12.5 male fetuses but differed from ESCs and differentiated germ cells.

We next performed ICSI with sorted in vitro-derived SLCs. Of 63 and 125 oocytes injected with SGPD-derived SLCs in two replicates, 51 and 107 developed to the two-cell stage after activation, respectively. Of 191 oocytes injected with BVSC-derived SLCs, 159 developed to the two-cell stage after activation (Table 1). From embryo transfers of SGPD-derived two-cell stage embryos, we obtained six full-term pups that were transgenic for *Prm1*-DsRed (Figures 5E and 5F). Analysis of one pup revealed a normal karyotype (Figure 5G). Bisulfite sequencing indicated a pattern associated with maternal and paternal genetic contributions (Figure 5H). The mouse developed normally to adulthood. Embryos resulting from ICSI with BVSC-derived SLCs (Figure S5A) expressed *Stella*-ECFP at the four-cell stage (Figure S5B). Embryo transfer resulted in three live pups that were positive for the *Blimp1*-mVenus and *Stella*-ECFP transgenes, had a normal karyotype, and exhibited normal weight gain after birth (Figures S5C–S5F). The birth rate following ICSI with round spermatids isolated from normal testis was 9.5% (Table 1). Genome-wide reduced representation bisulfite sequencing (RRBS) of tail tip fibroblast of a male and a female offspring derived from BVSC ESCs was performed to analyze the whole-genome methylation status of the offspring with normal male and female mice as controls. The genome-wide methylation data of sperm and oocyte by Shen et al. (2014) were included in the analysis. The proportions of high methylation sites (>80%)

Healthy Fertile Offspring Produced by ICSI with In Vitro-Derived SLCs

Sorted SLCs contained a cap-shaped acrosome (Figure 5A). To validate genome integrity, we performed 0.1 \times whole genome sequencing of single sorted SLCs. Of eight small and round cells selected for sequencing (Figure 5B), six were haploid cells with normal genome structure, one cell was haploid with chromosomal deletions, and one was diploid (Figure S4). The presence of diploid cells was not unexpected because the *Prm1*-DsRed-positive population contained diploid cells (Figure 4A), and we chose not to select cells for DNA content to avoid interference with sequencing. Bisulfite sequencing revealed male germ cell-specific differential methylation at the imprinted *H19* and *Snrpn* loci comparable with that of in vivo round spermatids (Figure 5C).

Embryos resulting from ICSI with BVSC-derived SLCs (Figure S5A) expressed *Stella*-ECFP at the four-cell stage (Figure S5B). Embryo transfer resulted in three live pups that were positive for the *Blimp1*-mVenus and *Stella*-ECFP transgenes, had a normal karyotype, and exhibited normal weight gain after birth (Figures S5C–S5F). The birth rate following ICSI with round spermatids isolated from normal testis was 9.5% (Table 1). Genome-wide reduced representation bisulfite sequencing (RRBS) of tail tip fibroblast of a male and a female offspring derived from BVSC ESCs was performed to analyze the whole-genome methylation status of the offspring with normal male and female mice as controls. The genome-wide methylation data of sperm and oocyte by Shen et al. (2014) were included in the analysis. The proportions of high methylation sites (>80%)

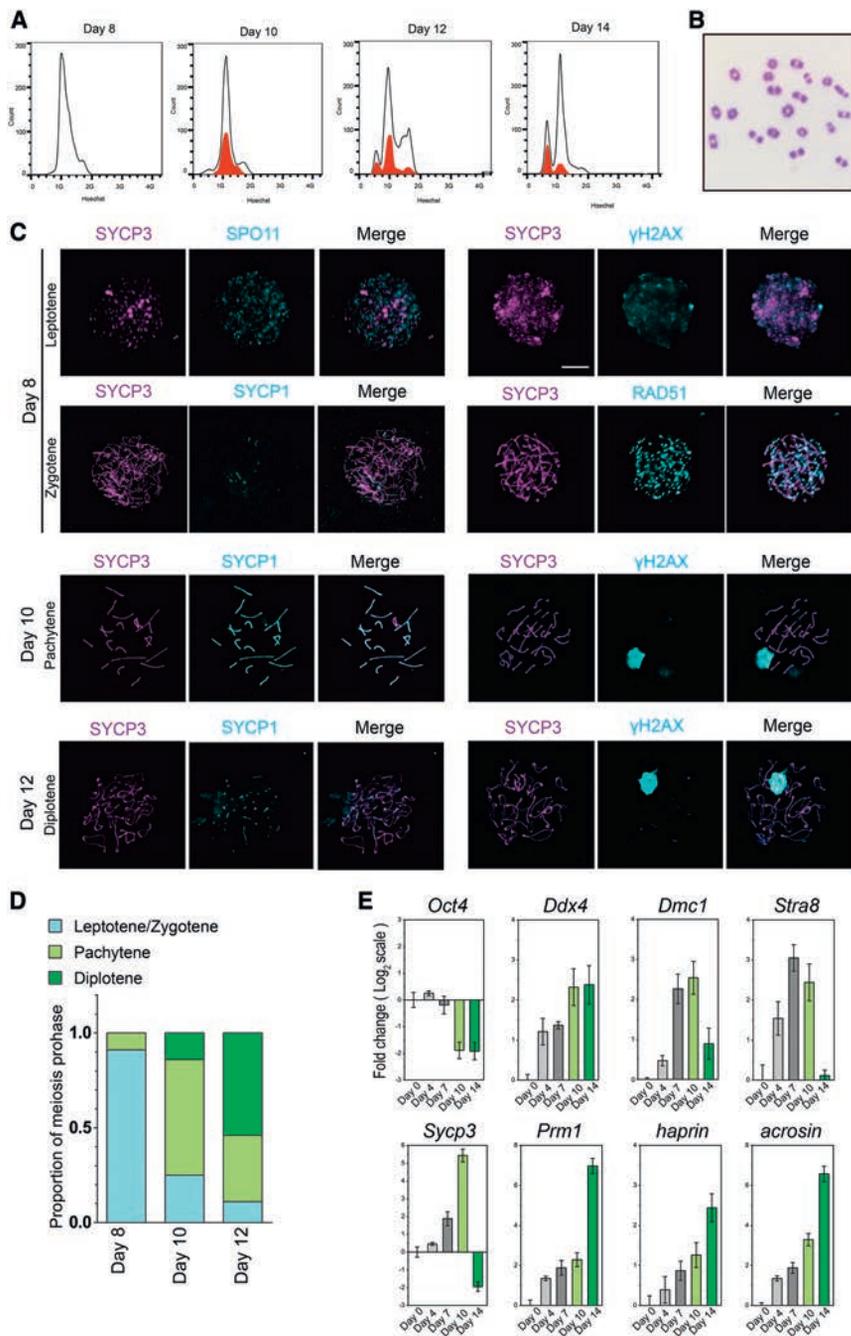


Figure 4. Chromosomal Synapsis and Recombination during In Vitro Meiosis

(A) FACS analysis of DNA contents in spermatogenic cultures from days 8–14. Red histograms reflect *Prm1*-DsRed-positive cells.

(B) Metaphase spread of a day 12 spermatocyte.

(C) Localization of SYCP3, SPO11, γ H2AX, SPO11, and RAD51 in nuclear spreads of germ cells on days 8–12 of co-culture. Scale bar, 100 μ m.

(D) Proportion of leptotene/zygotene-, pachytene-, and diplotene-stage spermatocytes in day 8–12 SLCs. The stages were identified according to SYCP3 distribution.

(E) qRT-PCR analysis of cultures from days 0–14. Relative gene expression levels were normalized to *Rps2* and reflect mean \pm SD from three independent biological replicates.

See also Figure S3.

summary, we describe the successful generation of ESC-derived spermatids in vitro that fully conform to the gold standards proposed for in vitro-derived germ cells (Handel et al., 2014).

DISCUSSION

Here we report the first successful generation of functional spermatids, conforming to the gold standards of in-vitro-derived germ cells, from pluripotent stem cells by stepwise differentiation in vitro. Specifically, we demonstrate that mouse ESC-derived PGCLCs entered meiosis in vitro, undergoing key processes of in vivo meiosis, including chromosomal synapsis and recombination, and finally differentiated into haploid SLCs. These SLCs successfully fertilized oocytes by ICSI, and the resulting embryos underwent embryonic development, resulting in fertile offspring that gave birth to the next generation. This unequivocally demonstrates the recapitulation of meiosis in a culture environment and proves the

of the BVSC-derived offspring and control mice were all lower than that of sperm but higher than that of the oocyte (Figure S5G), which is consistent with the lower methylation levels of somatic cells. The clustering analysis showed that the BVSC-derived male offspring was clustered together with the male control (Figure S5H). Additionally, imprinting regions from BVSC-derived offspring and control mice were analyzed, and all showed a 50% methylation level (Figure S5I). Therefore, BVSC ESC-derived offspring had a normal methylation level. All of these mice developed to adulthood and had offspring (Figure S5J). These data demonstrate the generation of functional spermatids in vitro from both SGPD and BVSC ESC lines. In

functionality of spermatids generated from pluripotent stem cells in vitro.

To establish an efficient approach for in vitro meiosis and gametogenesis, we adapted a method used to differentiate EpiLCs into PGCLCs from a protocol published previously (Hayashi et al., 2011) using N2B27 as basal medium for PGCLC induction. N2B27 is a chemically defined medium that contains the RA precursor vitamin A and insulin. This medium supports the ground-state pluripotency of ESCs when supplemented with 2i and is used frequently as a basal medium to promote neural differentiation from pluripotent cells. We presume that continued exposure to N2B27, which constituted the basal

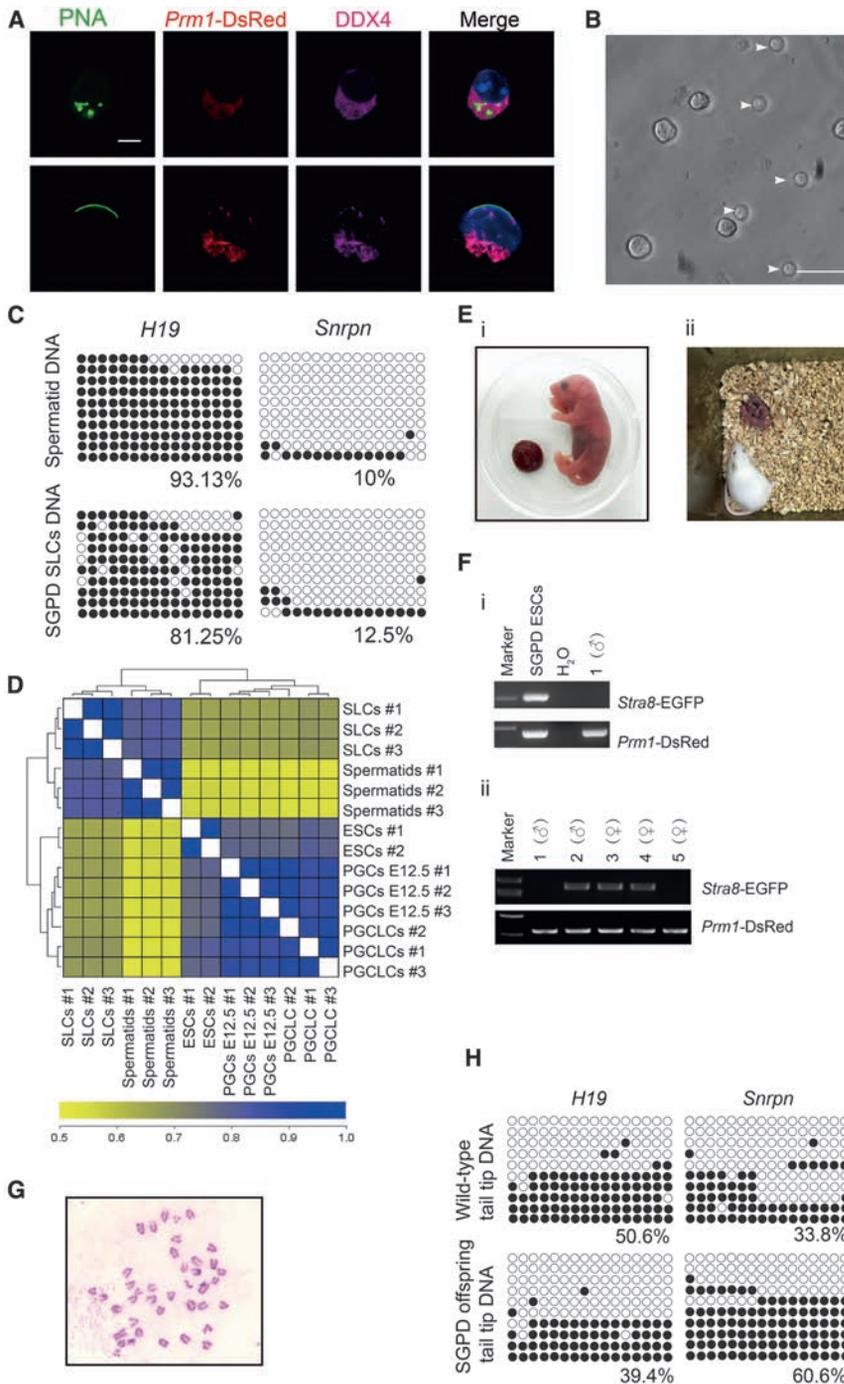


Figure 5. Healthy Offspring Produced from In Vitro-Derived SLCs

(A) Acrosin and PNA staining of SLCs. Scale bar, 5 μ m. (B) Bright-field image of SLCs sorted by FACS. The arrowheads mark small round cells used for single-cell sequencing and ICSI. Scale bar, 50 μ m. (C) Bisulfite sequencing of DMRs of the imprinted genes *Snrpn* and *H19* in wild-type round spermatids and in vitro-derived SLCs. White and black circles indicate unmethylated and methylated CpGs, respectively. Values indicate percent methylation of all CpGs assessed. (D) Unsupervised hierarchical clustering of ESCs, E12.5 male PGCs, day 6 PGCLCs, induced SLCs, and spermatids in vivo according to global gene expression. (E) Live pups obtained by ICSI with PGCLC-derived SLCs. (i) and (ii) show one and five pups from two replicates, respectively. (F) i and ii show genotyping for ESC-derived transgenes of (Ei) and (Eii), respectively. All pups were positive for *Prm1*-DsRed. (G) Metaphase spread confirming the normal karyotype of the pup shown in (Ei). (H) Bisulfite sequencing of *H19* and *Snrpn* DMRs in tail tip fibroblasts from the SLC-derived pup shown in (Ei) and a wild-type control. See also Figures S4 and S5 and Table 1.

(Zhou et al., 2008) because high levels of CYP26B1 expressed by somatic cells of the fetal male gonad degrade endogenous RA (Bowles et al., 2006; MacLean et al., 2007). Here we exposed PGCLCs resembling E12.5 PGCs to conditions supporting their entry into meiosis by identification of morphogens that induced upregulation of meiosis pathways and by providing a culture environment with postnatal somatic testicular cells expressing low CYP26B1 levels (Figure S3A).

We found that simultaneous exposure of PGCLCs to activin A, BMPs, and RA resulted in rapid silencing of *Blimp1* and *Stella* and subsequent upregulation of *Stra8* expression, resulting in initiation of meiosis and changes in gene expression that resemble those of in vivo differentiating germ cells (Figure 1; Kurimoto

et al., 2008; Saitou et al., 2002). Consistent with previous observations demonstrating that BMPs and activin A are required for the self-renewal (Hu et al., 2004; Puglisi et al., 2004) and proliferation of neonatal germ cells (Mithraprabhu et al., 2010), our results also suggest that BMPs and activin A are essential for the proliferation of meiosis-competent PGCLCs in culture, whereas RA is required to induce regulatory networks that lead to meiotic entry and differentiation.

Under our culture conditions, the ESC-derived PGCLCs failed to differentiate into SSCs capable of self-renewal in vitro, evident

medium for EpiLC induction, resulted in the formation of PGCLCs that were capable of undergoing meiosis in vitro. This is supported by our findings from global transcription profiling, which revealed that the ESC-derived PGCLCs generated in our study clustered closely with E12.5 PGCs (Figure 4 D), a developmental stage close to entry into meiosis in vivo. In the mouse germline, female PGCs enter meiosis at E13.5 (McLaren, 2003). This is triggered by the release of endogenous RA from the mesonephros (Anderson et al., 2008). In contrast, in the male, initiation of meiosis remains suppressed until after birth

Table 1. Development of Embryos Generated via ICSI with In Vitro-Differentiated Haploid Spermatids

| Origin of Sperm | Injected Oocytes (n) | Pronuclear Stage Embryos (n) (%) | Two-Cell Stage Embryos (n) (%) | Embryos Transferred (n) | Pups (n) | Pups Surviving to Adulthood (n) |
|-----------------|----------------------|----------------------------------|--------------------------------|-------------------------|----------|---------------------------------|
| Control | 160 | 148 (92.5) | – | 148 | 14 | 14 |
| SGPD (1) | 63 | 59 (93.6) | 51 (86.4) | 51 | 1 | 1 |
| SGPD (2) | 125 | 116 (92.8) | 107 (92.2) | 107 | 5 | 5 |
| BVSC | 191 | 171 (90) | 159 (92.9) | 159 | 3 | 3 |

See also Figure 5 and Figure S5.

from the absence of SSC-specific genes in PGCLC-derived germ cells. We presume that the in vitro culture system lacks features of the testicular microenvironment at the basement membrane required for SSC maintenance, including lack of growth factors required for SSC self-renewal, such as glial cell line-derived neurotrophic factor (GDNF), bFGF, or EGF, among others (Kanatsu-Shinohara et al., 2003; Kanatsu-Shinohara and Shinohara, 2013; Kubota et al., 2004).

We observed that, during induction of meiotic differentiation in co-cultures, *KIT^W/KIT^{W-V}* neonatal testicular cells formed colonies with PGCLCs by active cellular migration. Stimulated by RA, the PGCLCs synchronously passed a SSCs-like state and entered meiosis, reminiscent of PGCs in the E13.5 female genital ridge, which simultaneously enter meiosis at this stage (McLaren, 2003).

We found that differentiation of PGCLCs into male postmeiotic germ cells in vitro required simultaneous exposure to the sex hormones testosterone, FSH, and BPE. This reflects the dependence of in vivo spermatogenesis on pituitary FSH and locally produced testicular androgens, including testosterone. FSH supports Sertoli cell proliferation and stimulates mitotic division of spermatogonia, maintaining adequate cell counts (O'Shaughnessy, 2014). The requirement for BPE for spermatogenesis in vitro suggests that other pituitary factors promote meiotic progression and spermatid differentiation. These may include luteinizing hormone (LH), which normally stimulates the secretion of testosterone from Leydig cells but has also been implied in the maintenance of meiotic germ cells (O'Shaughnessy et al., 2009). The analysis and screening of pituitary tissues for factors affecting in vitro germ cell differentiation may further improve protocols for in vitro meiosis.

In summary, we demonstrate a robust approach toward the stepwise differentiation of pluripotent stem cells into haploid SLCs in vitro. The in vitro meiosis fully complies with the gold standards of meiosis, including erasure of imprints, synapsis, and recombination. Our findings could facilitate the generation of haploid human spermatids in vitro with the prospect of treating male infertility.

EXPERIMENTAL PROCEDURES

Derivation of BVSC and SGPD ESCs

BVSC transgenic mice (Ohinata et al., 2008) were provided by Mitinori Saitou. *Stra8*-EGFP transgenic mice and *Prrm1*-DsRed transgenic mice were generated by pronuclear injection of *Stra8*-EGFP and *Prrm1*-DsRed plasmids (Nayerina et al., 2006), a gift from Wolfgang Engel (University of Göttingen). ESCs were derived from blastocyst-stage embryos by standard culture on mouse feeder layers in 2i medium (Ying et al., 2008). For feeder-free culture, ESCs were maintained on dishes coated with poly-L-ornithine (0.01%; Sigma) and laminin (300 ng/ml, Invitrogen). All cell lines were negative for mycoplasma.

All animal experiments were performed in compliance with the guidelines of the Institute of Zoology, Chinese Academy of Sciences.

Induction of EpiLCs and PGCLCs

Differentiation of ESCs into EpiLCs and PGCLCs was induced by culture conditions adapted from a protocol published previously with minor modifications (Hayashi et al., 2011). For EpiLC differentiation, 1×10^5 ESCs/well were plated in a 12-well plate coated with 16.7 μ g/ml human plasma fibronectin in N2B27 medium supplemented with activin A, bFGF, and 1% knockout serum replacement (KSR). For PGCLC formation in floating culture, 2×10^3 EpiLCs/well were plated in a low cell-binding, U-bottom, 96-well plate (Corning Life Sciences) in modified N2B27 medium (N2B27 with 15% KSR, BMP-4, LIF, SCF, BMP-8a, and EGF). Cells were cultured in 5% CO₂ at 37°C. The medium was changed daily.

In Vitro Spermatogenesis

Testes of 2- to 8-day postpartum (dpp) *Kit^W/Kit^{W-V}* mice were harvested and digested by a two-step enzyme digestion method as described previously (Bellvé, 1993; Bellvé et al., 1977). Briefly, testes were dispersed with 1 mg/ml collagenase type IV at 37°C for 10 min, followed by digestion in 0.25% trypsin/1 mM EDTA for 10 min at 37°C. A single cell suspension was obtained after filtration through a 70- μ m cell strainer, and cells were collected by centrifugation. PGCLCs were mixed with *Kit^W/Kit^{W-V}* mouse testicular cells at a ratio of 1 to 1. From day 0 to day 6, cells were cultured in α MEM supplemented with 10% KSR, BMP-2/4/7 (20 ng/ml each, R&D Systems), retinoic acid (10^{-6} M, Sigma), and activin A (100 ng/ml, R&D Systems). From days 7–14, cells were cultured in α MEM containing 10% KSR, testosterone (10 μ M, Acros Organics), FSH (200 ng/ml, Sigma), and BPE (50 μ g/ml, Corning Life Sciences). The medium was changed every 2 days. Cells were cultured in 5% CO₂ at 37°C.

Immunostaining

Cells or seminiferous tubules were fixed for 15 min with 4% paraformaldehyde at room temperature, blocked for 30 min with 0.3% Triton X-100/2% BSA in PBS, and incubated with primary antibodies against OCT4 (Santa Cruz Biotechnology), NANOG (Millipore), SSEA1 (Millipore), DDX4 (Abcam), BLIMP1 (Abcam), STRA8 (Abcam), and GATA4 (Abcam). After overnight incubation at 4°C, samples were washed three times in PBS, followed by incubation with secondary antibodies or/and peanut agglutinin (PNA) (10 μ g/ml, Sigma) for 1 hr. Secondary antibodies were labeled with fluorescein isothiocyanate (FITC), Cy3, and Cy5 (Jackson ImmunoResearch). DNA was counterstained with 10 μ g/ml Hoechst 33342 for 15 min, followed by three washes with PBS. Images were captured with a Zeiss LSM780 Meta inverted confocal microscope.

Western Blotting

PGCLCs were collected into cell lysis buffer (10 mM Tris-HCl [pH 8.0], 10 mM NaCl, and 0.5% NP-40) containing protease inhibitor (Roche) for 30 min on ice. Lysates were centrifuged at 12,000 \times g for 20 min at 4°C, and the resulting supernatants separated by electrophoresis and western blotting using antibodies against H3K27me3 (Millipore), H3K9me2 (Millipore), and H3 (Millipore).

Karyotype

After culture in medium supplemented with 0.025% colchicine for 6–8 hr, cells were subject to hypotonic treatment with 1% sodium citrate for 30 min at room temperature (RT), followed by fixation in freshly prepared methanol/acetic acid

(3:1) for 2 hr with three replacements of fixative. The coverslip was removed using dry ice. Chromosomes were visualized by Giemsa staining. Images were captured on a Leica DM 6000 B microscope.

DNA and RNA Isolation and Real-Time PCR

DNA and RNA from mouse tail tips or cell pellets were extracted with a MicroElute genomic DNA kit (OMEGA) or an RNeasy micro/mini kit (QIAGEN), respectively. Reverse transcription was performed using a QuantiTect reverse transcription kit (QIAGEN). PCR and real-time PCR were performed with gene-specific primers (Table S1). All gene expression analyses were performed with samples from three independent differentiation experiments.

Flow Cytometry Analysis

PGCLCs were dissociated in 0.25% trypsin/1 mM EDTA, re-suspended in PBS supplemented with 1% BSA, filtered through a 40- μ m nylon mesh, and incubated with SSEA1-AF647-conjugated mouse monoclonal immunoglobulin M (IgM) (eBioscience) and Integrin β -FITC-conjugated mouse monoclonal IgG1 (BioLegend) for 30 min at 37°C. For ploidy analysis, single-cell suspensions were stained with 10 μ g/ml Hoechst 33342 for 20 min and washed three times with PBS. FACS analysis was performed using the FACS Calibur system (Becton Dickinson).

Bisulfite Sequencing

Sodium bisulfite treatment of DNA was performed using the EZ DNA methylation-direct kit (Zymo Research). PCR amplification was performed using hot start (HS) DNA polymerase (TAKARA) with specific primers for *H19* and *Snprn* DMR imprinting regions (Table S1). The PCR product was gel-extracted, subcloned into the pMD18T vector (TAKARA), and sequenced. The resulting data were analyzed using a web-based tool, Quantification Tool for Methylation Analysis (QUMA, <http://quma.cdb.riken.jp/>).

Chromosomal Spreads

Cultured cells were digested into single-cell suspensions. Chromosomal spreads were prepared using a hypotonic bursting technique (Peters et al., 1997). Primary antibodies were Sycp3 (Abcam), Sycp1 (Abcam), γ H2AX (Abcam), Rad 51 (Santa Cruz), and Spo11 (provided by Scott Keeney) (Lange et al., 2011). Secondary antibodies were FITC-, Cy3-, Cy5-, and DyLight 405-labeled (Jackson ImmunoResearch). Images were captured with Zeiss LSM780 Meta inverted confocal microscope. Super-resolution analysis was performed using a Zeiss Elyra PS.1 microscope system.

Global Expression Analysis

Global transcription profiles of SSEA1/integrin β double-positive, day 6 PGCLCs, SSEA1/integrin β double-positive PGCs from male E12.5 fetuses, and individual sorted SLCs and spermatids were determined by microarray as described previously (NCBI GEO: GSE71478) (Zhao et al., 2009). Published ESC data (NCBI GEO: GSE16925) (Zhao et al., 2009) were used for unsupervised hierarchical clustering.

Intracytoplasmic Sperm Injection and Embryo Transfer

ICSI was performed as described previously (Li et al., 2012). SLCs were exposed to 5 μ g/ml cytochalasin B in M2, and individual cells were injected into pre-activated mature oocytes with a Piezo-driven pipette, followed by culture in activation medium for 5 hr. Two-cell embryos were transferred to the oviduct of CD1 pseudopregnant females or further cultured to blastocyst stage in vitro, followed by embryo transfer into the uteri of recipient females. Full-term pups were delivered naturally or by cesarean section.

RRBS Library Preparation and Data Analysis

For RRBS, the libraries were generated and the data were analyzed as described previously (Gu et al., 2011; Shen et al., 2014; Yamaguchi et al., 2013). Paired-end sequencing was performed on an Illumina HiSeq 2500 sequencer (NCBI GEO: GSE76238). RRBS datasets of wild-type meiosis II (MI) oocyte and sperm were downloaded from the GEO database (GSE61331) (Shen et al., 2014). The sequencing reads were trimming by Trim Galore (Babraham Bioinformatics) with the “-rrbs” option and then mapped to the mouse genome (mm9 version) by Bismark v0.13.1 (Babraham Bioinformatics). The methylation levels of covered cytosine sites were calcu-

lated by dividing the number of reported C with the total number of reported C and T. Only the CpG sites that were covered by no less than ten reads were used for the next analysis. The hierarchical clustering was produced by the hcluster functions of R with the “euclidean” and “ward” Parameters. The Pearson correlation coefficients were generated using the lm function in R. The heatmaps were plotted by the heatmap.2 function in R. Histograms of methylation level distribution were drawn by ggplot2. Point plots were produced by the smoothScatter functions in R.

ACCESSION NUMBERS

The accession number for the RNA sequencing data reported in this paper is GEO: GSE71478. The accession number for the RRBS data reported in this paper is GEO: GSE76238.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, and two movies and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.01.017>.

AUTHOR CONTRIBUTIONS

Conceptualization, X.-Y.Z., J.S., and Q.Z.; Methodology, Q.S.; Investigation, Q.Z., Y.Y., Y.Z., M.W., X.W., R.F., H.W., M.X., M.L., and X.G.; Data Curation, G.F.; Writing, M.L., X.G., X.-Y.Z., J.S., and Q.Z., with all authors approving the final version; Funding Acquisition, X.-Y.Z., J.S., and Q.Z.; Supervision, X.-Y.Z., J.S., and Q.Z.

ACKNOWLEDGMENTS

We thank Wolfgang Engel for the *Stra8*-EGFP and *Prrm1*-DsRed plasmids, Mitinori Saitou for *Blimp1*-mVenus and *Stella*-ECFP (BVSC) transgenic mice, Scott Keeney for the SPO11 antibody, and Sigrid Eckardt and Xingxu Huang for help with manuscript preparation. This study was funded by the 973 Program (2011CB944304, 2012CBA01301, 2012CBA01300, and 2012CB966500).

Received: November 17, 2015

Revised: December 23, 2015

Accepted: January 21, 2016

Published: February 25, 2016

REFERENCES

- Anderson, E.L., Baltus, A.E., Roepers-Gajadien, H.L., Hassold, T.J., de Rooij, D.G., van Pelt, A.M., and Page, D.C. (2008). *Stra8* and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. *Proc. Natl. Acad. Sci. USA* *105*, 14976–14980.
- Aravind, L., Leipe, D.D., and Koonin, E.V. (1998). Toprim—a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res.* *26*, 4205–4213.
- Bellvé, A.R. (1993). Purification, culture, and fractionation of spermatogenic cells. *Methods Enzymol.* *225*, 84–113.
- Bellvé, A.R., Cavicchia, J.C., Millette, C.F., O'Brien, D.A., Bhatnagar, Y.M., and Dym, M. (1977). Spermatogenic cells of the prepubertal mouse. Isolation and morphological characterization. *J. Cell Biol.* *74*, 68–85.
- Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M.J., Rossant, J., et al. (2006). Retinoid signaling determines germ cell fate in mice. *Science* *312*, 596–600.
- Eguizabal, C., Montserrat, N., Vassena, R., Barragan, M., Garreta, E., Garcia-Quevedo, L., Vidal, F., Giorgetti, A., Veiga, A., and Izpisua Belmonte, J.C. (2011). Complete meiosis from human induced pluripotent stem cells. *Stem Cells* *29*, 1186–1195.
- Eijpe, M., Heyting, C., Gross, B., and Jessberger, R. (2000). Association of mammalian SMC1 and SMC3 proteins with meiotic chromosomes and synaptonemal complexes. *J. Cell Sci.* *113*, 673–682.

- Geijsen, N., Horoschak, M., Kim, K., Gribnau, J., Eggan, K., and Daley, G.Q. (2004). Derivation of embryonic germ cells and male gametes from embryonic stem cells. *Nature* 427, 148–154.
- Gu, H., Smith, Z.D., Bock, C., Boyle, P., Gnirke, A., and Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* 6, 468–481.
- Handel, M.A., Eppig, J.J., and Schimenti, J.C. (2014). Applying “gold standards” to in-vitro-derived germ cells. *Cell* 157, 1257–1261.
- Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* 146, 519–532.
- Hu, J., Chen, Y.X., Wang, D., Qi, X., Li, T.G., Hao, J., Mishina, Y., Garbers, D.L., and Zhao, G.Q. (2004). Developmental expression and function of Bmp4 in spermatogenesis and in maintaining epididymal integrity. *Dev. Biol.* 276, 158–171.
- Irie, N., Weinberger, L., Tang, W.W., Kobayashi, T., Viukov, S., Manor, Y.S., Dietmann, S., Hanna, J.H., and Surani, M.A. (2015). SOX17 is a critical specifier of human primordial germ cell fate. *Cell* 160, 253–268.
- Kanatsu-Shinohara, M., and Shinohara, T. (2013). Spermatogonial stem cell self-renewal and development. *Annu. Rev. Cell Dev. Biol.* 29, 163–187.
- Kanatsu-Shinohara, M., Ogonuki, N., Inoue, K., Miki, H., Ogura, A., Toyokuni, S., and Shinohara, T. (2003). Long-term proliferation in culture and germline transmission of mouse male germline stem cells. *Biol. Reprod.* 69, 612–616.
- Keeney, S., Giroux, C.N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* 88, 375–384.
- Koubova, J., Hu, Y.C., Bhattacharyya, T., Soh, Y.Q., Gill, M.E., Goodheart, M.L., Hogarth, C.A., Griswold, M.D., and Page, D.C. (2014). Retinoic acid activates two pathways required for meiosis in mice. *PLoS Genet.* 10, e1004541.
- Kubota, H., Avarbock, M.R., and Brinster, R.L. (2004). Growth factors essential for self-renewal and expansion of mouse spermatogonial stem cells. *Proc. Natl. Acad. Sci. USA* 101, 16489–16494.
- Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes Dev.* 22, 1617–1635.
- Lammers, J.H., Offenberger, H.H., van Aalderen, M., Vink, A.C., Dietrich, A.J., and Heyting, C. (1994). The gene encoding a major component of the lateral elements of synaptonemal complexes of the rat is related to X-linked lymphocyte-regulated genes. *Mol. Cell. Biol.* 14, 1137–1146.
- Lange, J., Pan, J., Cole, F., Thelen, M.P., Jasin, M., and Keeney, S. (2011). ATM controls meiotic double-strand-break formation. *Nature* 479, 237–240.
- Li, W., Dou, Z.Y., Hua, J.L., and Wang, H.Y. (2007). [Activation of Stra 8 gene during the differentiation of spermatogonial stem cells]. *Sheng Wu Gong Cheng Xue Bao* 23, 639–644.
- Li, W., Shuai, L., Wan, H., Dong, M., Wang, M., Sang, L., Feng, C., Luo, G.Z., Li, T., Li, X., et al. (2012). Androgenetic haploid embryonic stem cells produce live transgenic mice. *Nature* 490, 407–411.
- MacLean, G., Li, H., Metzger, D., Chambon, P., and Petkovich, M. (2007). Apoptotic extinction of germ cells in testes of Cyp26b1 knockout mice. *Endocrinology* 148, 4560–4567.
- Mahadevaiah, S.K., Turner, J.M., Baudat, F., Rogakou, E.P., de Boer, P., Blanco-Rodríguez, J., Jasin, M., Keeney, S., Bonner, W.M., and Burgoyne, P.S. (2001). Recombinational DNA double-strand breaks in mice precede synapsis. *Nat. Genet.* 27, 271–276.
- McLaren, A. (2003). Primordial germ cells in the mouse. *Dev. Biol.* 262, 1–15.
- Menke, D.B., Koubova, J., and Page, D.C. (2003). Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. *Dev. Biol.* 262, 303–312.
- Meuwissen, R.L., Offenberger, H.H., Dietrich, A.J., Riesewijk, A., van Iersel, M., and Heyting, C. (1992). A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. *EMBO J.* 11, 5091–5100.
- Mithraprabhu, S., Mendis, S., Meachem, S.J., Tubino, L., Matzuk, M.M., Brown, C.W., and Loveland, K.L. (2010). Activin bioactivity affects germ cell differentiation in the postnatal mouse testis in vivo. *Biol. Reprod.* 82, 980–990.
- Nayernia, K., Nolte, J., Michelmann, H.W., Lee, J.H., Rathsack, K., Drusenheimer, N., Dev, A., Wulf, G., Ehrmann, I.E., Elliott, D.J., et al. (2006). In vitro-differentiated embryonic stem cells give rise to male gametes that can generate offspring mice. *Dev. Cell* 11, 125–132.
- Neale, M.J., and Keeney, S. (2006). Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature* 442, 153–158.
- O’Shaughnessy, P.J. (2014). Hormonal control of germ cell development and spermatogenesis. *Semin. Cell Dev. Biol.* 29, 55–65.
- O’Shaughnessy, P.J., Morris, I.D., Huhtaniemi, I., Baker, P.J., and Abel, M.H. (2009). Role of androgen and gonadotrophins in the development and function of the Sertoli cells and Leydig cells: data from mutant and genetically modified mice. *Mol. Cell. Endocrinol.* 306, 2–8.
- Ohinata, Y., Sano, M., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). A comprehensive, non-invasive visualization of primordial germ cell development in mice by the Prdm1-mVenus and Dppa3-ECFP double transgenic reporter. *Reproduction* 136, 503–514.
- Peters, A.H., Plug, A.W., van Vugt, M.J., and de Boer, P. (1997). A drying-down technique for the spreading of mammalian meocytes from the male and female germline. *Chromosome Res.* 5, 66–68.
- Puglisi, R., Montanari, M., Chiarella, P., Stefanini, M., and Boitani, C. (2004). Regulatory role of BMP2 and BMP7 in spermatogonia and Sertoli cell proliferation in the immature mouse. *Eur. J. Endocrinol.* 151, 511–520.
- Saitou, M., Barton, S.C., and Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. *Nature* 418, 293–300.
- Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M., and Matsui, Y. (2005). Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev. Biol.* 278, 440–458.
- Shen, L., Inoue, A., He, J., Liu, Y., Lu, F., and Zhang, Y. (2014). Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* 15, 459–470.
- Sun, Y.C., Cheng, S.F., Sun, R., Zhao, Y., and Shen, W. (2014). Reconstitution of gametogenesis in vitro: meiosis is the biggest obstacle. *Yi Chuan Xue Bao* 41, 87–95.
- Yamaguchi, S., Shen, L., Liu, Y., Sandler, D., and Zhang, Y. (2013). Role of Tet1 in erasure of genomic imprinting. *Nature* 504, 460–464.
- Ying, Q.L., Wray, J., Nichols, J., Battle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* 453, 519–523.
- Yokonishi, T., Sato, T., Katagiri, K., Komeya, M., Kubota, Y., and Ogawa, T. (2013). In vitro reconstruction of mouse seminiferous tubules supporting germ cell differentiation. *Biol. Reprod.* 89, 15.
- Zhao, X.Y., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C.L., Ma, Q.W., Wang, L., et al. (2009). iPS cells produce viable mice through tetraploid complementation. *Nature* 461, 86–90.
- Zhou, Q., Li, Y., Nie, R., Friel, P., Mitchell, D., Evanoff, R.M., Pouchnik, D., Banasik, B., McCarrey, J.R., Small, C., and Griswold, M.D. (2008). Expression of stimulated by retinoic acid gene 8 (Stra8) and maturation of murine gonocytes and spermatogonia induced by retinoic acid in vitro. *Biol. Reprod.* 78, 537–545.

Mesenchymal Inflammation Drives Genotoxic Stress in Hematopoietic Stem Cells and Predicts Disease Evolution in Human Pre-leukemia

Noemi A. Zambetti,^{1,8} Zhen Ping,^{1,8} Si Chen,^{1,8} Keane J.G. Kenswil,¹ Maria A. Mylona,¹ Mathijs A. Sanders,¹ Remco M. Hoogenboezem,¹ Eric M.J. Bindels,¹ Maria N. Adisty,¹ Paulina M.H. Van Strien,¹ Cindy S. van der Leije,² Theresia M. Westers,³ Eline M.P. Cremers,³ Chiara Milanese,⁴ Pier G. Mastroberardino,⁴ Johannes P.T.M. van Leeuwen,² Bram C.J. van der Eerden,² Ivo P. Touw,¹ Taco W. Kuijpers,⁵ Roland Kanaar,⁶ Arjan A. van de Loosdrecht,³ Thomas Vogl,⁷ and Marc H.G.P. Raaijmakers^{1,9,*}

¹Department of Hematology, Erasmus MC Cancer Institute, Rotterdam 3015CN, the Netherlands

²Department of Internal Medicine, Erasmus MC Cancer Institute, Rotterdam 3015CN, the Netherlands

³Department of Hematology, VU University Medical Center, Cancer Center Amsterdam, Amsterdam 1081HV, the Netherlands

⁴Department of Molecular Genetics, Erasmus MC Cancer Institute, Rotterdam 3015CN, the Netherlands

⁵Department of Pediatric Hematology, Immunology and Infectious Diseases, Emma Children's Hospital, Academic Medical Centre (AMC), University of Amsterdam (UvA), Amsterdam 1105AZ, the Netherlands

⁶Departments of Genetics and Radiation Oncology, Cancer Genomics Center, Erasmus MC Cancer Institute, Rotterdam 3015CN, the Netherlands

⁷Institute of Immunology, University of Münster, Münster 48149, Germany

⁸Co-first author

⁹Lead Contact

*Correspondence: m.h.g.raaijmakers@erasmusmc.nl

<http://dx.doi.org/10.1016/j.stem.2016.08.021>

SUMMARY

Mesenchymal niche cells may drive tissue failure and malignant transformation in the hematopoietic system, but the underlying molecular mechanisms and relevance to human disease remain poorly defined. Here, we show that perturbation of mesenchymal cells in a mouse model of the pre-leukemic disorder Shwachman-Diamond syndrome (SDS) induces mitochondrial dysfunction, oxidative stress, and activation of DNA damage responses in hematopoietic stem and progenitor cells. Massive parallel RNA sequencing of highly purified mesenchymal cells in the SDS mouse model and a range of human pre-leukemic syndromes identified p53-S100A8/9-TLR inflammatory signaling as a common driving mechanism of genotoxic stress. Transcriptional activation of this signaling axis in the mesenchymal niche predicted leukemic evolution and progression-free survival in myelodysplastic syndrome (MDS), the principal leukemia predisposition syndrome. Collectively, our findings identify mesenchymal niche-induced genotoxic stress in heterotypic stem and progenitor cells through inflammatory signaling as a targetable determinant of disease outcome in human pre-leukemia.

INTRODUCTION

Genotoxic stress results in the accumulation of DNA lesions in hematopoietic stem and progenitor cells (HSPCs) over the life-

span of an organism, contributing to tissue failure and malignant transformation (Jaiswal et al., 2014; Rossi et al., 2007). The pathophysiological insults underlying genomic stress in HSPCs, however, remain incompletely understood. Perturbed signaling from their surrounding microenvironment may be implicated, but this has not been experimentally defined.

Components of the bone marrow microenvironment have emerged as key regulators of normal and malignant hematopoiesis (Arranz et al., 2014; Hanoun et al., 2014; Medyouf et al., 2014; Schepers et al., 2015; Walkley et al., 2007). We, and others, have shown that primary alterations of the mesenchymal niche can induce myelodysplasia and promote the emergence of acute myeloid leukemia (AML) with cytogenetic abnormalities in HSPCs (Kode et al., 2014; Raaijmakers et al., 2010), thus introducing a concept of niche-driven oncogenesis in the hematopoietic system.

To provide insights into the mechanisms that underlie this concept, as well as their relevance for human disease, we modeled the human leukemia predisposition disorder Shwachman-Diamond syndrome (SDS), caused by constitutive homozygous or compound heterozygous loss-of-function mutations in the *SBDS* gene, required for ribosome biogenesis (Boocock et al., 2003; Finch et al., 2011). SDS is characterized by skeletal defects in conjunction with a striking propensity to develop myelodysplastic syndrome (MDS) and AML at a young age, with a cumulative probability of >30% at the age of 30 years and a median onset at 18 years (Alter, 2007; Donadieu et al., 2012). Hematopoietic cell intrinsic loss of *Sbds* does not result in MDS or leukemia (Rawls et al., 2007; Zambetti et al., 2015), supporting the notion that cell-extrinsic factors contribute to malignant transformation. Deletion of *Sbds* from mesenchymal cells expressing the mesenchymal progenitor marker osterix (*Sp7*) in the bone marrow induced apoptosis in HSPCs and myelodysplasia, but the molecular mechanisms driving these observations

and their relevance for human disease remained to be defined (Raaijmakers et al., 2010).

Here, we identify the endogenous damage-associated molecular pattern (DAMP) molecules S100A8 and S100A9, secreted from mesenchymal niche cells, as drivers of mitochondrial dysfunction, oxidative stress, and DNA damage response (DDR) activation in HSPCs, with clinical relevance to the pathogenesis and prognosis of human bone marrow failure and leukemia predisposition syndromes.

RESULTS

Deletion of *Sbds* from Mesenchymal Progenitor Cells Recapitulates Skeletal Abnormalities of Human SDS

SDS is characterized by bone abnormalities including low-turnover osteoporosis with reduced trabecular bone volume, low numbers of osteoblasts, and reduced amount of osteoid, leading to increased risk of fractures (Toiviainen-Salo et al., 2007). The cellular subsets driving these abnormalities and the underlying molecular mechanisms have remained largely undefined. We have previously shown that Cre-mediated deletion of *Sbds* from osterix⁺ mesenchymal progenitor cells (MPCs) (*Sbds*^{ff} *Osx*^{cre/+} mice, hereafter OCS^{ff} or mutants) disrupts the architecture of the marrow and cortical bone (Raaijmakers et al., 2010). Here, we first sought to better define the skeletal defects in these mice and their relevance to human disease.

OCS^{ff} mice presented growth retardation and reduced femur length compared to control *Sbds*^{fl/+} *Osx*^{cre/+} (OCS^{fl/+}) mice (Figures 1A and 1B) as observed in human patients (Aggett et al., 1980; Ginzberg et al., 1999). The runted phenotype was associated with a significantly limited lifespan, with lethality observed after the age of 4 weeks. Analyses were therefore performed in 3-week-old mice. The femur trabecular area was profoundly reduced in OCS^{ff} mice, with decreased bone volume, low number of trabeculae, increased trabecular spacing, and reduced numbers of osteoblasts compared to controls (Figures 1C–1G and 1I). The cortical bone of OCS mutants was also affected, as indicated by low bone mineral density values (Figures 1C, 1D, and 1H), attenuating the mechanical properties of the bone, which was found less resistant to fracture in three-point bending test (Figure 1J). A tendency for reduced stiffness in the long bones was also observed (Figure 1K). Taken together, the structural and mechanical defects indicate that *Sbds* deficiency in MPCs causes osteoporosis with a propensity for fracturing, in line with observations in SDS patients (Ginzberg et al., 1999; Mäkitie et al., 2004; Toiviainen-Salo et al., 2007). Impaired osteogenesis did not reflect a contraction of the bone MPC pool as shown by frequency of CFU-F and *Osx*::GFP⁺ cells (Figures S1A and S1B), but rather impairment of terminal osteogenic differentiation as suggested by transcriptional profiling of prospectively isolated osterix-expressing (GFP⁺) cells (Figure S1C). Transcriptional data confirmed deregulated expression of genes related to ribosomal biogenesis and translation (Figures S1D and S1E), in line with the established role of *Sbds* in ribosome biogenesis. Collectively, these data support a view in which bone abnormalities in SDS are caused by deficiency of *Sbds* in MPCs, which attenuates terminal differentiation toward matrix-depositing osteoblastic cells with a compensatory increase in the most primitive mesenchymal compartment.

Sbds Deficiency in the Hematopoietic Niche Induces Mitochondrial Dysfunction, Oxidative Stress, and Activation of the DDR in HSPCs

Having established that the OCS mice represent a bona fide model for bone abnormalities in human disease, we next investigated the hematopoietic consequences of these environmental alterations. HSPC number was unaltered in OCS mice (Figures S2A–S2C), and HSPCs displayed global preservation of their transcriptional landscape after exposure to the *Sbds*-deficient environment (Figures S2D–S2F). Transcriptional network analysis, however, revealed significant overlap with signatures previously defined as predicting leukemic evolution of human CD34⁺ cells (Li et al., 2011), including pathways signaling mitochondrial abnormalities (Figure 2A; Table S1). Mitochondrial dysfunction was confirmed by measuring the mitochondrial membrane potential ($\Delta\psi$), indicating hyperpolarization of the mitochondria (Figures 2B and 2C). Mitochondrial hyperpolarization can result in reverse electron transfer, leading to the production of superoxide radicals, which can be further converted into other reactive oxygen species (ROS) (Murphy, 2009). In line with this, a marked increase in intracellular ROS levels was found in OCS mutant HSPCs (Figure 2D), more specifically superoxide radicals derived from mitochondria as shown by dihydroethidium (DHE) staining (Figure S2G) (Owusu-Ansah et al., 2008; Stowe and Camara, 2009). ROS can undermine the genomic integrity of HSPCs by inducing DNA damage (Ito et al., 2006; Walter et al., 2015; Yahata et al., 2011), to which normal HSPCs react by activating the DDR and DNA repair pathways (Rossi et al., 2007). Indeed, HSPCs (LKS-SLAM) from OCS^{ff} mice displayed accumulation of Ser139-phosphorylated H2AX histone (γ H2AX), which forms at the sites of DNA damage (Figures 2E and S2H). Treatment of OCS mutant animals with the ROS scavenger N-acetylcysteine (NAC) resulted in partial reduction in the accumulation of γ H2AX (Figures S2I and S2J). Congruent with genotoxic effects of the mutant microenvironment, HSPCs displayed transcriptional modulation of DDR and DNA repair pathways (Table S2), including nucleotide excision repair programs, associated with ROS-induced lesions (Curtin, 2012) and signatures related to the master regulator of DDR and cell-cycle checkpoint activation ataxia telangiectasia and Rad3-related (ATR). Activation of the G1-S cell-cycle checkpoint, resulting in cell-cycle arrest, was suggested by depletion of S-phase transcriptional signatures (Figure 2F; Table S1), in vivo BrdU/Ki67 labeling (Figures 2G, 2H, and S2K), and downregulation of the *Myc* pathway, a critical regulator for this restriction point and the coordination of S-G2-M progression (Figure 2I; Table S3). Apoptosis of mutant HSPCs, as an alternative outcome of checkpoint activation, was earlier demonstrated (Raaijmakers et al., 2010). Together, the data indicate that the *Sbds*-deficient environment induces mitochondrial dysfunction, oxidative stress, DNA damage, and genotoxic stress in HSPCs leading to activation of DDR pathways and G1-S checkpoint activation, reminiscent of a model in which mitochondrial dysfunction underlies an escalating cycle of increased ROS and genotoxic damage (Sahin and Depinho, 2010).

Short term exposure to the genotoxic environment did not attenuate HSPC function in DNA repair proficient cells, as demonstrated by competitive transplantation experiments

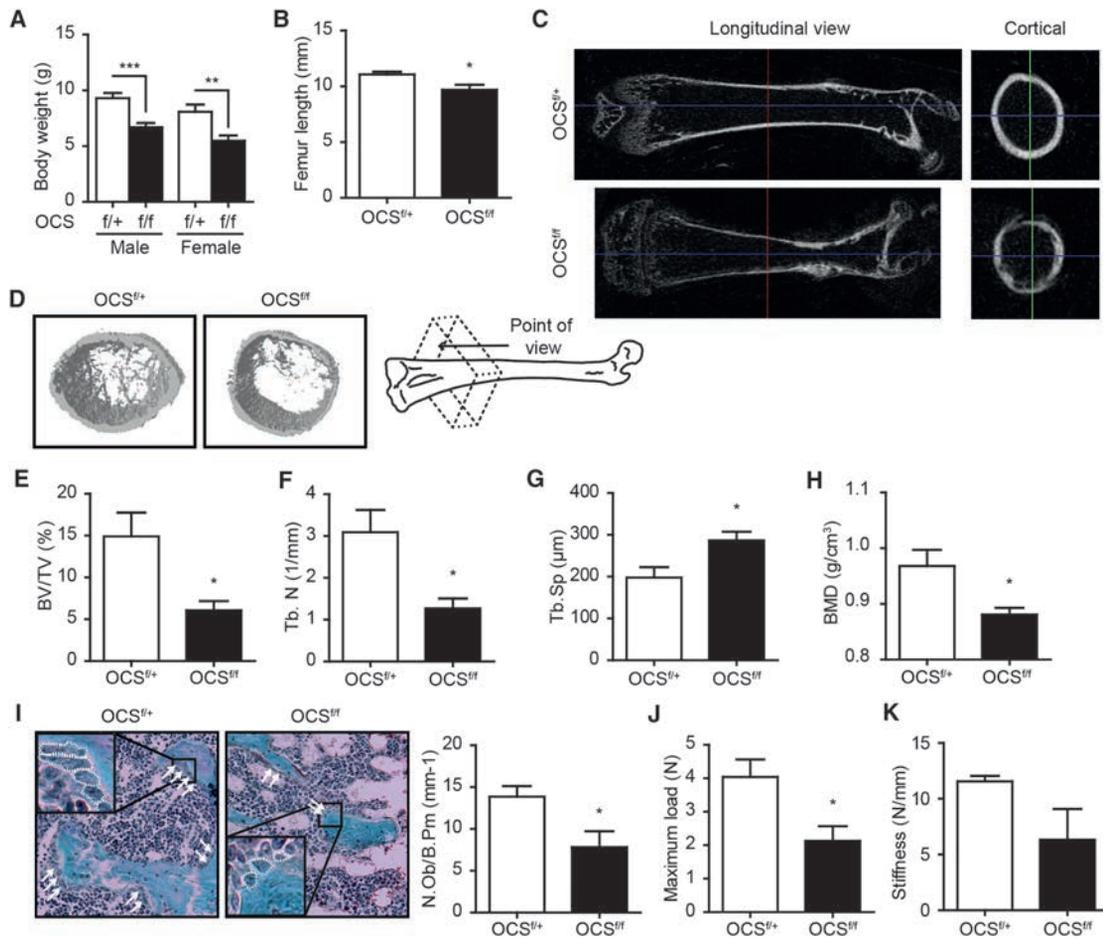


Figure 1. Deletion of *Sbds* in MPCs Recapitulates Skeletal Defects in Human SDS

(A and B) Impaired growth in *OCS^{f/f}* mice: body weight (n = 9) (A) and femur length (n = 5) (B).

(C–H) Femur μ CT analysis of *OCS^{f/+}* (n = 5) and *OCS^{f/f}* (n = 4) mice.

(C) Representative 2D images, left: longitudinal view and right: cortical bone.

(D) 3D image.

(E) Bone volume per tissue volume (BV/TV).

(F) Trabecular number (Tb. N).

(G) Trabecular spacing (Tb. Sp).

(H) Cortical bone mineral density (BMD).

(I) Goldner osteoblast staining (*OCS^{f/+}*, n = 6 and *OCS^{f/f}*, n = 8). The representative images (arrows: osteoblasts; with white dashed line in the magnified region) are shown (left). The number of osteoblasts per bone perimeter (N.Ob/B.Pm) is shown (right).

(J and K) 3-point bending test indicating reduced resistance to fracture (J) and decreased stiffness of *OCS^{f/f}* bone (K) (*OCS^{f/+}*, n = 5 and *OCS^{f/f}*, n = 4) (*p < 0.05, **p < 0.01, and ***p < 0.001). The data are mean \pm SEM.

See also Figure S1.

(Figures S3A–S3C), suggesting efficient DNA-repair or elimination of functionally impaired HSPCs by the DDR-driven apoptosis and cell-cycle arrest. Congruent with this notion, alkaline comet assays on sorted HSPCs failed to demonstrate structural DNA damage (Figures S3D and S3E).

Activation of the p53 Pathway Drives Bone Abnormalities and Genotoxic Stress in *OCS* Mice

Next, we sought to define the molecular programs underlying the bone and hematopoietic alterations in *OCS* mice. A proposed common molecular mechanism for the pathogenesis of ribosomopathies involves activation of the p53 tumor suppressor

pathway (Raiser et al., 2014). The p53 protein was overexpressed in GFP⁺ MPCs in *OCS* mutants, with activation of downstream transcriptional pathways and upregulation of canonical targets (Figures 3A–3C). To assess the pathophysiological role of p53 activation in MPCs, we intercrossed *OCS* with *Trp53*-floxed mice (Marino et al., 2000), generating a double conditional knockout model where the deletion of p53 is localized in the *Sbds*-deleted stromal compartment (*Sbds^{f/f} Trp53^{f/f} Osx^{cre/+}* mice; hence *OCS^{f/f} p53 Δ*) (Figure 3D). Genetic recombination of the *Trp53* locus was detected only in bone cells-containing samples, demonstrating the tissue specificity of p53 deletion in this model (Figure 3E). Genetic deletion of p53 from *Sbds*-deficient

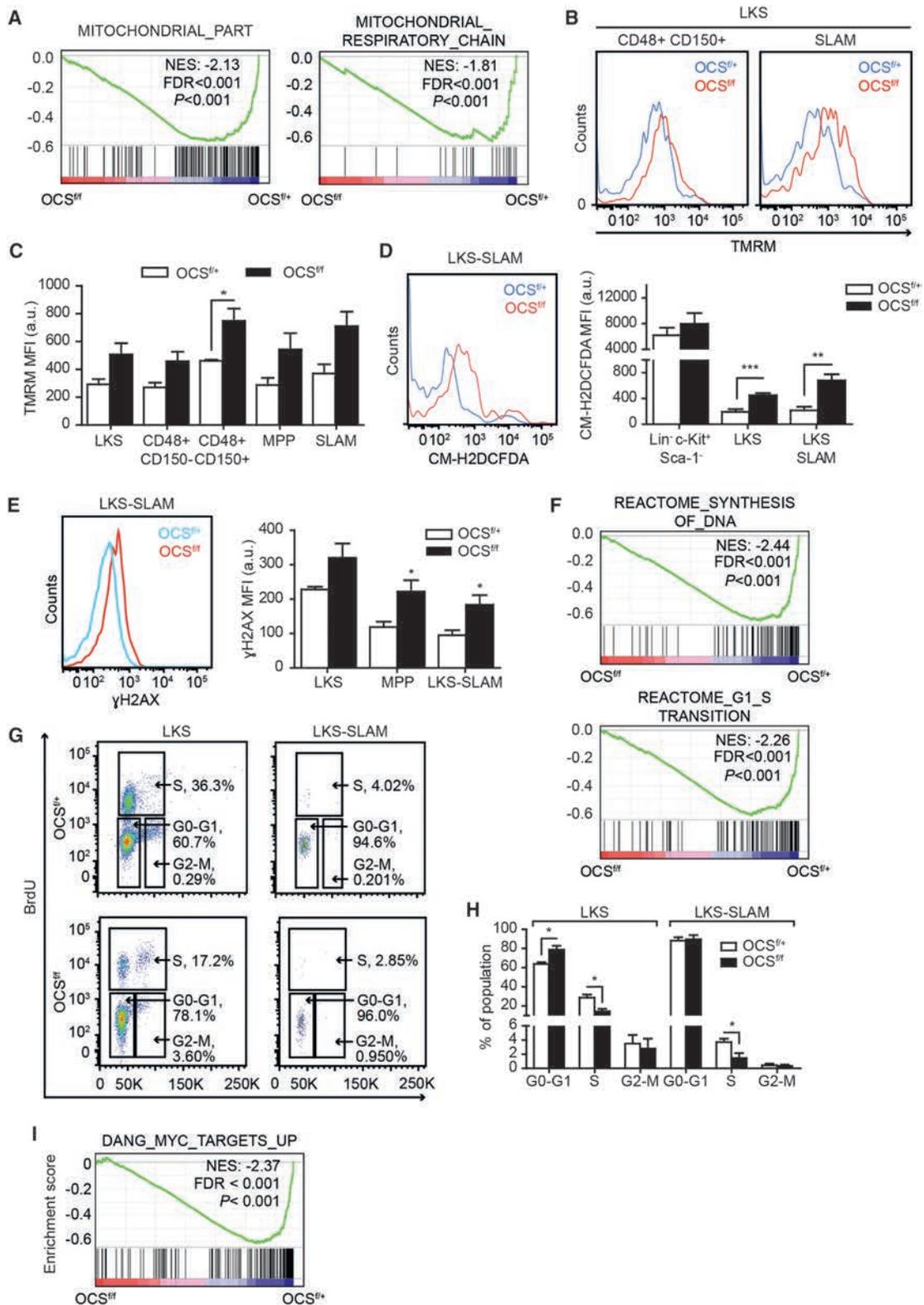


Figure 2. *Sbd5*-Deficient Mesenchymal Cells Induce Genotoxic Stress in HSPCs

(A) Transcriptional network analysis indicating mitochondrial dysregulation in mutant HSPCs. normalized enrichment score: NES. (B and C) Increased mitochondrial potential (TMRM) in HSPCs: representative plots (B); mean fluorescence intensity (MFI) (C) (n = 3). (D) ROS quantification by CM-H2DCFDA (OCS^{+/+}, n = 6 and OCS^{-/-}, n = 7).

(legend continued on next page)

MPCs rescued the osteoporotic phenotype (Figures 3F–3J), but not cortical bone mineralization (Figure 3K), while it had only modest effects on bone mass in OCS control mice (Figure S4), in line with earlier observations (Wang et al., 2006). Rescue of the skeletal phenotype was linked to amelioration of genotoxic stress in HSPCs, as demonstrated by a reduction of superoxide radicals derived from mitochondria and DNA damage (Figures 3L and 3M).

Identification of the DAMP Genes *S100A8* and *S100A9* as Candidate Niche Factors Driving Genotoxic Stress in Human Leukemia Predisposition Syndromes

To identify human disease-relevant niche factors, downstream of p53 activation, driving genomic stress in HSPCs, we compared the transcriptomes of GFP⁺ MPCs from OCS mice to those from prospectively fluorescence-activated cell sorting (FACS)-isolated mesenchymal CD271⁺ niche cells (Tormin et al., 2011) from human SDS patients (Figure 4A; Table S4). The mesenchymal nature of CD271⁺ cells was confirmed by CFU-F capacity and differential expression of mesenchymal, osteolineage, and HSPC-regulatory genes (Chen et al., 2016). RNA sequencing showed the presence of *SBDS* mutations (Figures 4B and S5; Table S4) associated with reduced *SBDS* expression (Figure 4C), confirming molecular aspects of SDS in previous studies (Finch et al., 2011; Woloszynek et al., 2004). Virtually identical transcriptional signatures of disrupted ribosome biogenesis and translation were found in human niche cells (Figure 4D) and in GFP⁺ cells from OCS mice (Figure S1E), confirming faithful recapitulation of human molecular disease characteristics in the mouse model. There were 40 genes that were differentially expressed both in the mouse model and human SDS, 25 of which were overexpressed, with a remarkable abundance of genes encoding proteins implicated in inflammation and innate immunity (Figure 4E).

To further delineate candidate genes driving genomic stress and leukemic evolution from this gene set, we performed whole transcriptome sequencing of CD271⁺ cells in two related human bone marrow failure and leukemia predisposition disorders: (1) low-risk MDS, the principal human pre-leukemic disorder in which cell-cycle exit (senescence), accumulation of ROS, DNA damage, and apoptosis have been described (Head et al., 2011; Peddie et al., 1997; Xiao et al., 2013), reminiscent of HSPC phenotypes in OCS mice and (2) DBA, like SDS, a ribosomopathy characterized by bone marrow failure, but with a much lower propensity to evolve into AML (<1% with longer latency than observed in SDS and MDS) (Vlachos et al., 2012) (Table S4). We reasoned that genes specifically overexpressed in mesenchymal niche cells from disorders with as strong propensity for leukemic evolution (SDS and MDS) might represent strong candidate drivers of genotoxic stress. We found 11 such genes (Figure 4F), among which

were the DAMP genes *S100A8* and *S100A9*, which were significantly ($p < 0.05$) differentially expressed in GFP⁺ cells from OCS mutant mice (Figures 4E–4G) and also represent a bona fide downstream transcriptional target of p53 (Li et al., 2009). Ex vivo small hairpin (sh)RNA experiments confirmed that upregulation of both p53 and *S100A8/9* are direct, cell-intrinsic consequences of *Sbds* downregulation in mesenchymal precursor (OP9) cells (Figure S6A).

Niche-Derived *S100A8/9* Induces Genotoxic Stress in Murine and Human HSPCs

S100A8 and *S100A9* belong to a subclass of proinflammatory molecules referred to as DAMP or alarmins. DAMPs are endogenous danger signals that are passively released or actively secreted in the microenvironment after cell death, damage, or stress and bind pattern recognition receptors (PRR) to regulate inflammation and tissue repair (Srikrishna and Freeze, 2009). *S100A8* and *S100A9* proteins were overexpressed in mouse *Sbds*-deficient MPCs (Figures 5A and 5B) and increased plasma concentration of *S100A8/9* indicated secretion of the heterodimer (Figure 5C). Its canonical receptor TLR4 (Vogl et al., 2007) is expressed in murine HSPCs (Figure S6B), and the canonical downstream signaling NF- κ B and MAPK pathways were activated in HSPCs from OCS^{fff} mice (Figure S6C).

Exposure of HSPCs (LKS) to recombinant murine *S100A8/9* resulted in increased DNA damage (number of γ H2AX and 53BP1 foci) (Figures 5D and S6D), which was replication independent (Figure 5E), and apoptosis (Figure 5F), associated with activation of TLR signaling (Figure 5G; Table S5), recapitulating the in vivo HSPC phenotype (Raaijmakers et al., 2010). In vivo, blockage of TLR4 by neutralizing antibodies resulted in a reduction of γ H2AX foci in LKS cells from OCS^{fff} mice (Figure 5H).

To provide formal experimental support for the view that *S100A8/9* production by ancillary cells in the bone marrow microenvironment is sufficient to drive genotoxic stress in HSPCs in a paracrine manner, we next transplanted CD45.1⁺ wild-type hematopoietic cells into *S100A9*-GFP transgenic (*S100A9Tg*) mice, overexpressing both *S100A8* and *S100A9* under control of the MHC class I H2K promoter (Cheng et al., 2008) (Figure 6A). *S100A8/9* (GFP) was expressed in a mesenchymal (CD45⁻ CD31⁻ Ter119⁻ CD51⁺ Sca1⁻) niche population, previously shown to contain the Osterix-expressing cells (Schepers et al., 2013) (Figures 6B and 6C). The *S100A8/9*⁺ microenvironment induced accumulation of superoxide radicals (DHE) and DNA-damage (γ H2AX) in wild-type (CD45.1⁺) HSPCs (Figures 6D–6F), in particular in immunophenotypic HSCs, indicating that secretion of *S100A8/9* from ancillary cells in the microenvironment is indeed sufficient to induce genotoxic stress in HSCs in a paracrine manner.

(E) Increased γ H2AX levels in mutant HSPCs. The representative FACS analysis is shown (left). The MFI values ($n = 4$) are shown (right).

(F–I) Activation of DDR in mutant HSPCs.

(F) Transcriptional repression of G1-S checkpoint progression.

(G and H) In vivo BrdU staining confirming impaired S-phase transition ($n = 4$).

(I) Downregulation of Myc signaling. GSEA data shown are from CD48⁻ LKS cells. a.u., arbitrary units. (†††FDR < 0.001, * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$). The data in bar graphs are mean \pm SEM.

See also Figures S2, S3, and S6 and Tables S1, S2, and S3.

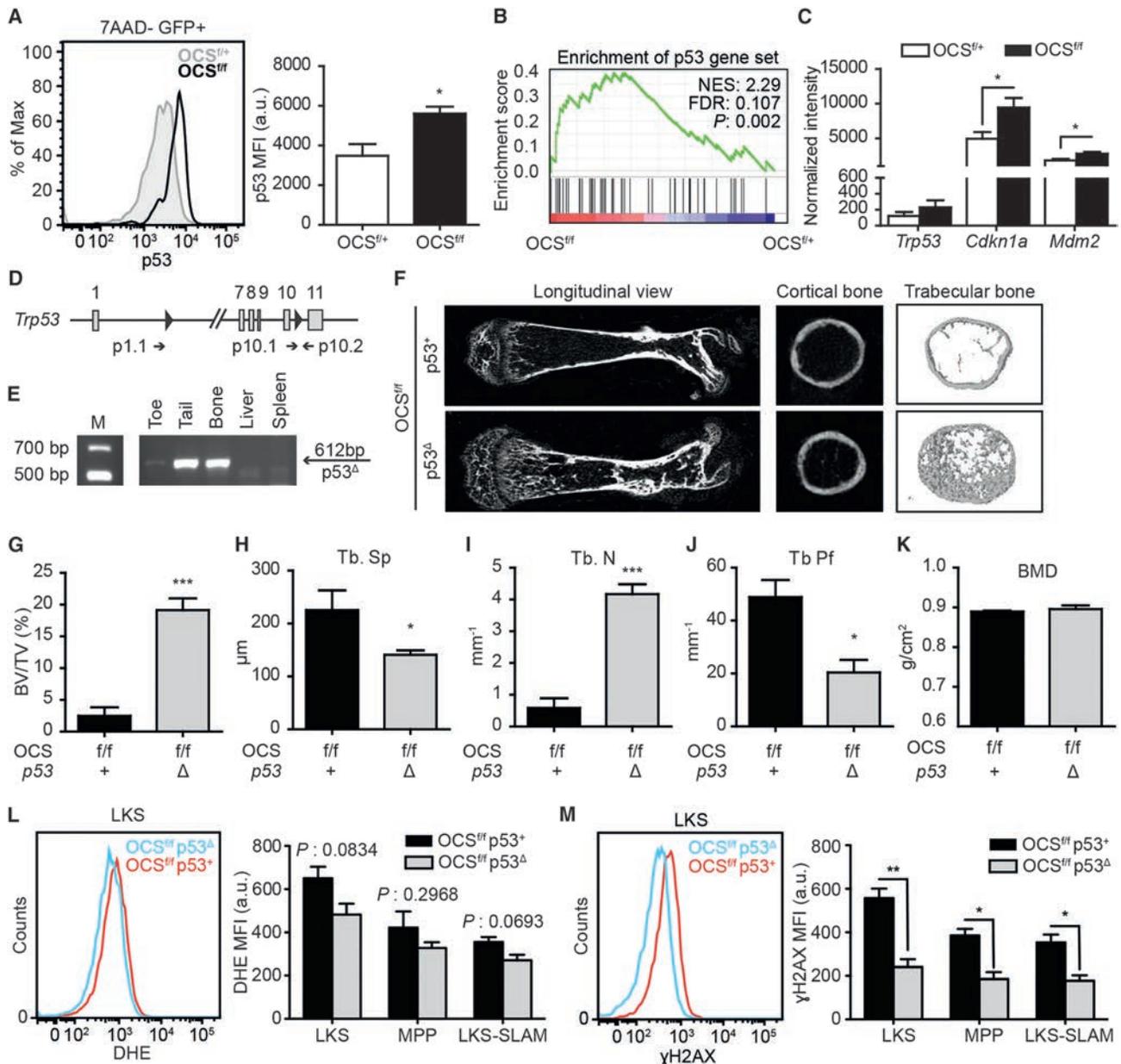


Figure 3. Activation of p53 in MPCs Drives Skeletal and Hematopoietic Abnormalities in OCS Mutants

(A) p53 protein (FACS) accumulates in GFP⁺ cells from OCS^{f/f} mice (n = 3).

(B and C) Activation of p53 in mutant GFP⁺ cells as demonstrated by enrichment of a p53 GSEA signature (B) and overexpression of canonical p53 targets (n = 3) (C).

(D) Schematic representation of the p53 floxed allele with indication of primers used to assess genotypes (p10.1–p10.2) and genomic deletion (p1.1–p10.2).

(E) Specific deletion of p53 in bone-containing tissue in OCS^{f/f} p53^Δ mice (genomic PCR).

(F–J) Non-compiled μ CT images indicating normalization of bone mass in OCS^{f/f} mice upon genetic deletion of p53 (p53⁺, n = 3 and p53^Δ, n = 5). trabecular bone pattern factor, Tb. Pf.

(K) Bone mineral density in OCS^{f/f} mice is not rescued by p53 deletion.

(L and M) Effects of p53 deficiency on OCS HSPCs. The tendency for reduced oxidative stress as assessed by DHE analysis (L) and significant reduction of γ H2AX levels (M) in HSPCs from OCS^{f/f} p53^Δ mice (n = 3). Mean fluorescence intensity, MFI and a.u. (*p < 0.05, **p < 0.01, and ***p < 0.001). The data are mean \pm SEM. See also Figure S4.

Translating these findings to human disease, exposure of human cord blood CD34⁺ HSPCs to human recombinant S100A8/9 at clinically relevant concentrations (Chen et al., 2013 and Supplemental Information) resulted in DNA damage (increased γ H2AX foci), apoptosis, and impaired HSPC function (CFU-C) (Figure S7).

Activation of the p53-S100A8/9-TLR Axis in Mesenchymal Niche Cells Predicts Leukemic Evolution and Clinical Outcome in Human Low-Risk MDS

To further define the biologic and clinical significance of these findings, we performed transcriptome sequencing of CD271⁺

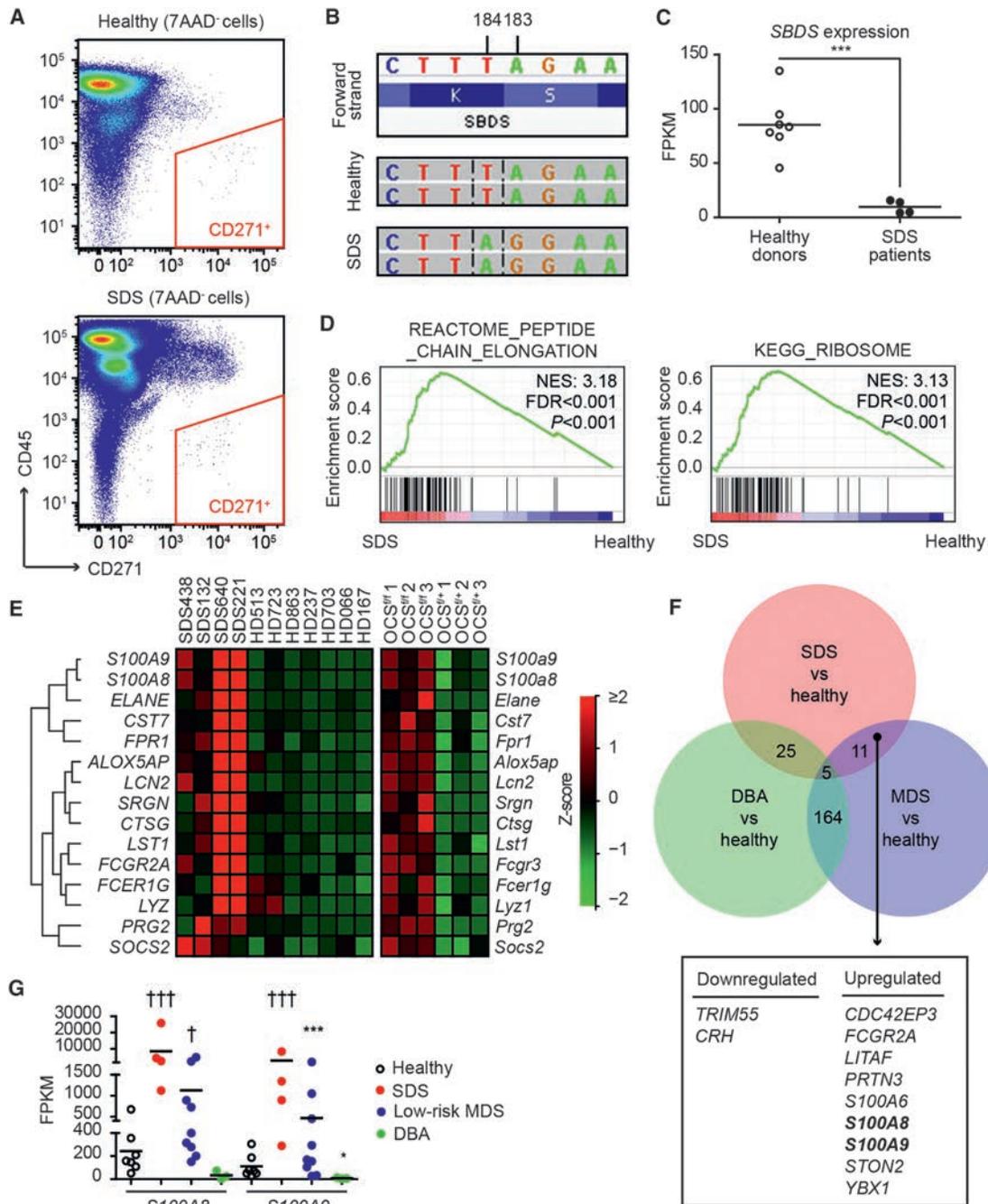


Figure 4. Identification of *S100A8* and *S100A9* as Candidate Drivers of Genotoxic Stress in Leukemia Predisposition Syndromes

(A) Representative mesenchymal CD271⁺ FACS gating.

(B) Pathognomonic 183–184 TA > CT mutation in niche cells from a representative SDS patient (IGV plot).

(C) Reduced *SBDS* expression in SDS niche cells.

(D) Disruption of ribosome biogenesis and translation in SDS CD271⁺ cells (GSEA).

(E) Inflammation-related transcripts are upregulated in niche cells from SDS patients and OCS^{fl/fl} mice.

(F) Significantly differentially expressed genes in SDS (n = 4), MDS (n = 9), and DBA (n = 3) in comparison to normal CD271⁺ cells.

(G) Expression of *S100A8* and *S100A9* in mesenchymal cells from SDS, low-risk MDS, and DBA patients (*p < 0.05, ***p < 0.001, and †††FDR-adjusted p < 0.001). See also Figures S1 and S5 and Table S4.

niche cells in a prospective, homogeneously treated cohort of low-risk MDS patients (n = 45; Figure 7A; Table S6). Expression of *S100A8* and *S100A9* was strongly correlated (Figures 7B and

7C), with a subgroup of MDS patients (17/45; 38%) demonstrating significant overexpression of *S100A8* and *S100A9* (Modified Thompson Tau outlier test) (Figures 7B and 7D),

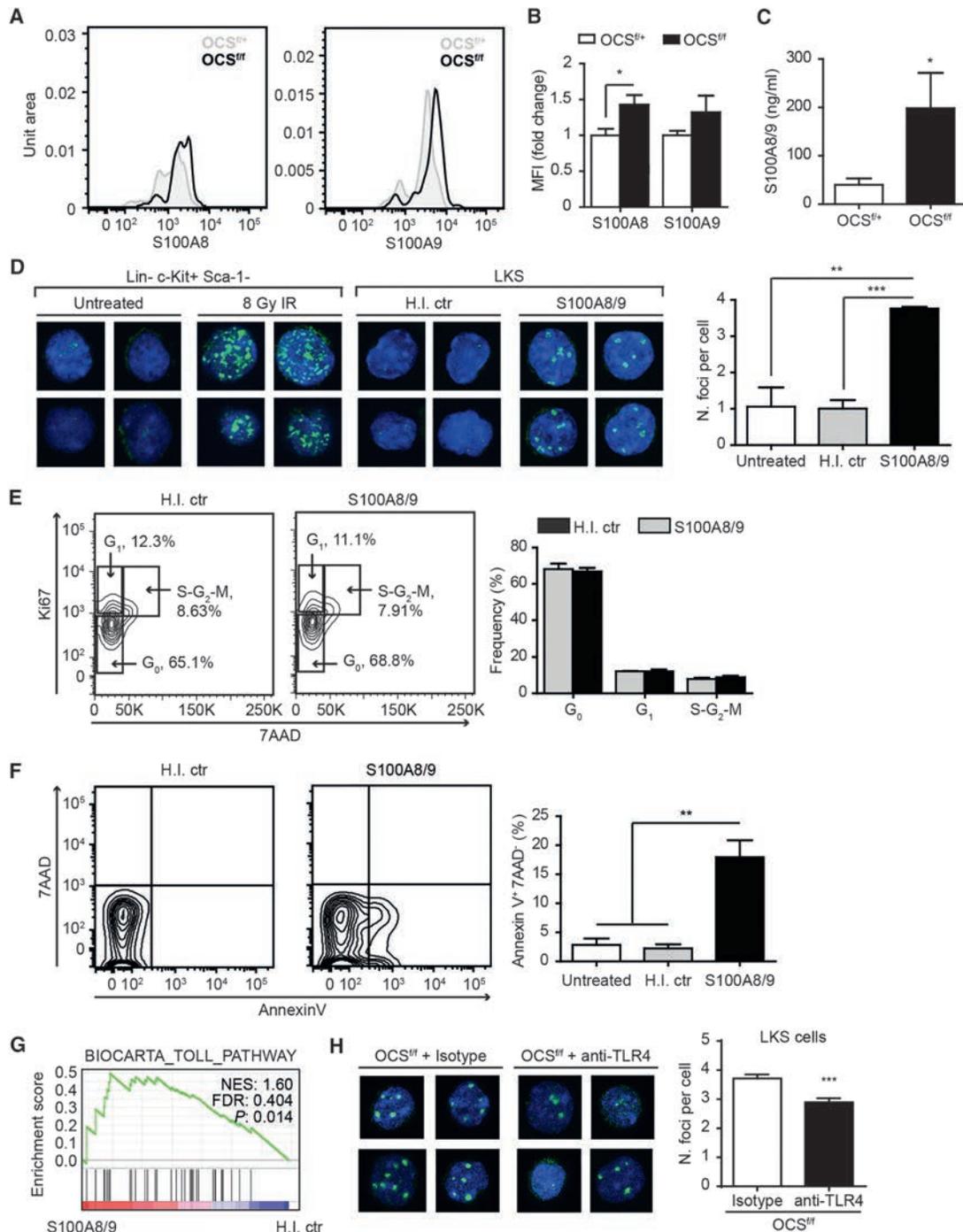


Figure 5. S100A8/9 Induces Genotoxic Stress in Murine HSPCs through TLR4 Signaling

(A and B) Increased S100A8 and S100A9 levels in OCS^{fl/fl} GFP⁺ cells.

(A) Representative plots.

(B) MFI values (n = 5).

(C) Increased plasma concentration of S100A8/9 by ELISA (OCS^{fl/+}, n = 5 and OCS^{fl/fl}, n = 4).

(D) Left: representative γ H2AX pictures after HSPCs in vitro exposure. Positive control: 8-Gy irradiated Lin⁻ c-Kit⁺ Sca-1⁻ cells and negative control: heat-inactivated S100A8/9 (H.I. ctr). The number of γ H2AX foci (n = 3) (right).

(E) S100A8/9 has no effect on cell cycle (n = 2).

(F) Increased apoptosis in S100A8/9-exposed LKS (n = 3).

(G) Activation of TLR signaling (GSEA).

(H) TLR4-blocking antibodies limit DNA damage in OCS^{fl/fl} mice (n = 4) (*p < 0.05, *p < 0.01, and ***p < 0.001). The data are mean \pm SEM.

See also Figures S6 and S7 and Table S5.

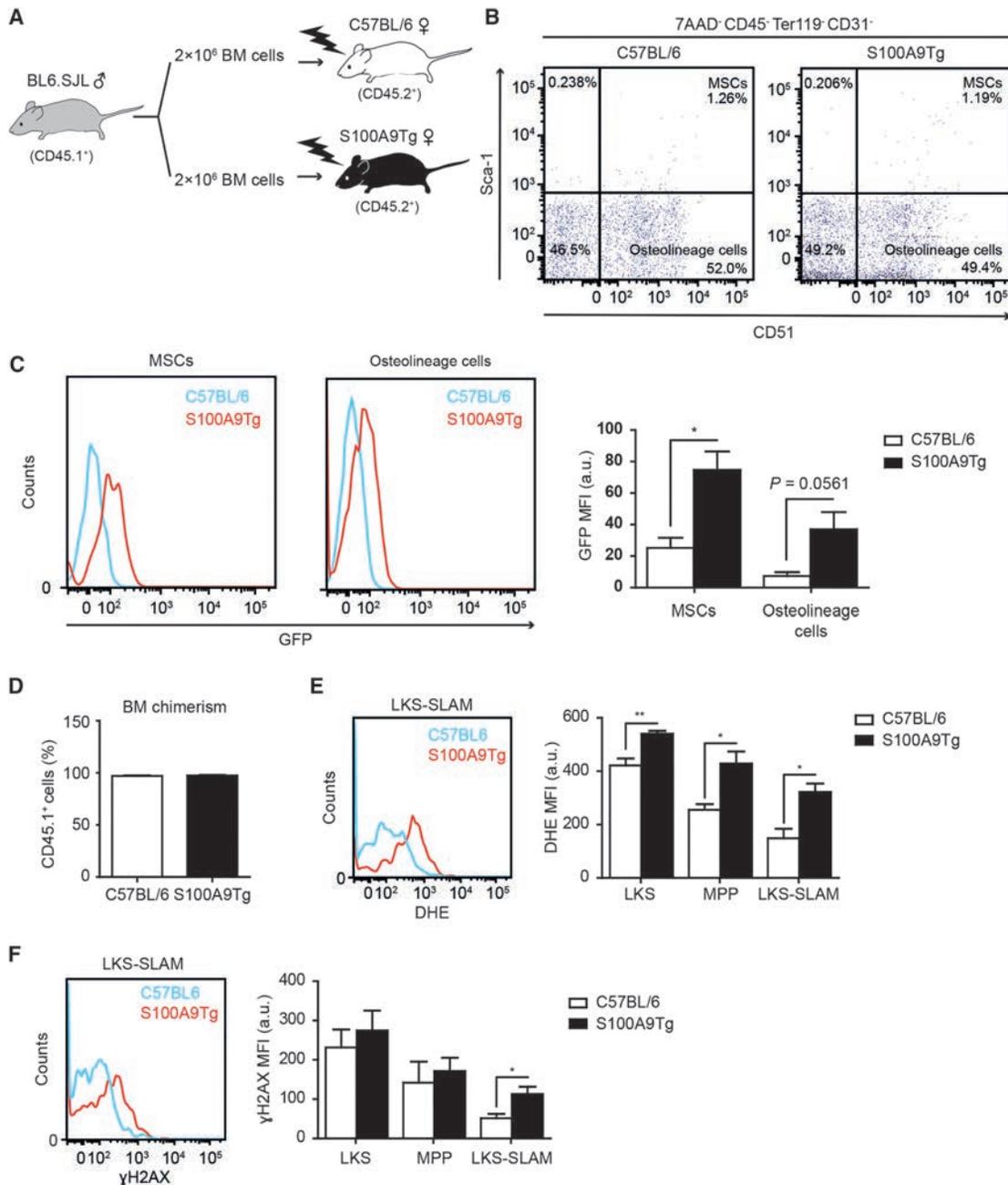


Figure 6. Niche-Derived S100A8/9 Induces Oxidative and Genotoxic Stress in HSPCs

(A) Schematic representation of wild-type HSPCs transplantation in S100A9Tg mice. Bone marrow, BM.

(B and C) Mesenchymal cells from S100A9Tg mice express the S100A9-IRES-GFP construct.

(B) Gating strategy defining CD51⁺ Sca-1⁺ mesenchymal “stem” cells (MSC) and CD51⁺ Sca-1⁻ osteolineage cells in the microenvironment.

(C) Expression of GFP in S100A9Tg-derived mesenchymal compartments (n = 3).

(D) Transplantation efficiency as assessed by CD45.1⁺ cell chimerism in the bone marrow (BM) of transplanted mice (n = 4).

(E) Accumulation of superoxide radicals in HSPCs exposed to S100A8/9-overexpressing microenvironment. The representative plot is shown (left). DHE MFI values are shown (n = 4) (right).

(F) Increased levels of γ H2AX in immunophenotypically defined HSCs. The representative plot is shown (left). The γ H2AX MFI values are shown (n = 4) (right) (*p < 0.05 and **p < 0.01). The data are mean \pm SEM.

independent of established prognostic factors as defined by the revised International Prognostic Scoring System (IPSS) and the MD Anderson risk score (LR-PSS) (Table S6). Transcriptional pathway analysis (GSEA) comparing mesenchymal cells over-

pressing S100A8/9 (n = 17) to those of niche S100A8/9⁻ patients (n = 28) revealed activation of p53 and TLR programs in S100A8/9⁺ mesenchymal cells (Figure 7E), in line with experimental data from the mouse model pointing at the existence of a

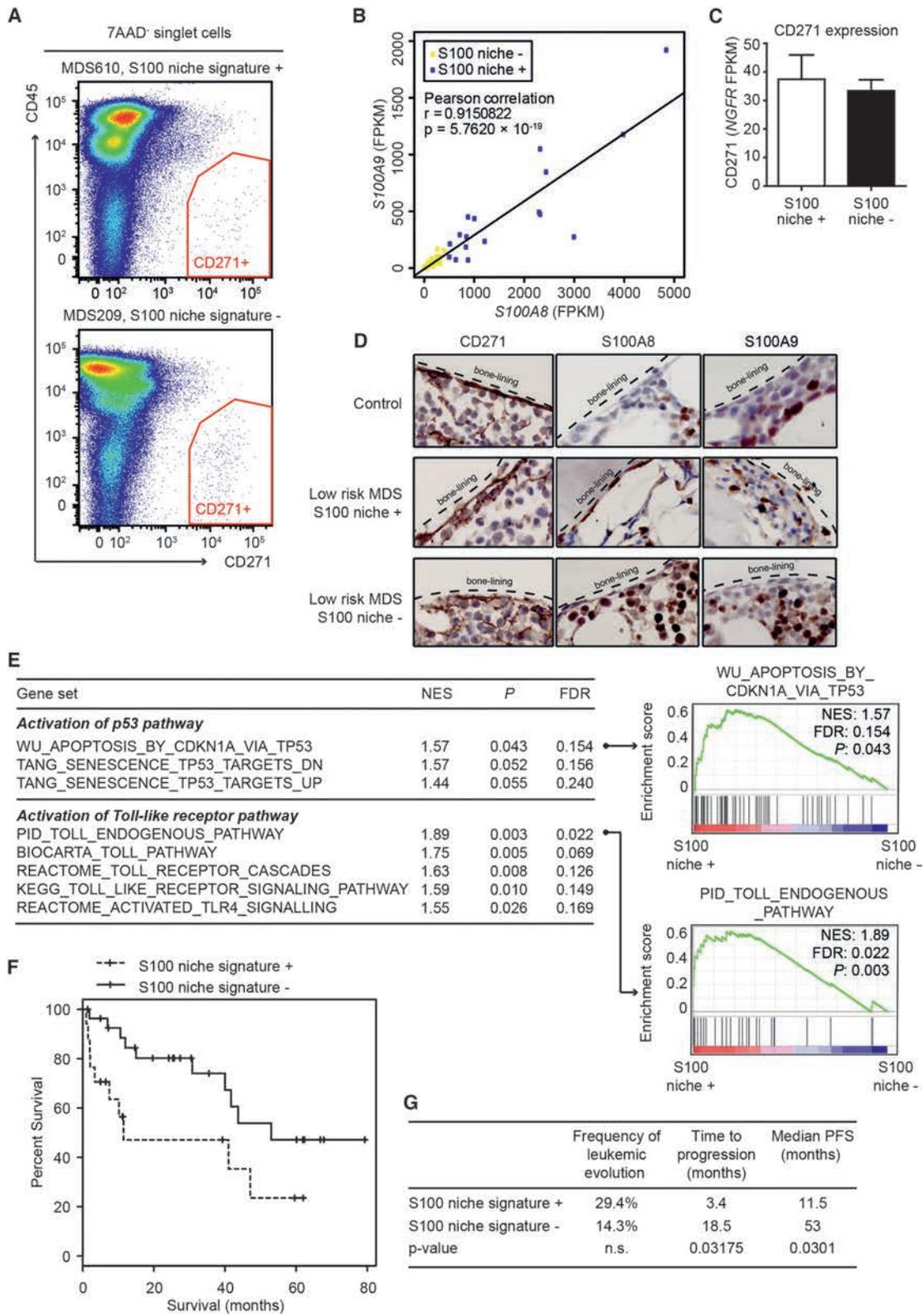


Figure 7. Activation of the p53-S100A8/9-TLR Axis in Mesenchymal Niche Cells Predicts Leukemic Evolution and Clinical Outcome in Human Low-Risk MDS

(A) Representative examples of FACS-isolated CD271⁺ niche cells in human low-risk MDS.
 (B) Correlation plot of S100A8 and S100A9 expression levels in human low-risk MDS (n = 45).

(legend continued on next page)

p53-S100A8/9-TLR axis. Leukemic evolution, defined as the development of frank AML or excess of blasts to World Health Organization (WHO) refractory anemia with excess of blasts (RAEB)1/ RAEB2, occurred in 5/17 (29.4%) of niche S100A8/9⁺ patients (three AML and two RAEB1/RAEB2) versus 4/28 (14.2%) in niche S100A8/9⁻ patients (two AML and 2 RAEB1/RAEB2). Time to leukemic evolution was significantly shorter in niche S100A8/9⁺ patients (average 3.4 [1–7.5] versus 18.5 [7–40] months; $p = 0.03$ by Exact Wilcoxon rank-sum test), resulting in a significantly shorter progression-free survival of niche S100A8/9⁺ patients (median 11.5 versus 53 months; $p = 0.03$) (Figures 7F and 7G). Collectively, the data establish activation of p53-S100A8/9 signaling in mesenchymal niche cells as an independent predictor of disease outcome in human MDS.

DISCUSSION

Genomic stress and the ensuing DNA damage play a pivotal role in the attenuation of normal hematopoiesis in aging and disease. Mutations accumulate in HSPCs over the lifespan of an organism, but the (patho)physiological sources of genomic stress in HSPCs and their relationship with human bone marrow failure remain incompletely understood. Here, we show that specific inflammatory signals from the mesenchymal niche can induce genotoxic stress in heterotypic stem/progenitor cells and relate this concept to the pathogenesis of two human bone marrow failure and leukemia predisposition syndromes, SDS and MDS.

The data indicate that the mesenchymal niche may actively contribute to the formation of a “mutagenic” environment, adding to our understanding of how a premalignant environment facilitates cancer initiation and evolution. The data argue that this may not only occur through facilitated selection and expansion of genetic clones that stochastically emerge in a permissive environment, but that the mesenchymal niche may be an active participant in driving the genotoxic stress underlying tissue failure and malignant transformation of parenchymal cells.

Notably, leukemic transformation was not observed in mice with targeted deficiency of *Sbds* in mesenchymal cells. Earlier, in a related mouse model of targeted *Dicer1* deletion in MPCs, leukemic transformation was a rare event (Raaijmakers et al., 2010). In the light of our current findings, these observations are likely explained by several factors. First, prolonged exposure to a mutagenic niche, beyond the limited lifespan of OCS mice, may be necessary for the accumulation of genetic damage required for full transformation. Additionally, the data argue that DNA repair-proficient HSPCs are able to cope with the mutagenic stress induced by their environment through activation of the DDR (as shown by molecular activation of cell-cycle checkpoints and apoptosis), preventing the accumulation of stable genetic damage (as demonstrated by comet assays

and maintaining the functional integrity of HSPCs (as shown by repopulation assays).

We propose that in SDS (and possibly other congenital bone marrow failure syndromes), genetically aberrant hematopoietic and niche elements cooperate in driving bone marrow failure and leukemic evolution. Our mouse models of SDS support a view in which hematopoietic cell autonomous loss of function of *Sbds* drives neutropenia (Zambetti et al., 2015), while niche alterations in this disease drive myelodysplastic alteration and genotoxic stress. It is conceivable that loss-of-function mutations in *Sbds* in HSPCs further sensitize HSPCs to the genotoxic effects of the *Sbds*-deficient environment, perhaps through attenuation of DNA damage repair mechanisms. It will thus be of considerable interest to test the hypothesis that a mutagenic environment cooperates with aberrant HSPCs, compromised in their ability to cope with inflammatory genotoxic stress, in leukemia evolution. In this context, the propensity of *Sbds*-deficient cells to accumulate ROS (Ambekar et al., 2010), and their reduced ability to cope with various cellular stressors such as mitotic spindle destabilizing agents, ER stress activators, topoisomerase inhibitors, and UV irradiation (Austin et al., 2008; Ball et al., 2009), is noteworthy.

The current findings add to emerging insights into the role of innate immune TLR-signaling in the pathogenesis of human MDS. TLR4 and other TLRs are overexpressed in HSPCs from MDS patients (Maratheftis et al., 2007; Wei et al., 2013), and TLR4 expression was shown to correlate with apoptosis in CD34⁺ hematopoietic cells. TLR signaling is constitutively activated in MDS mice with deletion of chromosome 5 (del5q) (Starczynowski et al., 2010), and multiple TLR downstream signaling pathways have been shown to be activated in MDS and related to loss of progenitor cell function (Gañán-Gómez et al., 2015).

Our findings implicate the DAMP S100A8/9 derived from the mesenchymal niche as a driver of TLR signaling in this disease. The unbiased identification of S100A8/9 seems to independently converge with an earlier report implicating S100A8/9 in the pathogenesis of MDS (Chen et al., 2013). In this study, it was shown that the plasma concentration of S100A9 was significantly increased in MDS patients (Chen et al., 2013), and S100A8/9 was shown to drive expansion and activation of myeloid-derived suppressor cells (MDSCs) that contributed to cytopenia and myelodysplasia in a murine model of S100A9 overexpression through secretion of suppressive cytokines. It is therefore an intriguing possibility that additional, indirect, biologic effects of S100A8/9 contribute to the hematopoietic phenotype of OCS mice. This may include engagement of other cognate receptors of the protein, including expansion of MDSCs through CD33 signaling (Chen et al., 2013).

In our study, S100A8/9 was aberrantly overexpressed in a rare population of mesenchymal niche cells, both in the mouse model

(C) Expression of the defining gene *NGFR* (CD271) in mesenchymal cells (S100 niche +, $n = 17$ and S100 niche -, $n = 28$).

(D) Representative staining of S100A8 and S100A9 in endosteal (CD271⁺) stromal cells. The intramedullary staining reflects expression of S100A8/9 in myeloid cells.

(E) GSEA analysis indicating enrichment of p53 and TLR signatures in S100A8/9-overexpressing CD271⁺ cells. Two representative GSEA plots are shown. Normalized enrichment score, NES.

(F) Kaplan-Meier survival curve showing progression-free survival.

(G) Statistical analysis indicating significantly reduced time to progression and progression-free survival (PFS). The data are mean \pm SEM.

See also Table S6.

and human disease. Typically, expression of the protein is found in myeloid cells, raising the question why S100A8/9 production by (rare) niche cells is more relevant to the biology of HSPCs than secretion from myeloid/erythroid cells. While the answer to this question remains speculative in the absence of *in vivo* targeted overexpression studies, it is noteworthy that, in contrast to most cytokines, chemokines, and other proinflammatory molecules, the local accumulation of S100A8/9 in the environment is very high (up to 100 $\mu\text{g}/\text{mL}$ and about 50- to 100-fold higher than systemic concentrations), likely caused by attachment to extracellular matrices such as proteoglycans (Vogl et al., 2014). This implicates that the exposure of HSPCs to S100A8/9 is projected to relate strongly to their anatomical proximity to a producing cell. CD271⁺ mesenchymal cells are directly adjacent to CD34⁺ HSPCs in human bone marrow (Flores-Figueroa et al., 2012). This notion of “spatial relevance” may also be congruent with recent observations that aberrant overexpression of S100A8/9 in hematopoietic (erythroid) cells within the erythroid island in a model of human 5q⁻ syndrome leads to a predominant erythroid, anemic, phenotype (Schneider et al., 2016).

The mechanisms of S100A8/9 induced DNA damage remain to be fully elucidated. Our experiments using NAC to reduce ROS burden suggest an incomplete association between oxidative stress and DNA damage, suggesting that S100A8/9 secretion may attenuate genomic integrity through additional mechanisms. Similarly, it is conceivable that other ligands secreted from mesenchymal cells contribute to the induction of DNA damage in HSPCs in the mouse model. We found a striking abundance of transcripts encoding other DAMPs and cytotoxic proteins in both the mouse model and mesenchymal elements isolated from SDS patients. Ongoing investigations will have to assess whether other selected ligands can evoke genomic stress in heterotypic HSPCs and in such a fashion contribute to the generation of a mutagenic environment in these disorders.

Finally, our findings establish molecular characteristics of the mesenchymal environment as an important determinant of disease outcome in humans. S100A8/9 expression, associated with activated p53 and TLR signaling, in mesenchymal cells predicted leukemic evolution and progression-free survival in a cohort of homogeneously treated low-risk MDS patients. This is of considerable clinical relevance because low-risk MDS is a heterogeneous disease-entity, with a subset of patients having a particular dismal prognosis not identified by current risk-stratification strategies (Bejar et al., 2012). Gene expression of S100A8/9 may identify a substantial subset of patients with a survival typically associated with “high-risk” patients and, if confirmed in larger independent cohorts, could guide therapeutic decision making in MDS. The data thus provide a strong rationale for niche-instructed therapeutic targeting of inflammatory signaling in human pre-leukemic disease.

EXPERIMENTAL PROCEDURES

Mice and In Vivo Procedures

OCS, *Trp53^{fl/fl}*, and S100A9Tg mice have been previously described (Cheng et al., 2008; Jonkers et al., 2001; Raaijmakers et al., 2010). *Ptprc⁹Pepc⁹/BoyCr1* (B6.SJL) mice were purchased from Charles River. Animals were maintained in specific pathogen free conditions in the Experimental Animal Center of Erasmus MC (EDC). For *in vivo* cell-cycle analysis, OCS mice received intraperitoneal injections of BrdU (1.5 mg in PBS, BD Biosciences) and sacrificed after

15 hr. For TLR4 studies, 2-week-old mice were intraperitoneally injected with a double dose (100 μg and 35 μg , 48 hr interval) of TLR4-neutralizing antibody (clone MTS510, eBioscience) or isotype control (clone eBR2a, eBioscience) and sacrificed after 60 hr. For NAC rescue studies, 2-week-old mice received daily intraperitoneal injections of NAC (320 mg/kg in saline, Sigma-Aldrich) until the day of the analysis and at least for 5 days. All mice were sacrificed by cervical dislocation. Animal studies were approved by the Animal Welfare/Ethics Committee of the EDC in accordance with legislation in the Netherlands (approval No. EMC 2067, 2714, 2892, 3062).

μCT Analysis

Femur bones were isolated, fixated in 3% PFA/PBS for 24 hr, and stored in 70% ethanol. μCT analysis was performed using a SkyScan 1172 system (Sky-Scan) using previously described settings (Tudpor et al., 2015). Bone micro-architectural parameters relative to the trabecular and the cortical area were determined in the distal metaphysis and the middiaphysis of each femur, respectively, using software packages from Bruker MicroCT (NRecon, CTAn, and DataViewer).

Human Bone Marrow Samples

Bone marrow aspirates were obtained from SDS and DBA patients during routine follow up. All MDS patients were treated with lenalidomide (10 mg/day, d 1–21 in a 4-week schedule) in the context of an ongoing prospective clinical trial for patients with low or intermediate-1 risk MDS according to IPSS criteria (HOVON89; <http://www.hovon.nl>; <http://www.trialregister.nl> as NTR1825; EudraCT No. 2008-002195-10). Bone marrow specimens were collected at study entry and disease diagnosis and staging confirmed by central board reviewing. Leukemic evolution was assessed according to WHO criteria; development of RAEB1 or RAEB-2 (if RAEB1 at entry) was considered progression of disease. Leukemia (AML) was diagnosed according to standard WHO criteria ($\geq 20\%$ myeloblasts in blood/bone marrow). Bone marrow cells from allogeneic transplantation donors were used as normal controls. Patients and healthy donor characteristics are described in Tables S4 and S6. All specimens were collected with informed consent, in accordance with the Declaration of Helsinki.

Gene Expression Profiling

Osx::GFP cells from bone cell suspensions of OCS mice were sorted in TRIzol Reagent (Life Technologies) and RNA was extracted according to the manufacturer's recommendations. Linear amplification of mRNA was performed using the Ovation Pico WTA System (NuGEN). cDNA was fragmented and labeled with Encore Biotin Module (NuGEN). The biotinylated cDNA was hybridized to the GeneChip Mouse Genome 430 2.0 Array (Affymetrix eBioscience). Signal was normalized and differential gene expression analysis was performed with the limma package (Ritchie et al., 2015). RNA sequencing experiments were performed as previously described (Zambetti et al., 2015). Human transcripts were aligned to the RefSeq transcriptome (hg19) and analyzed with DESeq2 (Love et al., 2014), while mouse transcripts were aligned to the Ensembl transcriptome (mm10) and analyzed with EdgeR (Robinson et al., 2010) in the R environment. Fragments per kilobase of transcript per million mapped reads (FPKM) values were calculated using Cufflinks (Trapnell et al., 2010). Principal component analysis was performed in the R environment on the raw fragment counts extracted from the BAM files by HTSeq-count (Anders et al., 2015). For gene set enrichment analysis (Subramanian et al., 2005) (GSEA, Broad Institute), normalized intensity values (microarray data) and FPKM values (RNA-seq) were compared to the curated gene sets (C2) and the gene ontology gene sets (C5) of the Molecular Signature Database (MsigDB) using the Signal2Noise metric and 1,000 gene set-based permutations. For HSPCs gene ontology-term analysis, genes with significantly differential expression ($p < 0.05$) were interrogated using g:Profiler web-based software (Reimand et al., 2007, 2011).

Immunofluorescence Microscopy

HSPCs were harvested in PBS + 0.5% FBS, cytospun on a glass slide for 3 min at 500 rpm using a Cytospin 4 centrifuge (Thermo Scientific), and fixed in 3% PFA/PBS for 15 min on ice. After three washing steps in PBS, cells were permeabilized for 2 min in 0.15% Triton X-100/PBS. Aspecific binding sites were blocked by incubation in 1% BSA/PBS for 1 hr at room temperature. Cells were

next stained overnight at 4°C with either anti-phospho-histone H2A.X (Ser139) mouse monoclonal antibody (clone JBW301, Merck Millipore, diluted 1:1,000 in 1% BSA/PBS) or with anti-53BP1 rabbit polyclonal antibody (Novus Biologicals, diluted 1:1,000 in 1% BSA/PBS). Slides were washed twice in PBS for 5 min and incubated for 1 hr at 37°C with either Alexa Fluor 488-conjugated goat anti-mouse antibody (Cat. A10667, Life Technologies) or goat anti-rabbit antibody (Cat. A11008, Life Technologies), both diluted 1:200 in 1% BSA/PBS. After two washes in PBS, slides were mounted in VECTASHIELD Mounting Medium with DAPI (Vector Laboratories). z series images were acquired with a Leica TCS SP5 confocal microscope (63× objective lens) using the LAS software (Leica Microsystems). γ H2AX and 53BP1 foci were counted manually from the maximum projection view.

Survival Analysis

The low-risk MDS patient subgroup with S100 niche signature was defined by the Modified Thompson Tau test for outlier detection. In brief, S100A8 statistics from the control cases were combined to define the rejection region, demarcating FPKM values to be considered as outliers. MDS cases with S100A8 FPKM values within the rejection region were thus defined as niche-signature*. To determine the significance difference in time to progression, we used the Wilcoxon signed-rank test accounting for tied observations. Event-free survival was determined by specifying leukemic progression or death as events. Patients experiencing a non-hematological related death (e.g., cardiac failure), were censored on the date of this event. Patients remaining alive were censored on the date of last consultation. Kaplan-Meier curves were used to estimate the survival functions through time. Statistical differences in the survival distributions were assessed with the Mantel-Cox log-rank test. All calculations were performed in the R environment.

Statistics

Statistical analysis was performed using Prism 5 (GraphPad Software). Unless otherwise specified, unpaired, two-tailed Student's t test (single test) or one-way ANOVA (multiple comparisons) were used to evaluate statistical significance, defined as $p < 0.05$. All results in bar graphs are mean value \pm SEM.

ACCESSION NUMBERS

The accession number for the gene expression array data derived from murine Osterix::GFP cells is ArrayExpress: E-MTAB-5023. The accession number for the RNA-seq data derived from murine CD48⁻/CD48⁺ LSK cells is European Nucleotide Archive, which is hosted by the EBI: PRJEB15060. The accession number for the RNA-seq data derived from human LR-MDS, SDS, and DBA specimens is European Genome-phenome Archive, which is hosted by the EBI: EGAS00001001926.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and six tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.08.021>.

AUTHOR CONTRIBUTIONS

Conceptualization: N.A.Z. and M.H.G.P.R.; Methodology: N.A.Z., Z.P., S.C., K.J.G.K., M.A.M., E.M.J.B., B.C.J.v.d.E., J.P.T.M.v.L., R.K., T.V., and M.H.G.P.R.; Investigation: N.A.Z., Z.P., S.C., M.A.M., E.M.J.B., M.N.A., P.M.H.V.S., C.S.v.d.L., B.C.J.v.d.E., and T.V.; Resources: T.M.W., E.M.P.C., C.M., P.G.M., J.P.T.M.v.L., I.P.T., T.W.K., R.K., A.A.v.d.L., and T.V.; Data Curation: M.A.S., R.M.H., T.M.W., E.M.P.C., T.W.K., and A.A.v.d.L.; Writing: N.A.Z. and M.H.G.P.R.; Visualization: N.A.Z., Z.P., S.C., and R.M.H.; and Supervision and Funding Acquisition: M.H.G.P.R.

ACKNOWLEDGMENTS

Professor Johanna M. Rommens donated *Sbds*-floxed mice; Dr. Eric Braakman and Mariette ter Borg provided CD34⁺ cells; Dr. Elwin Rombouts, Onno Roovers, Peter van Geel, Kirsten van Lom, Marijke Koedam, Nicole van Vliet,

Charlie Laffeber, and Gert-Jan Kremers provided technical assistance; Dr. Marc Bierings and Dr. Valerie de Haas on behalf of the "Stichting Kinderoncologie Nederland (SKION)" provided DBA samples; Pearl F.M. Mau Asam helped with SDS bone marrow collection; members of the myeloid working party of the HOVON collected MDS bone marrow samples; Dr. Dana Chitu of the HOVON data center helped with clinical data; members of the Erasmus MC Department of Hematology provided scientific discussion; and members of the Erasmus MC animal core facility EDC helped with animal care. This work was supported by grants from the Dutch Cancer Society (KWF Kankerbestrijding), Amsterdam, the Netherlands (grant EMCR 2010-4733 to M.H.G.P.R.), the ErasmusMC (Fellowship to M.H.G.P.R.), the Netherlands Organization of Scientific Research (NWO 90700422 to M.H.G.P.R.), and the Netherlands Genomics Initiative (Zenith grant no 40-41009-98-11062 to M.H.G.P.R.).

Received: November 23, 2015

Revised: July 6, 2016

Accepted: August 22, 2016

Published: September 22, 2016

REFERENCES

- Aggett, P.J., Cavanagh, N.P.C., Matthew, D.J., Pincott, J.R., Sutcliffe, J., and Harries, J.T. (1980). Shwachman's syndrome. A review of 21 cases. *Arch. Dis. Child.* **55**, 331–347.
- Alter, B.P. (2007). Diagnosis, genetics, and management of inherited bone marrow failure syndromes. *Hematology (Am. Soc. Hematol. Educ. Program)* **1**, 29–39.
- Ambekar, C., Das, B., Yeger, H., and Dror, Y. (2010). SBDS-deficiency results in deregulation of reactive oxygen species leading to increased cell death and decreased cell growth. *Pediatr. Blood Cancer* **55**, 1138–1144.
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169.
- Arranz, L., Sánchez-Aguilera, A., Martín-Pérez, D., Isern, J., Langa, X., Tzankov, A., Lundberg, P., Muntión, S., Tzeng, Y.S., Lai, D.M., et al. (2014). Neuropathy of haematopoietic stem cell niche is essential for myeloproliferative neoplasms. *Nature* **512**, 78–81.
- Austin, K.M., Gupta, M.L., Jr., Coats, S.A., Tulpule, A., Mostoslavsky, G., Balazs, A.B., Mulligan, R.C., Daley, G., Pellman, D., and Shimamura, A. (2008). Mitotic spindle destabilization and genomic instability in Shwachman-Diamond syndrome. *J. Clin. Invest.* **118**, 1511–1518.
- Ball, H.L., Zhang, B., Riches, J.J., Gandhi, R., Li, J., Rommens, J.M., and Myers, J.S. (2009). Shwachman-Bodian Diamond syndrome is a multi-functional protein implicated in cellular stress responses. *Hum. Mol. Genet.* **18**, 3684–3695.
- Bejar, R., Stevenson, K.E., Caughey, B.A., Abdel-Wahab, O., Steensma, D.P., Gallli, N., Raza, A., Kantarjian, H., Levine, R.L., Neuberger, D., et al. (2012). Validation of a prognostic model and the impact of mutations in patients with lower-risk myelodysplastic syndromes. *J. Clin. Oncol.* **30**, 3376–3382.
- Boocock, G.R., Morrison, J.A., Popovic, M., Richards, N., Ellis, L., Durie, P.R., and Rommens, J.M. (2003). Mutations in SBDS are associated with Shwachman-Diamond syndrome. *Nat. Genet.* **33**, 97–101.
- Chen, X., Eksioğlu, E.A., Zhou, J., Zhang, L., Djeu, J., Fortenberry, N., Epling-Burnette, P., Van Bijnen, S., Dolstra, H., Cannon, J., et al. (2013). Induction of myelodysplasia by myeloid-derived suppressor cells. *J. Clin. Invest.* **123**, 4595–4611.
- Chen, S., Zambetti, N.A., Bindels, E.M., Kenswill, K., Mylona, A.M., Adisty, N.M., Hoogenboezem, R.M., Sanders, M.A., Cremers, E.M., Westers, T.M., et al. (2016). Massive parallel RNA sequencing of highly purified mesenchymal elements in low-risk MDS reveals tissue-context-dependent activation of inflammatory programs. *Leukemia*, Published online June 3, 2016. <http://dx.doi.org/10.1038/leu.2016.91>.
- Cheng, P., Corzo, C.A., Luetteke, N., Yu, B., Nagaraj, S., Bui, M.M., Ortiz, M., Nacken, W., Sorg, C., Vogl, T., et al. (2008). Inhibition of dendritic cell differentiation and accumulation of myeloid-derived suppressor cells in cancer is regulated by S100A9 protein. *J. Exp. Med.* **205**, 2235–2249.

- Curtin, N.J. (2012). DNA repair dysregulation from cancer driver to therapeutic target. *Nat. Rev. Cancer* *12*, 801–817.
- Donadieu, J., Fenneteau, O., Beaupain, B., Beauvils, S., Bellanger, F., Mahlaoui, N., Lambilliotte, A., Aladjidi, N., Bertrand, Y., Mialou, V., et al.; Associated investigators of the French Severe Chronic Neutropenia Registry* (2012). Classification of and risk factors for hematologic complications in a French national cohort of 102 patients with Shwachman-Diamond syndrome. *Haematologica* *97*, 1312–1319.
- Finch, A.J., Hilcenko, C., Basse, N., Drynan, L.F., Goyenechea, B., Menne, T.F., González Fernández, A., Simpson, P., D'Santos, C.S., Arends, M.J., et al. (2011). Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. *Genes Dev.* *25*, 917–929.
- Flores-Figueroa, E., Varma, S., Montgomery, K., Greenberg, P.L., and Gratzinger, D. (2012). Distinctive contact between CD34+ hematopoietic progenitors and CXCL12+ CD271+ mesenchymal stromal cells in benign and myelodysplastic bone marrow. *Lab. Invest.* *92*, 1330–1341.
- Gañán-Gómez, I., Wei, Y., Starczynowski, D.T., Colla, S., Yang, H., Cabrero-Calvo, M., Bohannon, Z.S., Verma, A., Steidl, U., and Garcia-Manero, G. (2015). Deregulation of innate immune and inflammatory signaling in myelodysplastic syndromes. *Leukemia* *29*, 1458–1469.
- Ginzberg, H., Shin, J., Ellis, L., Morrison, J., Ip, W., Dror, Y., Freedman, M., Heitlinger, L.A., Belt, M.A., Corey, M., et al. (1999). Shwachman syndrome: phenotypic manifestations of sibling sets and isolated cases in a large patient cohort are similar. *J. Pediatr.* *135*, 81–88.
- Hanoun, M., Zhang, D., Mizoguchi, T., Pinho, S., Pierce, H., Kunisaki, Y., Lacombe, J., Armstrong, S.A., Dührsen, U., and Frenette, P.S. (2014). Acute myelogenous leukemia-induced sympathetic neuropathy promotes malignancy in an altered hematopoietic stem cell niche. *Cell Stem Cell* *15*, 365–375.
- Head, D.R., Jacobberger, J.W., Mosse, C., Jagasia, M., Dupont, W., Goodman, S., Flye, L., Shinar, A., McClintock-Treep, S., Stelzer, G., et al. (2011). Innovative analyses support a role for DNA damage and an aberrant cell cycle in myelodysplastic syndrome pathogenesis. *Bone Marrow Res.* *2011*, 950934.
- Ito, K., Hirao, A., Arai, F., Takubo, K., Matsuoka, S., Miyamoto, K., Ohmura, M., Naka, K., Hosokawa, K., Ikeda, Y., and Suda, T. (2006). Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nat. Med.* *12*, 446–451.
- Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burt, N., Chavez, A., et al. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* *371*, 2488–2498.
- Jonkers, J., Meuwissen, R., van der Gulden, H., Peterse, H., van der Valk, M., and Berns, A. (2001). Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat. Genet.* *29*, 418–425.
- Kode, A., Manavalan, J.S., Mosialou, I., Bhagat, G., Rathinam, C.V., Luo, N., Khiabani, H., Lee, A., Murty, V.V., Friedman, R., et al. (2014). Leukaemogenesis induced by an activating β -catenin mutation in osteoblasts. *Nature* *506*, 240–244.
- Li, C., Chen, H., Ding, F., Zhang, Y., Luo, A., Wang, M., and Liu, Z. (2009). A novel p53 target gene, S100A9, induces p53-dependent cellular apoptosis and mediates the p53 apoptosis pathway. *Biochem. J.* *422*, 363–372.
- Li, L., Li, M., Sun, C., Francisco, L., Chakraborty, S., Sabado, M., McDonald, T., Gyorfy, J., Chang, K., Wang, S., et al. (2011). Altered hematopoietic cell gene expression precedes development of therapy-related myelodysplasia/acute myeloid leukemia and identifies patients at risk. *Cancer Cell* *20*, 591–605.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* *15*, 550.
- Mäkitie, O., Ellis, L., Durie, P.R., Morrison, J.A., Sochett, E.B., Rommens, J.M., and Cole, W.G. (2004). Skeletal phenotype in patients with Shwachman-Diamond syndrome and mutations in SBDS. *Clin. Genet.* *65*, 101–112.
- Maratheftis, C.I., Andreaskos, E., Moutsopoulos, H.M., and Voulgarelis, M. (2007). Toll-like receptor-4 is up-regulated in hematopoietic progenitor cells and contributes to increased apoptosis in myelodysplastic syndromes. *Clin. Cancer Res.* *13*, 1154–1160.
- Marino, S., Vooijs, M., van Der Gulden, H., Jonkers, J., and Berns, A. (2000). Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum. *Genes Dev.* *14*, 994–1004.
- Medyouf, H., Mossner, M., Jann, J.C., Nolte, F., Raffel, S., Herrmann, C., Lier, A., Eisen, C., Nowak, V., Zens, B., et al. (2014). Myelodysplastic cells in patients reprogram mesenchymal stromal cells to establish a transplantable stem cell niche disease unit. *Cell Stem Cell* *14*, 824–837.
- Murphy, M.P. (2009). How mitochondria produce reactive oxygen species. *Biochem. J.* *417*, 1–13.
- Owusu-Ansah, E., Yavari, A., Mandal, S., and Banerjee, U. (2008). Distinct mitochondrial retrograde signals control the G1-S cell cycle checkpoint. *Nat. Genet.* *40*, 356–361.
- Peddie, C.M., Wolf, C.R., McLellan, L.I., Collins, A.R., and Bowen, D.T. (1997). Oxidative DNA damage in CD34+ myelodysplastic cells is associated with intracellular redox changes and elevated plasma tumour necrosis factor- α concentration. *Br. J. Haematol.* *99*, 625–631.
- Raaijmakers, M.H., Mukherjee, S., Guo, S., Zhang, S., Kobayashi, T., Schoonmaker, J.A., Ebert, B.L., Al-Shahrour, F., Hasserjian, R.P., Scadden, E.O., et al. (2010). Bone progenitor dysfunction induces myelodysplasia and secondary leukaemia. *Nature* *464*, 852–857.
- Raiser, D.M., Narla, A., and Ebert, B.L. (2014). The emerging importance of ribosomal dysfunction in the pathogenesis of hematologic disorders. *Leuk. Lymphoma* *55*, 491–500.
- Rawls, A.S., Gregory, A.D., Woloszynek, J.R., Liu, F., and Link, D.C. (2007). Lentiviral-mediated RNAi inhibition of Sbds in murine hematopoietic progenitors impairs their hematopoietic potential. *Blood* *110*, 2414–2422.
- Reimand, J., Kull, M., Peterson, H., Hansen, J., and Vilo, J. (2007). g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Res.* *35*, W193–W200.
- Reimand, J., Arak, T., and Vilo, J. (2011). g:Profiler—a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Res.* *39*, W307–W315.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* *43*, e47.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* *26*, 139–140.
- Rossi, D.J., Bryder, D., Seita, J., Nussenzweig, A., Hoeijmakers, J., and Weissman, I.L. (2007). Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* *447*, 725–729.
- Sahin, E., and Depinho, R.A. (2010). Linking functional decline of telomeres, mitochondria and stem cells during ageing. *Nature* *464*, 520–528.
- Schepers, K., Pietras, E.M., Reynaud, D., Flach, J., Binnewies, M., Garg, T., Wagers, A.J., Hsiao, E.C., and Passegué, E. (2013). Myeloproliferative neoplasia remodels the endosteal bone marrow niche into a self-reinforcing leukemic niche. *Cell Stem Cell* *13*, 285–299.
- Schepers, K., Campbell, T.B., and Passegué, E. (2015). Normal and leukemic stem cell niches: insights and therapeutic opportunities. *Cell Stem Cell* *16*, 254–267.
- Schneider, R.K., Schenone, M., Ferreira, M.V., Kramann, R., Joyce, C.E., Martanyi, C., Beier, F., Brümmendorf, T.H., Germing, U., Platzbecker, U., et al. (2016). Rps14 haploinsufficiency causes a block in erythroid differentiation mediated by S100A8 and S100A9. *Nat. Med.* *22*, 288–297.
- Srikrishna, G., and Freeze, H.H. (2009). Endogenous damage-associated molecular pattern molecules at the crossroads of inflammation and cancer. *Neoplasia* *11*, 615–628.
- Starczynowski, D.T., Kuchenbauer, F., Argiropoulos, B., Sung, S., Morin, R., Muranyi, A., Hirst, M., Hogge, D., Marra, M., Wells, R.A., et al. (2010). Identification of miR-145 and miR-146a as mediators of the 5q- syndrome phenotype. *Nat. Med.* *16*, 49–58.

- Stowe, D.F., and Camara, A.K.S. (2009). Mitochondrial reactive oxygen species production in excitable cells: modulators of mitochondrial and cell function. *Antioxid. Redox Signal.* *11*, 1373–1414.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* *102*, 15545–15550.
- Toiviainen-Salo, S., Mäyränpää, M.K., Durie, P.R., Richards, N., Grynpas, M., Ellis, L., Ikegawa, S., Cole, W.G., Rommens, J., Marttinen, E., et al. (2007). Shwachman-Diamond syndrome is associated with low-turnover osteoporosis. *Bone* *41*, 965–972.
- Tormin, A., Li, O., Brune, J.C., Walsh, S., Schütz, B., Ehinger, M., Ditzel, N., Kassem, M., and Scheduling, S. (2011). CD146 expression on primary nonhematopoietic bone marrow stem cells is correlated with in situ localization. *Blood* *117*, 5067–5077.
- Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* *28*, 511–515.
- Tudpor, K., van der Eerden, B.C., Jongwattanapisan, P., Roelofs, J.J., van Leeuwen, J.P., Bindels, R.J., and Hoenderop, J.G. (2015). Thrombin receptor deficiency leads to a high bone mass phenotype by decreasing the RANKL/OPG ratio. *Bone* *72*, 14–22.
- Vlachos, A., Rosenberg, P.S., Atsidaftos, E., Alter, B.P., and Lipton, J.M. (2012). Incidence of neoplasia in Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry. *Blood* *119*, 3815–3819.
- 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.
- Vogl, T., Eisenblätter, M., Völler, T., Zenker, S., Hermann, S., van Lent, P., Faust, A., Geyer, C., Petersen, B., Roebrock, K., et al. (2014). Alarmin S100A8/S100A9 as a biomarker for molecular imaging of local inflammatory activity. *Nat. Commun.* *5*, 4593.
- Walkley, C.R., Olsen, G.H., Dworkin, S., Fabb, S.A., Swann, J., McArthur, G.A., Westmoreland, S.V., Chambon, P., Scadden, D.T., and Purton, L.E. (2007). A microenvironment-induced myeloproliferative syndrome caused by retinoic acid receptor gamma deficiency. *Cell* *129*, 1097–1110.
- Walter, D., Lier, A., Geiselhart, A., Thalheimer, F.B., Huntscha, S., Sobotta, M.C., Moehrl, B., Brocks, D., Bayindir, I., Kaschutnig, P., et al. (2015). Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature* *520*, 549–552.
- Wang, X., Kua, H.Y., Hu, Y., Guo, K., Zeng, Q., Wu, Q., Ng, H.H., Karsenty, G., de Crombrughe, B., Yeh, J., and Li, B. (2006). p53 functions as a negative regulator of osteoblastogenesis, osteoblast-dependent osteoclastogenesis, and bone remodeling. *J. Cell Biol.* *172*, 115–125.
- Wei, Y., Dimicoli, S., Bueso-Ramos, C., Chen, R., Yang, H., Neuberger, D., Pierce, S., Jia, Y., Zheng, H., Wang, H., et al. (2013). Toll-like receptor alterations in myelodysplastic syndrome. *Leukemia* *27*, 1832–1840.
- Woloszynek, J.R., Rothbaum, R.J., Rawls, A.S., Minx, P.J., Wilson, R.K., Mason, P.J., Bessler, M., and Link, D.C. (2004). Mutations of the SBDS gene are present in most patients with Shwachman-Diamond syndrome. *Blood* *104*, 3588–3590.
- Xiao, Y., Wang, J., Song, H., Zou, P., Zhou, D., and Liu, L. (2013). CD34+ cells from patients with myelodysplastic syndrome present different p21 dependent premature senescence. *Leuk. Res.* *37*, 333–340.
- Yahata, T., Takahashi, T., Muguruma, Y., Ibrahim, A.A., Matsuzawa, H., Uno, T., Sheng, Y., Onizuka, M., Ito, M., Kato, S., and Ando, K. (2011). Accumulation of oxidative DNA damage restricts the self-renewal capacity of human hematopoietic stem cells. *Blood* *118*, 2941–2950.
- Zambetti, N.A., Bindels, E.M., Van Strien, P.M., Valkhof, M.G., Adisty, M.N., Hoogenboezem, R.M., Sanders, M.A., Rommens, J.M., Touw, I.P., and Raaijmakers, M.H. (2015). Deficiency of the ribosome biogenesis gene *Sbds* in hematopoietic stem and progenitor cells causes neutropenia in mice by attenuating lineage progression in myelocytes. *Haematologica* *100*, 1285–1293.

Zika Virus Disrupts Neural Progenitor Development and Leads to Microcephaly in Mice

Cui Li,^{1,2,7} Dan Xu,^{5,7} Qing Ye,^{3,7} Shuai Hong,^{1,2,7} Yisheng Jiang,¹ Xinyi Liu,^{1,2} Nana Zhang,^{3,6} Lei Shi,¹ Cheng-Feng Qin,^{3,*} and Zhiheng Xu^{1,4,*}

¹State Key Laboratory of Molecular Developmental Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

²University of Chinese Academy of Sciences, Beijing 100101, China

³Department of Virology, State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, China

⁴Parkinson's Disease Center, Beijing Institute for Brain Disorders, Beijing 100101, China

⁵Institute of Life Sciences, Fuzhou University, Fuzhou 350116, China

⁶Guangxi Medical University, Nanning 530021, China

⁷Co-first author

*Correspondence: qincf@bmi.ac.cn (C.-F.Q.), zhxu@genetics.ac.cn (Z.X.)

<http://dx.doi.org/10.1016/j.stem.2016.04.017>

SUMMARY

The link between Zika virus (ZIKV) infection and microcephaly has raised urgent global alarm. The historical African ZIKV MR766 was recently shown to infect cultured human neural precursor cells (NPCs), but unlike the contemporary ZIKV strains, it is not believed to cause microcephaly. Here we investigated whether the Asian ZIKV strain SZ01 could infect NPCs *in vivo* and affect brain development. We found that SZ01 replicates efficiently in embryonic mouse brain by directly targeting different neuronal lineages. ZIKV infection leads to cell-cycle arrest, apoptosis, and inhibition of NPC differentiation, resulting in cortical thinning and microcephaly. Global gene expression analysis of infected brains reveals upregulation of candidate flavivirus entry receptors and dysregulation of genes associated with immune response, apoptosis, and microcephaly. Our model provides evidence for a direct link between Zika virus infection and microcephaly, with potential for further exploration of the underlying mechanisms and management of ZIKV-related pathological effects during brain development.

INTRODUCTION

Recent world attention has been drawn to a global Zika virus (ZIKV) outbreak and its link with devastating cases of microcephaly. ZIKV infection is spreading rapidly within the Americas after originating from an outbreak in Brazil. So far 31 countries and territories in South and Central America have reported the circulation of this type of mosquito-borne flavivirus (Heukelbach et al., 2016). There is mounting concern about the association of ZIKV infection with approximately 5,000 fetal and newborn microcephaly cases and with serious neurological complications in adults, such as Guillain-Barré syndrome. In November 2015,

the Brazilian Ministry of Health reported a 20-fold increase in cases of neonatal microcephaly, which corresponds geographically and temporally to the ZIKV outbreak (Marrs et al., 2016). Due to this global threat, WHO declared a public health emergency of international concern on February 1 (Heymann et al., 2016; Marrs et al., 2016).

Precise timing of proliferation/self-renewal of neural progenitor cells (NPCs) and of their differentiation, neuronal migration, and maturation are essential for normal mammalian brain development. Disturbance of these processes leads to developmental brain disorders including microcephaly (Kriegstein and Alvarez-Buylla, 2009; Manzini and Walsh, 2011; Nowakowski et al., 2016; Thornton and Woods, 2009).

A causal association between ZIKV infection and microcephaly was proposed based on an increased incidence of microcephaly coinciding with the ZIKV outbreak and the detection of ZIKV in both microcephalic fetal brain tissues and the amniotic fluid of pregnant women with microcephalic fetuses (Calvet et al., 2016; Driggers et al., 2016; Marrs et al., 2016; Mlakar et al., 2016). In addition, ZIKV strain MR766 has been shown to be capable of infecting NPCs derived from human induced pluripotent stem cells (hiPSCs) (Tang et al., 2016). ZIKV infection induces cell death and deregulation of cell-cycle progression of hiPSCs, reducing their viability and growth as neurospheres and brain organoids (Garcez et al., 2016; Tang et al., 2016; Qian et al., 2016). However, there is still an urgent demand for direct evidence from mammalian animal models that ZIKV infection can cause microcephaly. Here we used an Asian ZIKV strain, SZ01, isolated from a patient infected in Samoa (Deng et al., 2016) to investigate whether ZIKV infects the embryonic mouse brain and affects brain development.

RESULTS

ZIKV Replicates Efficiently in Embryonic Mouse Brain and Causes Microcephaly

To test whether the contemporary ZIKV strain can infect the embryonic mouse brain, different titers of ZIKV SZ01 or culture medium (mock) were injected into one side of the cerebroventricular space/lateral ventricle (LV) of embryonic day 13.5 (E13.5)

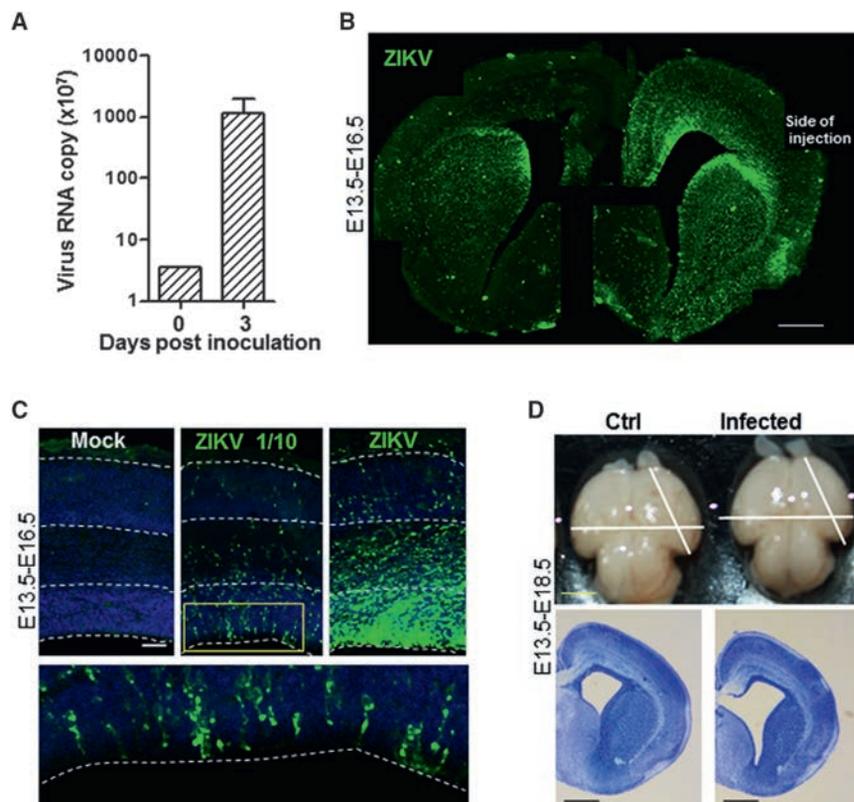


Figure 1. ZIKV Replicates Efficiently in Embryonic Mouse Brain and Causes Microcephaly

(A) Viral RNA copies were determined by real-time RT-PCR of whole brains. Data are means \pm SEM from two independent experiments.

(B–D) Embryonic brains were injected with ZIKV in the lateral ventricle at E13.5 and inspected at E16.5 (B and C) or E18.5 (D). (B) Coronal section of a whole brain slice stained with ZIKV serum. Scale bar, 400 μ m. (C) Images of coronal sections stained with DAPI to label nuclei (blue) and ZIKV antiserum (green). Left: mock-infected (control), middle: 1/10 diluted, right: undiluted virus. Lower panel: High-magnification of the area outlined by the yellow box. Scale bar, 40 μ m. (D) Images of E18.5 infected or mock-infected littermate brains. Lower panel: similar position of coronal sections of cortices with Nissl staining. Scale bar, 5 mm (upper panel), 1 mm (lower panel). See also Figure S1.

ZIKV Infects NPCs and Causes Cell Death

Increased cell death has been shown as one of the causes of microcephaly (Xu et al., 2014). In order to determine whether cell death contributes to the smaller size of infected brains, we infected or mock-infected littermate brains at E13.5 and inspected them 3 or 5 days later.

Compared to mock-infected brains, many cells in the intermediate zone (IZ) and CP were strongly positive for the activated form of caspase3 3 days after infection, although a weak signal was also detected in the SVZ (Figure S2A). Interestingly, there were many more strongly positive cells 5 days after infection and these were mostly located in the CP (Figures 2D). This suggests that ZIKV infection induces cell death mainly in immature neurons and mature neurons.

To identify which cell types were infected by ZIKV in the embryonic brain, we examined E16.5 brain slices with cell markers including Sox2 and Pax6, markers for apical progenitor cells and radial glial cells; Tbr2, a marker for intermediate or basal progenitor cells (IPCs/BPCs); and Tuj1 and Dcx, markers for immature neurons. As shown in Figures 3A and S3, infected cells were positive for Sox2, Pax6, or Tbr2, suggesting that ZIKV mainly infects NPCs or IPCs/BPCs. Some infected cells were also positive for Tuj1 or Dcx in the IZ and CP, which had possibly differentiated from NPCs infected earlier (Figures 3B and S3C). Together, these results establish that NPCs in the developing brain are the main direct targets of ZIKV. However, almost all cells in the brain including those in the CP were also positive for ZIKV 5 days after infection (Figures 2B–2D and S2B), indicating that post-mitotic neurons can also be the target of ZIKV.

ZIKV Infection Leads to Dysregulation of NPC Cell Cycle and Differentiation

We went on to determine whether ZIKV infection would affect NPC proliferation and differentiation. It was interesting to notice that there were significantly fewer mitotic cells in the VZ of the E16.5 brains infected at E13.5 (Figure S3E). This was

littermate brains and inspected 3 or 5 days later. Infection of the embryonic brains was verified by real-time PCR and with immunocytochemistry. The brains were readily infected since an \sim 300-fold increase of viral RNA copies was detected 3 days after infection (Figure 1A). Convalescent serum from a ZIKV-infected patient was used for immunocytochemistry staining and co-stained well with ZIKV envelope antibody (Figure S1). The numbers of ZIKV-infected cells corresponded to the amount of virus applied, with most of the infected cells located in the ventricular and subventricular zones (VZ and SVZ), where NPCs are located (Figures 1B and 1C). There were far fewer infected cells in the midbrain or the cortical plate (CP), where post-mitotic neurons are located during the infection window. Together with the random diffusion of virus injected in the LV, this may account for the variability of the infection pattern in Figure 1B.

Importantly, brains of smaller sizes compared to those of their mock infected littermates were detected 5 days after infection (Figure 1D). In addition, we noticed enlarged LVs and a thinner CP and VZ/SVZ in the infected brains (Figures 1D and 2A). We also stained the cortex with different cortical layer markers at E18.5, including Tbr1 and Foxp2, and found that ZIKV infection led to a thinning of cortical layers without apparent disturbance of the lamination (Figures 2A–2C). We were also interested in examining postnatal stages to confirm that viral infection of SZ01 leads to microcephalic brains. Unfortunately, we tried twice to infect and mock-infect littermates at E13.5 and found that the infected newborn pups were eaten by their lactating mothers within 2 days. This prevents the analysis of postnatal stages for the time being, and we therefore focused on analyzing infected brains at embryonic stages in the present study.

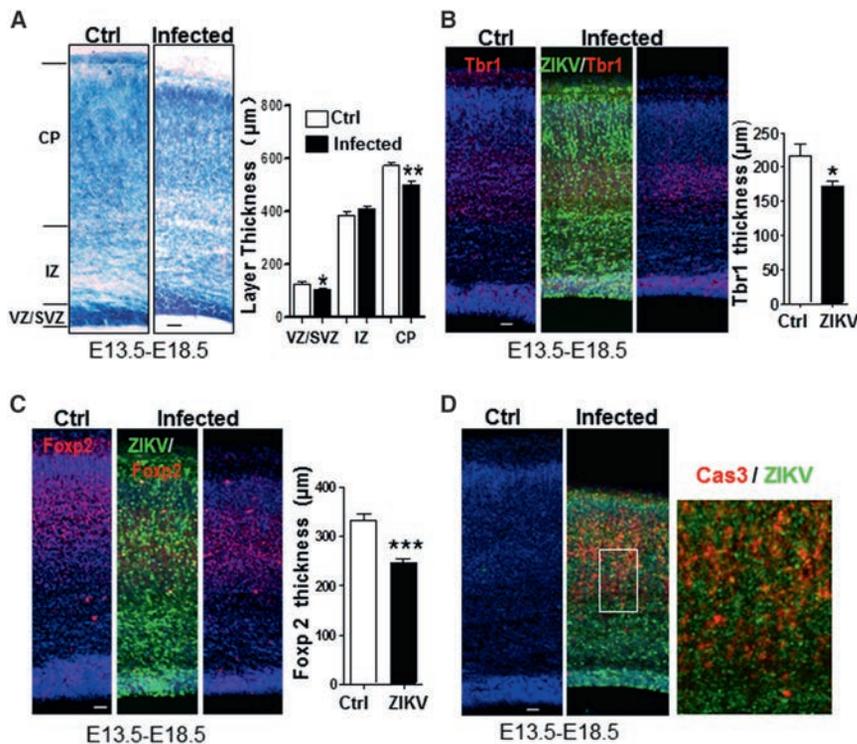


Figure 2. ZIKV Infection Leads to Thinner Cortex and Increased Cell Death

Embryonic littermate brains were infected at E13.5 and inspected at E18.5.

(A) Nissl staining of infected and mock-infected cortices. Right panel: quantification of layer thickness. Ctrl and Infected: $n = 12$. n : brain slices in two controls or three ZIKV-infected brains.

(B and C) Images of cortices stained for Tbr1 (B) or Foxp2 (C) and ZIKV (green). Right panels: quantification of thickness measured with individual markers. Ctrl: $n = 5$. Infected: $n = 7$. n : brain slices in two controls or six ZIKV-infected brains.

(D) Coronal sections of cortices stained with the activated form of caspase3 (Cas3, red) and ZIKV (green). Right panel: high magnification of the area outlined by the white box.

All data are means \pm SEM, *t* test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All scale bars, 40 μ m. See also Figure S2.

accompanied by more centrosomes at the ventricular surface that were facing away from the nuclei. We therefore examined phosphorylated H3, a marker for cells in the M phase. As shown in Figure 3C, there were substantially fewer cells positive for phosphorylated H3 in the VZ/SVZ of ZIKV-infected brains, indicating a decreased number of NPCs in M phase. In addition, more Pax6⁺ cells were found in infected brains 1 hr after BrdU injection, while significantly fewer Tbr2⁺ cells were found 24 hr after BrdU injection (Figures 3D and S3D). Intriguingly, 24 hr after BrdU labeling, the infected cells positive for Ki67 or BrdU were more concentrated in the VZ (Figure 3E). Furthermore, most of the infected cells in the VZ showed the morphology of S, G1, or G2 phase cells with the somas away from the ventricular surface but tethered there through elongated end-feet (Figure 1C). Finally, the percentage of cells in the infected brains positive for both Ki67 and BrdU was increased after 24 hr of BrdU labeling (Figure 3E), suggesting that ZIKV infection inhibits NPC cell-cycle exit and differentiation. The above results indicate that ZIKV infection suppresses NPC proliferation, the transition of Pax6⁺ radial glial cells to Tbr2⁺ IPCs, and NPC differentiation. They are in agreement with the finding that ZIKV infection leads to S phase arrest of hiPSCs (Tang et al., 2016) and that proliferating NPCs exhibit a much longer S phase than those committed to neuronal differentiation (Arai et al., 2011).

ZIKV Infection Induces Immune Response and Deregulation of Microcephaly-Associated Genes

To investigate the global impact of ZIKV infection on the whole developing brain at the molecular level, we carried out global transcriptome analyses (RNA-seq). Genome-wide analyses identified a large number of differentially expressed genes at 3 days after viral infection (Table S2). Gene Ontology analyses revealed a

particular enrichment of upregulated genes in immune-response-related and apoptosis pathways (Figure 4A and Table S2). Very notable were the genes related to cytokine production and the response to cytokines, suggesting that cytokines play a critical role in the pathogenesis of ZIKV infection (Figure 4A and Table S2). Many antiviral response genes were reported to be induced by ZIKV in human skin fibroblasts (Hamel et al., 2015). Although there is a much longer list of related genes in our dataset, we have confirmed all of those reported by Hamel et al., including Tlr3, Ddx58, Ifih1, Oas2, Isg15, and Mx1, in addition to Ccl5, Cxcl10, and Ifnb1, which were not detected in controls but were expressed in infected brains (Figure 4C and Table S2). Interestingly, many candidate flavivirus entry receptors were also induced, most notably AXL, which has been predicted as a ZIKV receptor (Nowakowski et al., 2016) (Figure 4D and Table S2). In contrast, many genes involved in cell proliferation, differentiation, migration, and organ development were downregulated (Figure 4B and Table S2). It was of interest that most of the microcephaly-associated genes were significantly downregulated (Figure S4) in the dataset of ZIKV-infected hNPCs (Tang et al., 2016). We confirmed the significant downregulation of seven of them, including ASPM, CASC5, CENPF, MCPH1, RBBP8, STIL, and TBR2, in our dataset or by real-time PCR of the virus-injected sides of the brains (Figure 4E). The downregulation of microcephaly-associated genes that we found largely overlaps with that found in the hNPC dataset, although the downregulation of CEP152 and WDR62 was not very significant. Therefore, these global transcriptome datasets not only support our cell biology findings but also provide a useful resource for the exploration of the underlying pathogenesis and for the potential treatment of ZIKV infection when combined with a dataset from hiPSCs (Tang et al., 2016).

DISCUSSION

Our results demonstrate that ZIKV can directly infect different lineages of NPCs and immature neurons in vivo in the beginning

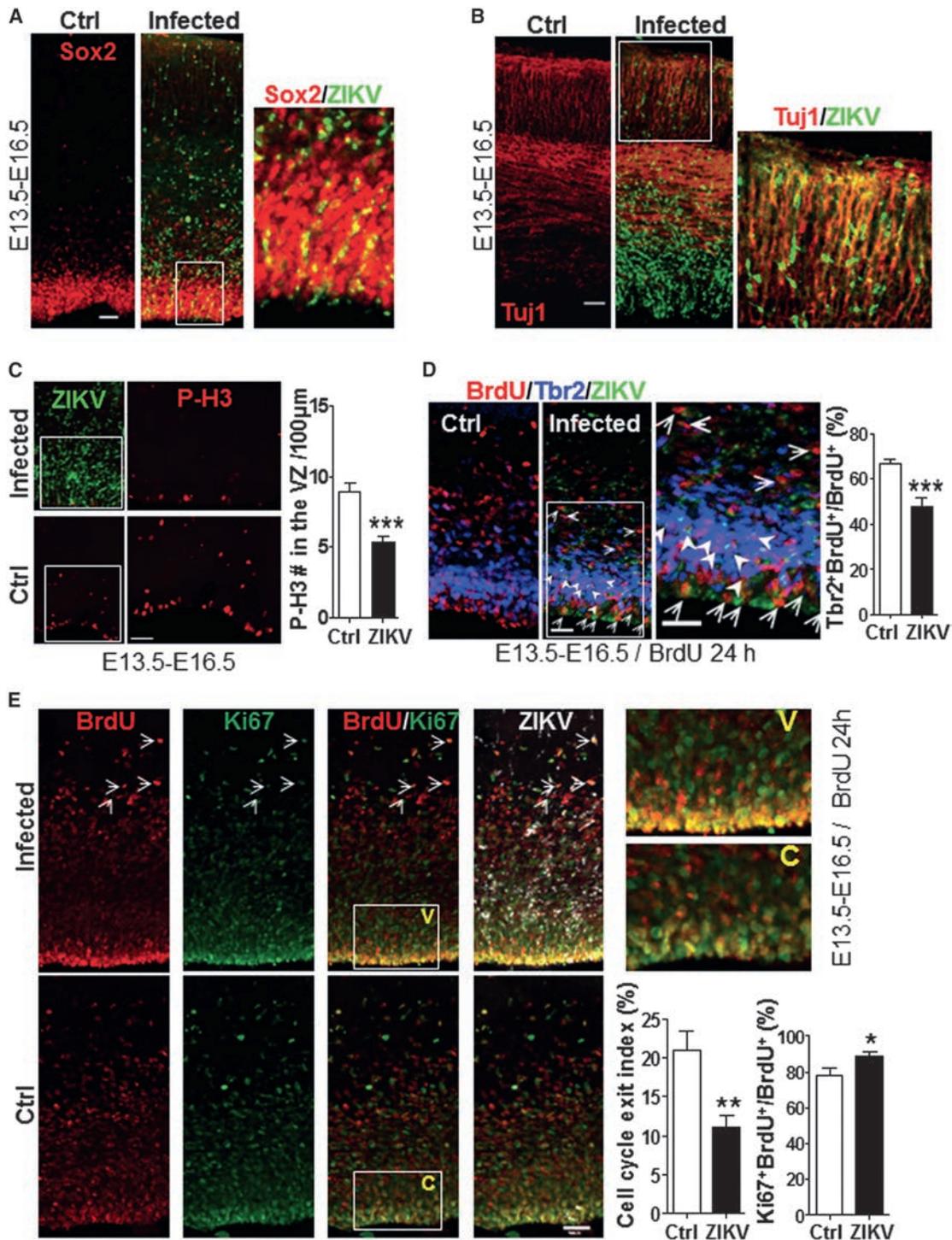


Figure 3. ZIKV Infects Different Lineages of Neural Progenitor Cells

Coronal sections of E16.5 cortices infected or mock-infected at E13.5 are shown.

(A–C) Sections were stained with antibodies for Sox2 (A), Tuj1 (B), Phospho-Histone H3 (C) (red), and ZIKV (green). Right panel in (A) and (B) and middle panel in (C): high magnification of the areas outlined by the white boxes. Right panel in (C): quantification of P-H3⁺ cells in the VZ. $n = 7$.

(D and E) Coronal sections from infected or mock-infected E16.5 littermates 24 hr after BrdU labeling. (D) Sections were stained for ZIKV (green), Tbr2 (blue), and BrdU (red). White arrows: ZIKV and BrdU double-positive cells; arrowheads: Tbr2, BrdU, and ZIKV triple-positive cells. Right panel: quantification of Tbr2 and BrdU double-positive cells per total BrdU⁺ cells within ZIKV-positive or -negative cells (control). $n = 7$. (E) Sections stained for ZIKV (white), Ki67 (green), and BrdU

(legend continued on next page)

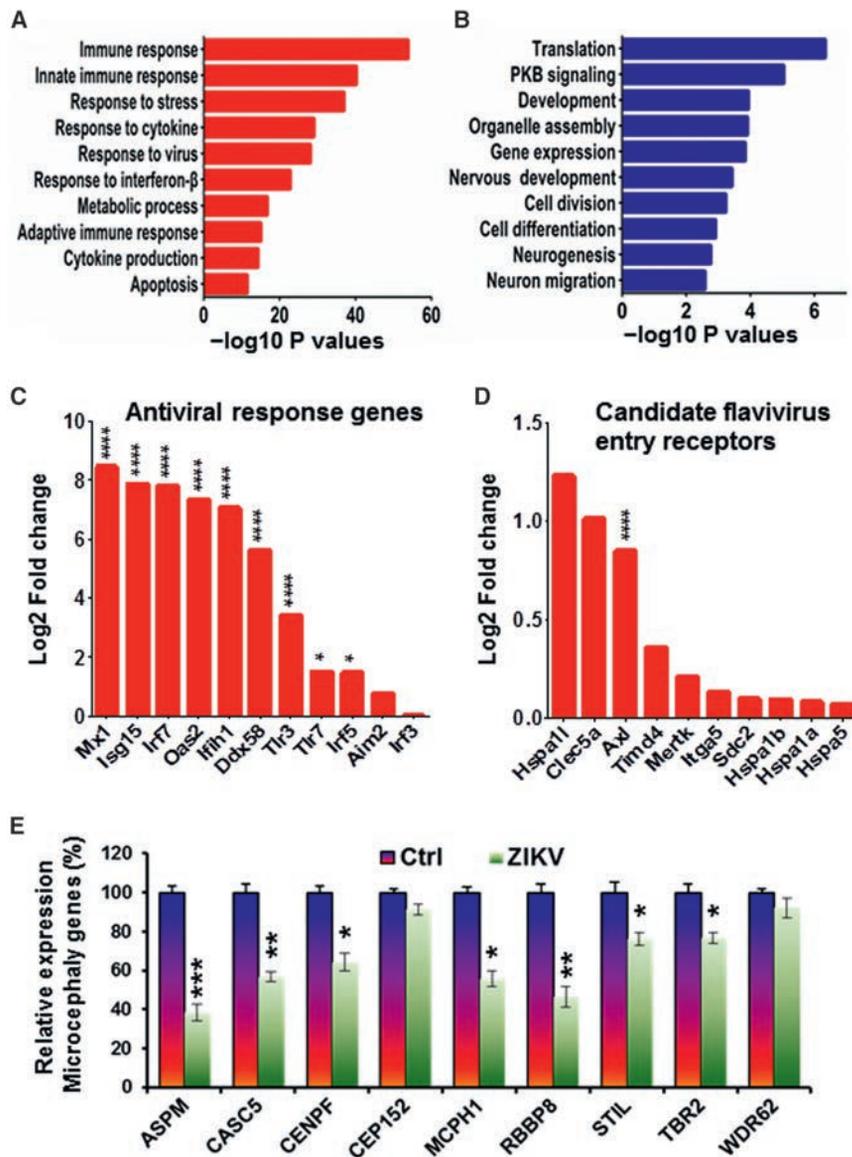


Figure 4. ZIKV Infection Causes Global Deregulation of Gene Expression

(A and B) RNA-seq analysis of whole ZIKV-infected and mock-infected littermate brains (E13.5–E16.5). Genes with significant differences in expression were subjected to GO analyses. The typical 20 most significant terms are selected for upregulated (A) and downregulated (B) genes, respectively. The $-\log_{10} p$ values are indicated by bar plots.

(C and D) RNA expression analysis of RNA-seq results in Table S2. Bar plots indicate induction of antiviral response genes (C) or the candidate flavivirus entry receptors (D) in the ZIKV-infected brains.

(E) Expression of microcephaly-associated genes between viral injection side of brains and controls as determined by real-time PCR.

All data are means \pm SD. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, **** $p < 0.0001$. See also Figure S4 and Table S2.

because they were very sick. Therefore, in the future it will be desirable to lower the titer of ZIKV infection to determine whether the pups will survive longer, because if they do, it will permit determination of the long-term consequences of ZIKV infection. It will also be intriguing to investigate whether ZIKV infects adult neural stem cells and affects adult neurogenesis.

In summary, our results show direct effects of ZIKV on NPC development, including proliferation, differentiation, and cell death, which may link ZIKV with the development of microcephaly. Our study also provides insights into indirect effect of ZIKV infection on induced immune responses, including cytokine production and the effects of these cytokines

of infection and replicates in these cells with high efficiency. Due to the dramatically higher level of ZIKV several days after infection, neurons are also targeted at E18.5, indicating that neurons are infected at much lower efficiency than NPCs. Infection of NPCs leads to attenuated NPC expansion through virally induced apoptosis and cell-cycle dysregulation. Together with ZIKV's effects on NPC differentiation, these effects are likely to account for microcephaly in human fetuses or newborn babies. Our study also indicates that more attention should be focused on NPCs and immature neurons regarding the roles of these cells in ZIKV-associated neuropathology. It was unfortunate that our infected pups did not survive for a long time after birth. It is very likely that they were eaten by their lactating mothers

on neural development. Moreover, the combination and comparison of our global transcriptome datasets of infected brains with that of hiPSCs (Tang et al., 2016) will provide valuable resources for further investigation of the underlying cellular and molecular mechanisms and management of ZIKV-related pathological effects during neural development.

EXPERIMENTAL PROCEDURES

ZIKV Preparation and Animal Infection

1 μ l of ZIKV SZ01 (GenBank accession number: GEO: KU866423) virus stock (6.5×10^5 PFU/ml) (Deng et al., 2016) or stock diluted 10-fold was injected into one side of the cerebroventricular space/LV of the E13.5 ICR mouse brains and inspected 3–5 days later or after birth. Diluted virus was used only in Figure 1A

(red). Right upper panel: high magnification of the area outlined by the white box. Right lower panels: quantification of cell-cycle exit index (Ki67⁺ BrdU⁺/BrdU⁺) (n = 12) and Ki67⁺ BrdU⁺ cells in BrdU-labeled cells within ZIKV-infected or uninfected control cells (n = 9). Arrows: Ki67⁺ BrdU⁺ in ZIKV-infected cells.

All data are means \pm SEM; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. n: brain slice numbers from four or five different ZIKV or control littermate brains. All scale bars, 40 μ m. See also Figure S3.

while all others were undiluted. For each pregnant dam, around two-thirds of littermates were injected with ZIKV while one-third were injected with culture medium (mock) for proper controls. For BrdU labeling experiments, pregnant dams were injected with 50 mg/kg (body weight) of BrdU in solution at either 1 hr or 24 hr before scarification. All experimental procedures involved were performed according to protocols approved by the Institutional Animal Care and Use Committee at Beijing Institute of Microbiology and Epidemiology.

Immunohistochemistry and Antibodies

For cryosections, tissues were fixed in 4% PFA, dehydrated in 30% sucrose, and frozen in TFM (tissue freezing medium). Sections (thickness: E16.5, 50 μ m; E18.5, 40 μ m) were used for immunofluorescence staining as described previously (Xu et al., 2014; Zhang et al., 2014). The antibodies used for immunostaining were Sox2 (Abcam, ab97959, 1:1000), Pax6 (Covance, PRB-278P, 1:1000), Tbr2 (Millipore, ab15894, 1:1000), β -III Tubulin (Abcam, ab7751, 1:1000), γ -Tubulin (Abcam, ab11316, 1:1000), Phospho-Histone 3 (P-H3) (Abcam, ab10543, 1:1000), Ki67 (Abcam, ab15580, 1:1000), BrdU (Abcam, ab6326, 1:500), Activated-caspase3 (CST, 9664s, 1:500), Dcx (CST, 4604s, 1:1000), Tbr1 (Abcam, ab31940, 1:500), Foxp2 (Abcam, ab16046, 1:500), and ZIKV serum from a patient (1:100). Control serum for ZIKV was from a healthy person (1:100), and mouse serum immunized with recombinant ZIKV envelope (E) protein was from GenBank (GEO: JN860885) (1:500, used only in Figure S1B). Nuclei were stained with DAPI (Invitrogen). Sections stained for BrdU detection were subject to 10 min 1N HCl on ice and 30 min 2N HCl at 37°C prior to blocking.

Nissl Staining

Brian slices were stained with 0.1% toluidine blue for 20 min and dehydrated in turn by 70%, 96%, and 99% ethanol (45 s, twice for each). Finally, slices were hyalinized by Xylene for more than 30 min.

Confocal Imaging and Quantification

Slices were imaged on an LSM 700 (Carl Zeiss) confocal microscope, and the images were analyzed with Imaris, ImageJ, and Photoshop as described previously (Xu et al., 2014; Zhang et al., 2014). All data were analyzed using Prism software (GraphPad) or Excel. Statistical evaluation was performed by Student's unpaired t test. Data are presented as mean \pm SEM.

RNA-Seq Analyses

Whole E16.5 brains 3 days after ZIKV or mock infection (two for each group) were used for global transcriptome analysis by Annoroad Co. Significantly differentially expressed genes were identified when we compared Normalized Reads Count between ZIKV and mock infection groups with $p < 0.05$ and $|\text{Log}_2\text{FoldChange}| > 0.263$. Significance of Gene Ontology term enrichment was estimated with Fisher's Exact Test (p value).

Real-Time PCR

For viral RNA copies: total RNA was extracted from whole embryo brains and viral RNA copies were determined by real-time PCR (Johnson et al., 2005). Primers and fluorogenic probes for ZIKV detection were shown in Table S1. 2 μ l of RNA samples was mixed with each primer of a final concentration of 0.2 μ M to prepare the reaction mixtures in accordance with the One Step PrimeScript RT-PCR Kit instructions, and they were moved to a reverse transcription reaction at 42°C for 5 min, followed by PCR amplification using a 60°C annealing temperature for 40 cycles. The standard curve of viral RNA copies was determined from 10-fold dilutions of ZIKV RNA with known concentrations, and viral RNA copies in embryo brains were calculated.

Determining expression of microcephaly-associated genes between viral injection side of the brains and mock-infected control (E13.5–16.5) was performed as described previously (Xu et al., 2014), and primers used are listed in Table S1.

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes four figures and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.04.017>.

AUTHOR CONTRIBUTIONS

Z.X. and C.-F.Q. conceived of the research and Z.X. designed the study and wrote the manuscript. D.X., C.L., and S.H. designed and performed most of the experiments. Q.Y. performed ZIKV real-time PCR assays and took care of the mice. Y.J. did bioinformatics analysis. X.L. did Nissl staining. Everyone contributed to the writing.

ACKNOWLEDGMENTS

We thank Drs. Fu Gao, Xia Jin, Weixiang Guo, Lloyd Greene, and Feng Zhang for valuable advice, and we thank Fu-Chun Zhang for serum from the ZIKV patient. This work was supported by grants from NSF (China) (31430037/31271156/81522025/8151101191) and the MOST (China) "973" program (2014CB942801/2013DFA31990/2012YQ03026006).

Received: April 15, 2016

Revised: April 25, 2016

Accepted: April 29, 2016

Published: May 11, 2016

REFERENCES

- Arai, Y., Pulvers, J.N., Haffner, C., Schilling, B., Nüsslein, I., Calegari, F., and Huttner, W.B. (2011). Neural stem and progenitor cells shorten S-phase on commitment to neuron production. *Nat. Commun.* 2, 154.
- Calvet, G., Aguiar, R.S., Melo, A.S., Sampaio, S.A., de Filippis, I., Fabri, A., Araujo, E.S., de Sequeira, P.C., de Mendonça, M.C., de Oliveira, L., et al. (2016). Detection and sequencing of Zika virus from amniotic fluid of fetuses with microcephaly in Brazil: a case study. *Lancet Infect. Dis.* S1473-3099(16) 00095-5. [http://dx.doi.org/10.1016/S1473-3099\(16\)00095-5](http://dx.doi.org/10.1016/S1473-3099(16)00095-5).
- Deng, Y.Q., Zhao, H., Li, X.F., Zhang, N.N., Liu, Z.Y., Jiang, T., Gu, D.Y., Shi, L., He, J.A., Wang, H.J., et al. (2016). Isolation, identification and genomic characterization of the Asian lineage Zika virus imported to China. *Sci. China Life Sci.* 59, 428–430. <http://dx.doi.org/10.1007/s11427-016-5043-4>.
- Driggers, R.W., Ho, C.Y., Korhonen, E.M., Kuivanen, S., Jääskeläinen, A.J., Smura, T., Rosenberg, A., Hill, D.A., DeBiasi, R.L., Vezina, G., et al. (2016). Zika Virus Infection with Prolonged Maternal Viremia and Fetal Brain Abnormalities. *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1601824>.
- Garcez, P.P., Loiola, E.C., Madeiro da Costa, R., Higa, L.M., Trindade, P., Delvecchio, R., Nascimento, J.M., Brindeiro, R., Tanuri, A., and Rehen, S.K. (2016). Zika virus impairs growth in human neurospheres and brain organoids. *Science*, aaf6116. <http://dx.doi.org/10.1126/science.aaf6116>.
- Hamel, R., Dejamac, O., Wichit, S., Ekcharyawat, P., Neyret, A., Luplertlop, N., Perera-Lecoin, M., Surasombatpattana, P., Talignani, L., Thomas, F., et al. (2015). Biology of Zika Virus Infection in Human Skin Cells. *J. Virol.* 89, 8880–8896. <http://dx.doi.org/10.1128/JVI.00354-15>.
- Heukelbach, J., Alencar, C.H., Kelvin, A.A., De Oliveira, W.K., and Pamplona de Góes Cavalcanti, L. (2016). Zika virus outbreak in Brazil. *J. Infect. Dev. Ctries.* 10, 116–120.
- Heymann, D.L., Hodgson, A., Sall, A.A., Freedman, D.O., Staples, J.E., Althabe, F., Baruah, K., Mahmud, G., Kandun, N., Vasconcelos, P.F., et al. (2016). Zika virus and microcephaly: why is this situation a PHEIC? *Lancet* 387, 719–721.
- Johnson, B.W., Russell, B.J., and Lanciotti, R.S. (2005). Serotype-specific detection of dengue viruses in a fourplex real-time reverse transcriptase PCR assay. *J. Clin. Microbiol.* 43, 4977–4983.
- Kriegstein, A., and Alvarez-Buylla, A. (2009). The glial nature of embryonic and adult neural stem cells. *Annu. Rev. Neurosci.* 32, 149–184.
- Manzini, M.C., and Walsh, C.A. (2011). What disorders of cortical development tell us about the cortex: one plus one does not always make two. *Curr. Opin. Genet. Dev.* 21, 333–339.
- Marrs, C., Olson, G., Saade, G., Hankins, G., Wen, T., Patel, J., and Weaver, S. (2016). Zika Virus and Pregnancy: A Review of the Literature and Clinical Considerations. *Am. J. Perinatol.* <http://dx.doi.org/10.1055/s-0036-1580089>.

- Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., Kolenc, M., Resman Rus, K., Vesnaver Vipotnik, T., Fabjan Vodusek, V., et al. (2016). Zika Virus Associated with Microcephaly. *N. Engl. J. Med.* 374, 951–958.
- Nowakowski, T.J., Pollen, A.A., Di Lullo, E., Sandoval-Espinosa, C., Bershteyn, M., and Kriegstein, A.R. (2016). Expression Analysis Highlights AXL as a Candidate Zika Virus Entry Receptor in Neural Stem Cells. *Cell Stem Cell*. S1934-5909(16)00118-1. <http://dx.doi.org/10.1016/j.stem.2016.03.012>.
- Qian, X., Nguyen, H.N., Song, M.M., Hadiono, C., Ogden, S.C., Hammack, C., Yao, B., Hamersky, G.R., Jacob, F., Zhong, C., et al. (2016). Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell*, in press. Published online April 21, 2016. <http://dx.doi.org/10.1016/j.cell.2016.04.032>.
- Tang, H., Hammack, C., Ogden, S.C., Wen, Z., Qian, X., Li, Y., Yao, B., Shin, J., Zhang, F., Lee, E.M., et al. (2016). Zika Virus Infects Human Cortical Neural Progenitors and Attenuates Their Growth. *Cell Stem Cell*. S1934-5909(16)00106-5. <http://dx.doi.org/10.1016/j.stem.2016.02.016>.
- Thornton, G.K., and Woods, C.G. (2009). Primary microcephaly: do all roads lead to Rome? *Trends Genet.* 25, 501–510.
- Xu, D., Zhang, F., Wang, Y., Sun, Y., and Xu, Z. (2014). Microcephaly-associated protein WDR62 regulates neurogenesis through JNK1 in the developing neocortex. *Cell Rep.* 6, 104–116.
- Zhang, F., Xu, D., Yuan, L., Sun, Y., and Xu, Z. (2014). Epigenetic regulation of Atrophia1 by lysine-specific demethylase 1 is required for cortical progenitor maintenance. *Nat. Commun.* 5, 5815.

Induction of Expandable Tissue-Specific Stem/Progenitor Cells through Transient Expression of YAP/TAZ

Tito Panciera,^{1,5} Luca Azzolin,^{1,5} Atsushi Fujimura,^{1,5} Daniele Di Biagio,¹ Chiara Frasson,² Silvia Bresolin,² Sandra Soligo,¹ Giuseppe Basso,² Silvio Bicciato,³ Antonio Rosato,⁴ Michelangelo Cordenonsi,¹ and Stefano Piccolo^{1,6,*}

¹Department of Molecular Medicine, University of Padua School of Medicine, viale Colombo 3, 35126 Padua, Italy

²Department of Woman and Child Health, Haemato-Oncology Laboratory, University of Padua, via Giustiniani 3, 35128 Padua, Italy

³Center for Genome Research, Department of Life Sciences, University of Modena and Reggio Emilia, via G. Campi 287, 41100 Modena, Italy

⁴Istituto Oncologico Veneto IOV-IRCCS and Department of Surgery, Oncology, and Gastroenterology, University of Padua School of Medicine, via Gattamelata 64, 35128 Padua, Italy

⁵Co-first author

⁶Lead Contact

*Correspondence: piccolo@bio.unipd.it

<http://dx.doi.org/10.1016/j.stem.2016.08.009>

SUMMARY

The ability to induce autologous tissue-specific stem cells in culture could have a variety of applications in regenerative medicine and disease modeling. Here we show that transient expression of exogenous YAP or its closely related paralogue TAZ in primary differentiated mouse cells can induce conversion to a tissue-specific stem/progenitor cell state. Differentiated mammary gland, neuronal, and pancreatic exocrine cells, identified using a combination of cell sorting and lineage tracing approaches, efficiently convert to proliferating cells with properties of stem/progenitor cells of their respective tissues after YAP induction. YAP-induced mammary stem/progenitor cells show molecular and functional properties similar to endogenous MaSCs, including organoid formation and mammary gland reconstitution after transplantation. Because YAP/TAZ function is also important for self-renewal of endogenous stem cells in culture, our findings have implications for understanding the molecular determinants of the somatic stem cell state.

INTRODUCTION

Stem cells (SCs) display the capacity to self-renew when they divide and to generate a differentiated progeny. Somatic SCs operate in multiple adult organs for continuous tissue renewal or repair after injury. However, these cells are still mainly defined by operational definitions and cell surface markers, rather than the molecular traits that govern their special status (Blanpain and Fuchs, 2014). Unlimited availability of normal, somatic SCs will be critical for effective organ repopulation in regenerative medicine applications to understand SC biology and for disease modeling in the Petri dish. However, these efforts remain limited by the fact that SCs are rare and difficult to purify or expand from native tissues.

Direct conversion of terminally differentiated cells back into their corresponding tissue-specific SCs may represent an attractive approach to obtain somatic SCs. Indeed, several reports have recently highlighted a surprising plasticity in somatic cell fates because differentiated cells can return to a SC status under special conditions, such as tissue damage (Blanpain and Fuchs, 2014). However, the identity of the factors able to control the somatic SC status remains poorly understood, limiting the exploitation of such plasticity.

Dysregulation of the Hippo signaling pathway has been recently associated with cell fate plasticity. In tumors, activation of the Hippo pathway transcriptional effectors YAP/TAZ can reprogram non-stem cancer cells into cancer SCs (Cordenonsi et al., 2011). Genetic inactivation of the Hippo cascade induces liver overgrowth, in part by influencing liver cell fate and zonation (Lee et al., 2016; Yimlamai et al., 2014; Fitamant et al., 2015). Nuclear YAP/TAZ proteins are found in anatomical locations enriched in tissue SCs, possibly as a consequence of their regulation by structural and chemical signals associated with the SC niche (Piccolo et al., 2014). That said, genetic ablation of YAP and/or TAZ from several adult organs in mice, such as liver, pancreas, intestine, and mammary gland, revealed that these factors are surprisingly dispensable during normal tissue homeostasis. Strikingly, however, in those same tissues, YAP/TAZ become essential for organ regrowth after tissue damage or oncogenic transformation (Azzolin et al., 2014; Zanconato et al., 2016). This suggests that YAP/TAZ remain latent—and, thus, apparently dispensable—in normal adult tissues but are called into action to generate new stem cells.

With this background in mind, we considered the possibility that ectopic expression of YAP or TAZ may be instrumental to turn differentiated cells into somatic stem-like cells. Here we tested this hypothesis *in vitro* by using distinct paradigms of terminal differentiation; that is, luminal mammary gland cells, neurons, and pancreatic exocrine cells. We found that transient expression of YAP/TAZ indeed converts differentiated cells into cells displaying multiple features of their corresponding tissue-specific SCs. Notably, the ability of YAP/TAZ to impart an SC state pairs with their endogenous function in isolated native SCs, where they are essential for preserving organoid-forming

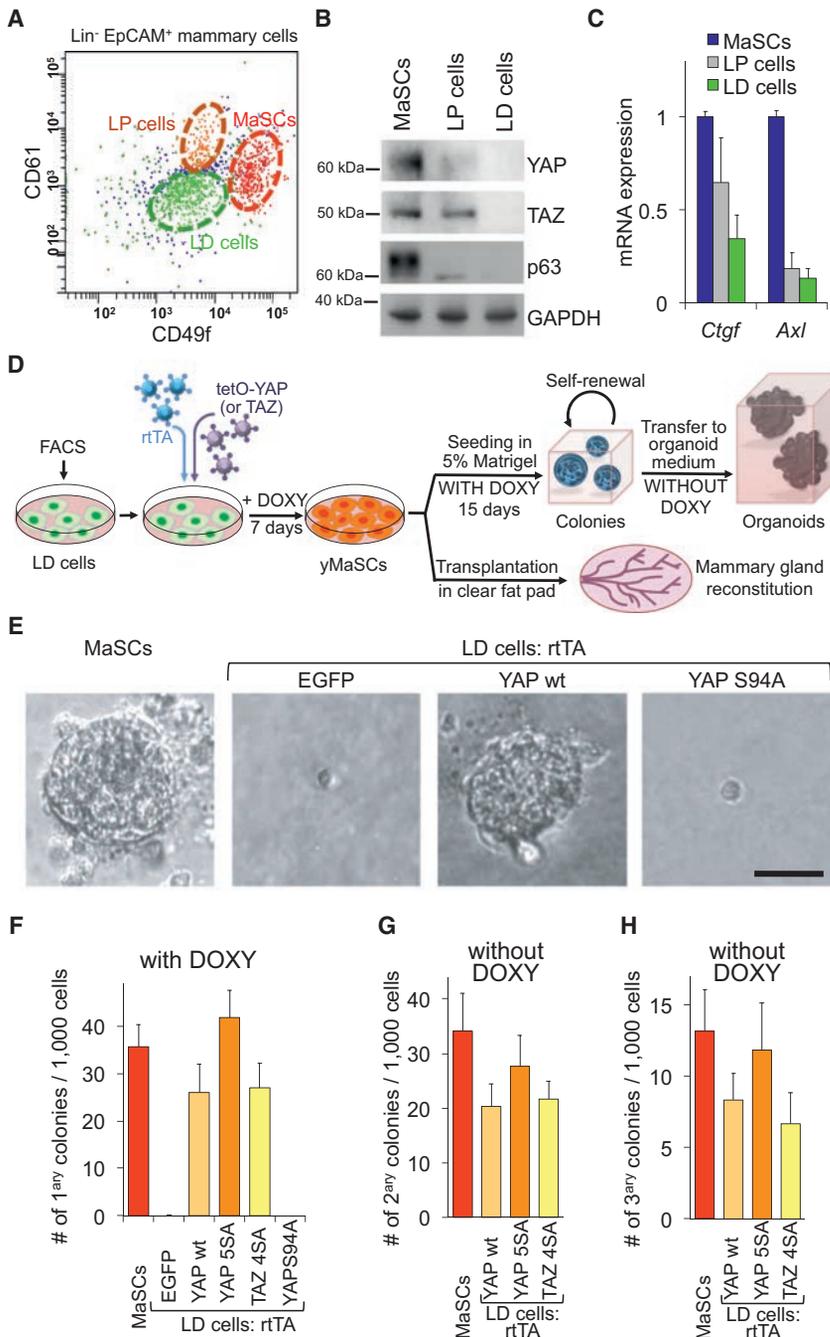


Figure 1. YAP and TAZ Convert Luminal Differentiated Cells in yMaSCs

(A) FACS profile showing the distribution of Lin⁻/EpCAM⁺ mammary cells according to their CD49f/CD61 antigenic profile. Three subpopulations were separated: a MaSC-enriched fraction (EpCAM^{low}CD49f^{high}CD61⁺), luminal progenitors (EpCAM^{high}CD49f^{low}CD61⁺), and luminal differentiated cells (EpCAM^{high}CD49f^{low}CD61⁻). See also Figure S1A for the FACS pattern of EpCAM and Figures S1B–S1E for characterization of the three subpopulations.

(B) Western blots for YAP, TAZ, and the MaSC marker p63 in the indicated purified populations. GAPDH served as a loading control.

(C) qRT-PCRs for *Ctgf* and *Axl* in the indicated cell populations (mean + SD). The results are representative of three independent experiments (each using mammary glands from 20 mice) performed in triplicate.

(D) Schematic of the experiments performed with LD cells.

(E and F) Representative images (E) and quantifications (F) of mammary colonies formed by the indicated cells 15 days after seeding in mammary colony medium. The data in (F) are presented as mean + SD and are representative of five independent experiments, each with six technical replicates.

(G and H) Quantifications of secondary (G) and tertiary (H) colonies formed by primary mammary colonies after dissociation and re-seeding in mammary colony medium without doxycycline. The data are representative of three independent experiments performed with six technical replicates and presented as mean + SD.

See also Figure S1.

potential. Our work therefore reveals that a single factor can induce somatic stem cell features in cells of different lineages.

RESULTS

YAP/TAZ Revert Differentiated Cells of the Mammary Gland into MaSC-like Cells

The mammary gland represents a classic model system for the study of epithelial SCs and tissue regeneration. Remarkably, implantation of mammary gland SCs (MaSCs) into the mammary fat pad is sufficient to regenerate an entire ductal tree, with MaSCs

contributing to both the luminal and myoepithelial lineages (Blanpain and Fuchs, 2014). Given that YAP/TAZ can reprogram non-stem mammary tumor cells into their corresponding cancer stem cells (CSCs) (Cordenonsi et al., 2011), we hypothesized that expression of YAP/TAZ may bestow stem-like characteristics on normal mammary cells as well.

To address this, we used fluorescence-activated cell sorting (FACS) to isolate terminally differentiated luminal cells (LD, EpCAM^{high}CD49f^{low}CD61⁻) from dissected

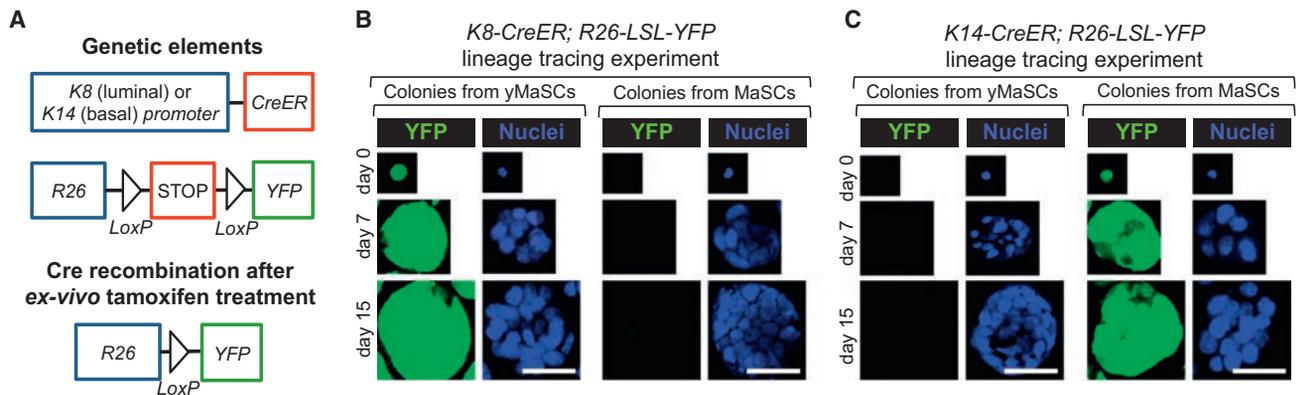


Figure 2. Lineage Tracing of yMaSCs

(A) Schematic of the genetic lineage tracing strategies to trace different mammary cell lineages.

(B and C) Immunostainings with anti-YFP of K8-CreER/R26-YFP-traced (B) or K14-CreER/R26-YFP-traced (C) cells during colony formation. Days indicate the time in mammary colony medium. Scale bars, 62 μ m.

See also Figure S2.

were plated on collagen-coated dishes and transduced with doxycycline (Doxy)-inducible lentiviral vectors encoding for wild-type (WT) YAP or the activated versions of YAP and TAZ (i.e., YAP5SA or TAZ4SA, lacking inhibitory phosphorylation sites) (see the diagram in Figure 1D). As a control, cells were infected with an inducible EGFP vector. Transduced cells were cultured for 7 days in doxycycline-containing medium and then plated at clonogenic density in three-dimensional 5% Matrigel cultures (Experimental Procedures). Strikingly, cells expressing either YAP or TAZ formed solid colonies indistinguishable from those generated by MaSCs (Figures 1E and 1F) and very distinct from the cysts generated by LP cells (Figure S1D). EGFP-expressing control cells invariably remained as single cells without ever originating even a single colony in 33 experiments. As a further control, the expression of transcriptionally deficient YAPS94A (i.e., unable to interact with its DNA-binding partner TEAD) also had no effect.

We then asked whether YAP/TAZ expression converted luminal differentiated cells to a MaSC-like state. This includes the ability to form colonies that can be serially passaged. Indeed, YAP/TAZ-induced colonies, similarly to those generated from MaSCs, could form additional generations of colonies after single-cell dissociation (Figures 1G and 1H). Notably, colonies could be passaged even after expression of ectopic YAP had been turned off (by removing doxycycline) (Figures 1G, 1H and S3A). This suggests that transient expression of YAP/TAZ is sufficient to stably endow self-renewal potential to differentiated mammary cells. We thus designated the YAP/TAZ-induced “MaSC-like” cells as “yMaSCs.”

To verify whether the switch from LD to yMaSC could be recapitulated at the single-cell level, individual LD cells were seeded in 96-well plates (visually verified) and induced to express YAP. By monitoring the resulting outgrowths, we found that these individual cells formed solid colonies with high frequency (Figure S1F; 18.5% on average in the three independent experiments). From this experiment, we also noticed that this frequency of conversion, combined with the lack of colony-forming cells in controls (0%), argues against the hypothesis

that yMaSCs arise from rare, contaminating, pre-existing stem/progenitors in our LD preparations.

Of note, we also found that overexpressing YAP in the endogenous MaSC-enriched cell population does not increase its colony-forming capacity (Figure S1G). In other words, even if rare contaminant MaSCs were present, then these would remain rare and not be expanded by YAP expression.

Validation of LD-to-yMaSC Conversion by Lineage Tracing

To validate the notion that YAP expression converts differentiated cells to an SC fate, we carried out reprogramming of LD cells purified from *K8-CreERT2; R26-LSL-YFP* mice (Figure 2A), allowing for a lineage tracing strategy to genetically label luminal cells (Van Keymeulen et al., 2011). For this experiment, we first FACS-purified LD cells (as in Figure 1A). After plating, cells were exposed to a pulse of tamoxifen to activate the YFP tracer exclusively in K8-positive cells and then infected with empty or YAP-expressing vectors. Colonies generated by YAP reprogramming of LD cells were entirely YFP-positive, confirming their origin from the luminal lineage (Figures 2B and S2A). As a control, we validated that the K8-CreERT2 tracing was restricted to luminal cells. Tamoxifen-treated MaSCs from the *K8-CreERT2; R26-LSL-YFP* mammary gland formed colonies that were exclusively YFP-negative ($n = 154$, 0% YFP+) (Figure 2B). These results also argue against the possibility that YFP-labeled yMaSCs could emerge from contaminating endogenous MaSCs.

The same conclusion was further validated by a complementary experiment in which we genetically labeled the basal/MaSC-enriched cell population by using *K14-CreERT2; R26-LSL-YFP* mice (Van Keymeulen et al., 2011; Figure 2A). LD cells and MaSCs were sorted by FACS as above and then treated with a pulse of tamoxifen to label K14-positive cells. YAP-reprogrammed LD cells purified from this genetic setup generated yMaSC colonies that were invariably void of any YFP expression (0%, $n = 122$) (Figure 2C). As a positive control, endogenous MaSCs traced by *K14-CreERT2* labeling formed YFP-positive

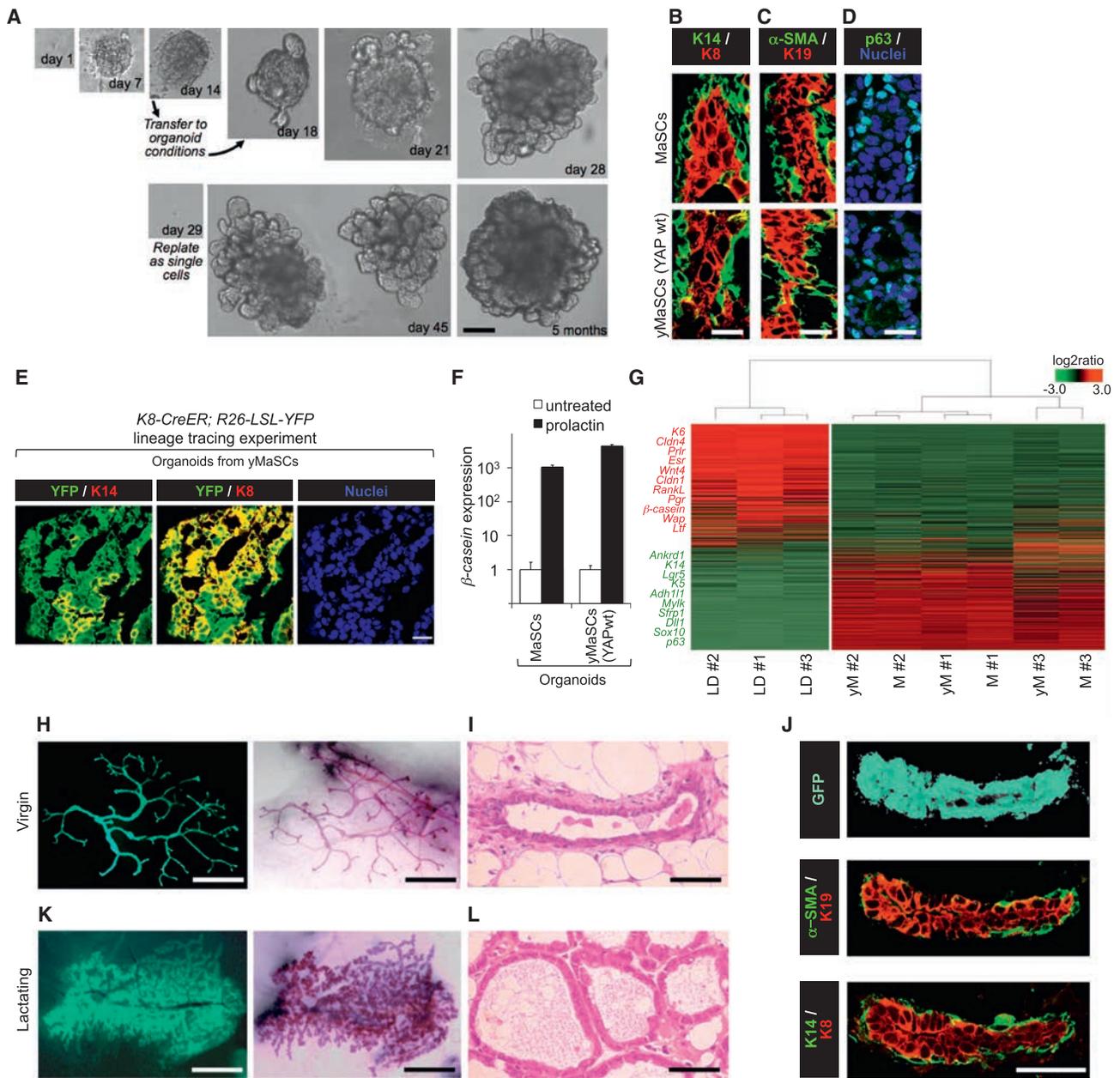


Figure 3. Characterization of yMaSCs

(A) Representative images of yMaSC outgrowths at the indicated time points. Until day 14, cultures were in mammary colony medium. After transfer to organoid conditions (see scheme in Figure 1D), 64%–75% of yMaSC colonies evolved as organoids and were maintained and passaged without doxycycline. Scale bar, 250 μ m. See also Figure S3A for transgene expression.

(B–D) Organoids from MaSCs and yMaSCs (from YAP WT) expressed basal and/or stem (α -SMA, K14, p63) and luminal markers (K8, K19). Scale bars in IF pictures, 17 μ m. See also Figures S3D–S3F for yMaSCs from phosphomutant YAP/TAZ.

(E) Immunostainings with anti-YFP combined with either anti-K14 or anti-K8 antibodies of yMaSC-derived organoids obtained from K8-CreER/R26-YFP-traced LD cells as in Figure 2A and Figure S2A. Scale bars, 49 μ m.

(F) Organoids from MaSCs and yMaSCs (from YAP WT) expressed β -casein (qRT-PCR) when treated with prolactin. Data were normalized to *Gapdh* expression and are presented as mean + SD. The results are representative of two independent experiments performed in triplicate. See also Figure S3G.

(G) Unsupervised hierarchical clustering of gene expression profiles in LD cells, organoids from MaSCs (M), and organoids from yMaSCs (yM). Each column represents one separated biological sample. Genes are ordered according to the decreasing average expression level in LD cells. Representative genes up-regulated in LD cells (red) or in MaSC- and yMaSC-derived organoids (green) are shown on the left.

(H–J) Mammary gland reconstitution generated by stably GFP-expressing yMaSCs (from YAP WT) in virgin females. (H) Whole-mount images (left, native GFP fluorescence; right, hematoxylin staining). (I) Histological section. (J) Representative sections stained for GFP and the indicated markers. See Figures S3I and S3J for controls. Scale bars, 0.5 cm in (H) and 21 μ m in (I) and (J).

(legend continued on next page)

colonies (Figure 2C). Collectively, several lines of evidence indicate that yMaSCs do not emerge from rare contaminating MaSCs pre-existing in our LD preparation. We conclude from these lineage-tracing experiments that YAP acts in differentiated cells to reprogram them into a MaSC-like state.

The Expansion, Differentiation, and Regenerative Potential of yMaSCs

We then examined whether yMaSCs are functionally equivalent to mammary SCs, as determined by additional characteristic properties of SCs, such as the ability to self-organize in vitro into mammary tissue-like structures, to differentiate along distinct lineages, and to regenerate a mammary tree in vivo after injection into a cleared mammary fat pad. For this, we sought to establish a long-term culture system that allows mammary SCs to form mammary gland-like structures in vitro. MaSC- and yMaSC-derived colonies were transferred and embedded into 100% Matrigel and overlaid with “organoid” medium (Sato et al., 2009) in the absence of doxycycline. Under these conditions, colonies underwent extensive budding and, by 2 weeks, grew into large epithelial organoids (Figure 3A; Figures S3B and S3C; see the legend of Figure 3A for quantification). Organoids derived from yMaSCs were indistinguishable in growth pattern, size, and frequency from those generated by natural MaSCs. yMaSC-derived organoids were dissociated, replated as single cells every 2 weeks, and cultured for at least 12 months (i.e., >25 passages) without changes in growth pattern, plating efficiency, and differentiation potential.

By histological examination, both MaSC and yMaSC organoids were composed by a stratified epithelium reminiscent of the histology of the mammary gland (E-cadherin-positive; Figure S3C). Internal cells surrounding a lumen-like cavity expressed differentiated luminal markers such as K8 and K19 (Figures 3B and 3C). Outer cells (either at external surfaces or bordering inner folds) displayed expression of basal, myoepithelial, and SC markers (K14, α -smooth muscle actin [α -SMA], p63) (Figures 3B–3D). yMaSCs derived from *K8-CreERT2*; *R26-LSL-YFP*-traced LD cells (as in Figure 2B) generated organoids with YFP-positive and K14-positive basal cells (Figures 3E and S2A), attesting to their origin from reprogrammed LD cells. Furthermore, addition of a lactogenic stimulus triggered expression of α - and β -casein, indicative of alveolar (milk-producing) cell differentiation (Figures 3F and S3G).

To molecularly characterize yMaSCs, we FACS-purified yMaSCs from organoids and evaluated the expression of luminal and basal and/or stem cell markers (Figure S3H). yMaSCs were remarkably similar to native MaSCs (freshly purified from the mammary gland) because they express basal markers (including myoepithelial markers such as α -Sma and *Myh11*) but not luminal markers (*Claudin1*, *K8*, *K18*, and *K19*). yMaSCs also express genes previously associated with various types of adult mammary SCs, including $\Delta Np63$, *Lgr4/5/6*, and *Procr*, and all to levels comparable with MaSCs (Chakrabarti et al., 2014; Plaks et al., 2013; Wang et al., 2015).

By unsupervised hierarchical clustering of gene expression profiles, organoids from MaSCs or yMaSCs could not be distinguished (Figure 3G). Taken together, the results indicate that, similar to authentic MaSCs, yMaSCs display self-renewal potential, generate self-organizing epithelial structures reminiscent of the normal mammary gland, and retain multilineage differentiation ability.

Next we tested whether yMaSCs displayed mammary gland reconstituting activity. For this, FACS-purified LD cells were transduced with vectors encoding for EGFP and inducible wild-type YAP. Cells were treated with doxycycline for 7 days and then transplanted (10^3 – 10^4 cells) into the cleared mammary fat pads of non-obese diabetic (NOD)-severe combined immunodeficiency (SCID) mice. Strikingly, cells that had experienced transient expression of wild-type YAP had also acquired the ability to regenerate the mammary gland (25%, $n = 16$) (Figures 3H and 3I). Ductal tree and terminal end buds were regenerated when as few as 100 YAP-reprogrammed LD cells were implanted (33%, $n = 6$). As a control, LD cells transduced with the sole EGFP vector did not display any reconstituting activity at any inoculum dose (0%, $n = 28$, 10^2 – 5×10^4) (Figure S3I). Histological analyses revealed that the epithelial outgrowths obtained from yMaSCs were EGFP-positive and morphologically indistinguishable from those generated by endogenous MaSCs and consisted of a bilayered epithelium composed of a basal and/or myoepithelial layer (positive for K14 and α -SMA) overlaid by luminal cells (positive for K8 and K19) (Figures 3J and S3J).

To explore the reconstituting potential of a single yMaSC, we injected in the cleared fat pads single-cell derived organoids and found that these were also able to regenerate the mammary gland (33%, $n = 6$). Notably, when these mice were impregnated, reconstituted mammary glands generated a dense ductal system ending in clusters of milk-secreting alveoli, indicating that yMaSCs retain full differentiation potential in vivo (Figures 3K and 3L). We conclude from this collective set of experiments that transient expression of YAP/TAZ in differentiated cells of the mammary gland is able to convert them into bona fide MaSCs.

YAP/TAZ are not only instrumental for reprogramming of differentiated mammary cells but also endogenously required in MaSCs for preserving their self-renewal potential. Indeed, we generated organoids from either endogenous MaSCs or yMaSCs obtained from *Yap^{fl/fl}*; *Taz^{fl/fl}* mice; deletion of endogenous YAP/TAZ by Adeno-Cre severely affected the ability of organoids to self-renew upon passaging (Figure S3K). Consistently, we also found that conversion to the ySC state was accompanied by activation of endogenous YAP/TAZ proteins. As shown in Figure S3L, induction of exogenous YAP in LD cells turned on expression of endogenous YAP and TAZ that remained expressed in ySC-derived organoids after ectopic YAP expression had been turned off (see Figure S3A for on-off control of exogenous YAP expression). We conclude that transient exposure to YAP/TAZ is sufficient to empower a self-sustaining loop of endogenous YAP/TAZ expression.

(K and L) Mammary gland reconstitution generated by yMaSC organoids in an impregnated female. (K) Whole-mount images (left, native GFP fluorescence; right, hematoxylin staining). (L) Histological section. Note that, upon gestation and lactation, the mammary gland is constituted by alveoli filled with milk. Scale bars, 0.5 cm in (K) and 21 μ M in (L). See also Figure S3.

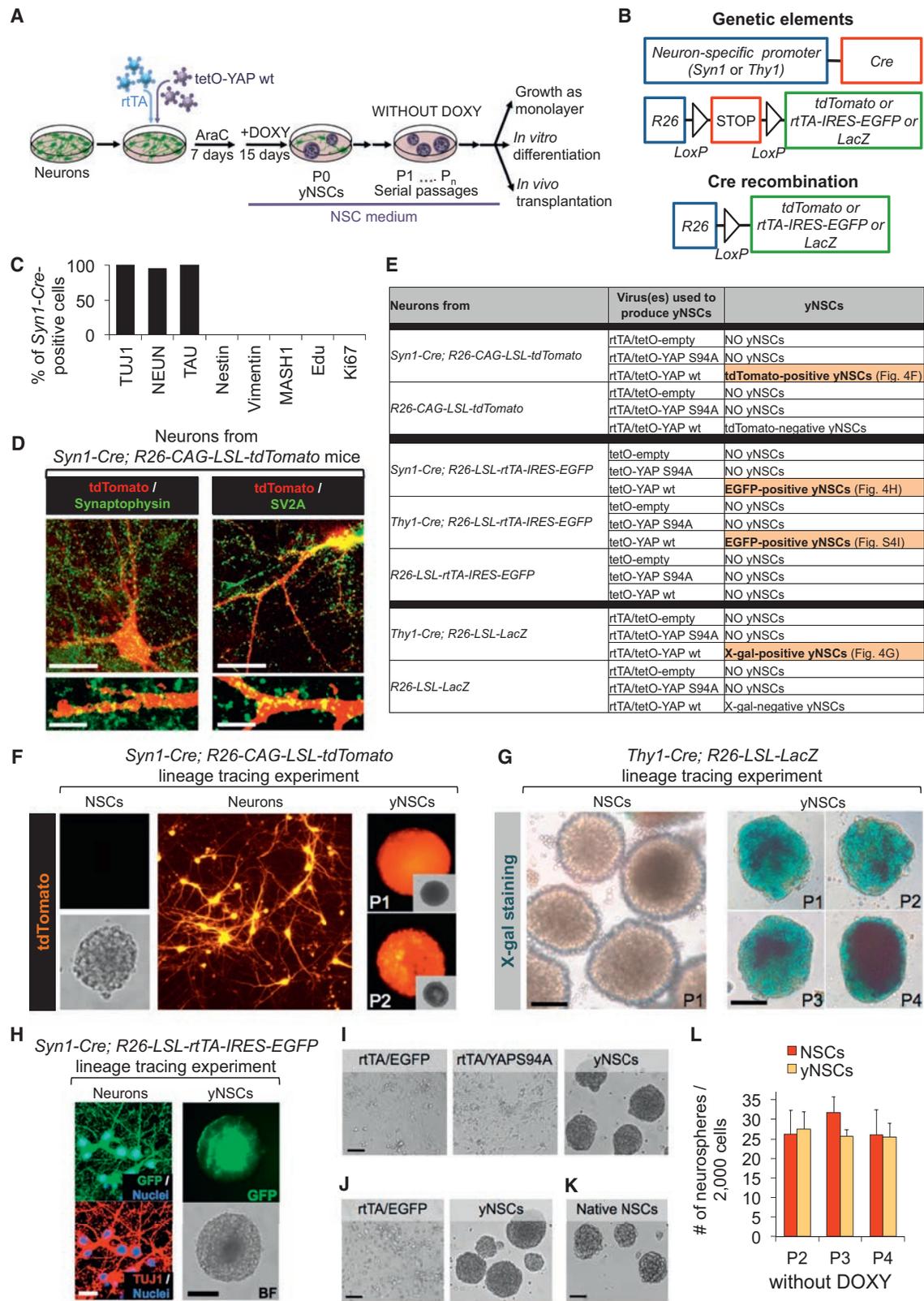


Figure 4. YAP Converts Neurons in yNSCs

(A) Schematic of the experiments performed with hippocampal or cortical neurons.
 (B) Schematic of the genetic lineage tracing strategy used to trace neurons ex vivo.

(legend continued on next page)

YAP Turns Neurons into Neural SC-like Cells

Neurons are post-mitotic cells that have been long considered refractory to any cell fate change. Endogenous YAP and TAZ proteins are highly expressed and transcriptionally active in neural stem cells (NSCs) but absent in neurons (Figures S4A–S4E), prompting us to determine whether ectopic expression of YAP in neurons was sufficient to convert them into NSCs. Primary cultures enriched in post-mitotic neurons (Figures 5A and S5A; see below) were obtained from the hippocampus or cortex from late mouse embryos (embryonic day 19 [E19]) or newborns. Cells were cultured in serum-free neuronal differentiation medium (Neurobasal with B27 containing vitamin A) also containing cytosine β -D-arabinofuranoside (AraC) for 7 days to support neuronal differentiation and eliminate proliferating cells (Figure 4A).

To follow the fate of post-mitotic neurons during YAP reprogramming, we adopted redundant lineage tracing strategies to stably label post-mitotic neurons. For this we prepared primary neurons from the brain of *Syn1-Cre* or *Thy1-Cre* transgenic mice expressing the Cre recombinase exclusively in post-mitotic neurons under the control of a neuron-specific promoter (Zhu et al., 2001; Dewachter et al., 2002) also bearing Cre-inducible tracers from the R26 locus (Figure 4B). We validated that expression of the lineage tracer was restricted to cells displaying unambiguous neuronal morphology and markers, whereas none of these traced cells was ever positive for NSCs and/or progenitor markers, as detailed in Figures 4C, 4D, and 4H and Figures S4F–S4H.

Primary cells were infected with lentiviral vectors encoding for reverse tetracycline transactivator (rtTA) and inducible wild-type YAP (Experimental Procedures). After AraC, neurons were shifted to NSC medium in the presence of doxycycline (see experimental outline in Figure 4A). Remarkably, after 2 weeks, neurosphere-like structures emerged from YAP-expressing neurons (P0 spheres) but never from neurons transduced with rtTA alone or rtTA combined with empty vectors or transcriptionally inactive YAPS94A (Figure 4E). Lineage-traced neurons from *Syn1-Cre; R26-CAG-LSL-tdTomato* or *Thy1-Cre; R26-LSL-*

LacZ transgenic mice gave rise to tdTomato-positive or β -galactosidase (β -gal)-positive neurospheres, respectively (Figures 4F and 4G). Similarly, infection with tetO-YAP lentiviruses of neurons from mice bearing *Syn1-Cre* or *Thy1-Cre* and the *R26-LSL-rtTA-IRES-EGFP* reporter generated EGFP-positive yNSCs (Figures 4H and S4G–S4I). In contrast, NSCs derived from the same strains were invariably unlabeled (Figures 4F, left panel, and 4G, left panel). This confirms that, as originally described (Zhu et al., 2001; Dewachter et al., 2002), these drivers are not active in natural NSCs; hence, YAP-induced NSCs, or “yNSCs,” originate from neurons rather than through amplification of pre-existing, contaminating, endogenous NSCs.

By monitoring the process more closely using lineage-traced neurons from *Syn1-Cre; R26-CAG-LSL-tdTomato*, we found that Doxy addition triggered YAP expression in about 30% of these neurons, and that, of these, 25% became positive for the NSC marker Nestin 2–3 days after switching them to the NSC medium (Figure S4J). This frequency of conversion, combined with the fact that no Nestin-positive neurons were present before YAP induction and no Nestin-positive neurons ever appeared in control cultures, argues against the possibility that yNSCs arose from rare pre-existing stem/progenitor cells in our neuronal preparations. This is in line with the above conclusions drawn from lineage tracing experiments.

Of note, we also found that overexpressing YAP in endogenous NSCs does not increase neurosphere-forming capacity (Figure S4K). Thus, even if rare contaminant NSCs were present, then it seems these would remain rare and not be expanded by YAP expression.

P0 spheres of yNSCs were transferred to new plates for further growth and could then be propagated for several passages as clonal outgrowths after single-cell dissociation, similarly to native NSCs (Figures 4I–4L). Lineage tracing was retained upon passaging (Figures 4F and 4G), indicating that YAP-reprogrammed neurons had acquired self-renewing properties.

The propagation of yNSCs as neurospheres did not require addition of doxycycline, indicating that transient exposure to

(C) We confirmed, in primary neuronal cultures prepared from the brain of *Syn1-Cre; R26-LSL-rtTA-IRES-EGFP* or *Syn1-Cre; R26-CAG-LSL-tdTomato* transgenic mice, that expression of the lineage tracer is indeed restricted to post-mitotic neurons expressing TUJ1 (100%, $n = 635$ cells), TAU (100%, $n = 140$ cells), and NEUN (95%, $n = 252$ cells). None of these traced cells were ever positive for the NSCs and/or progenitor markers Nestin, Vimentin, or MASH1 or for proliferation markers (5-Ethynyl-2'-deoxyridine [EdU], Ki67) (0%, $n > 4,000$ cells). See Figure S4F for representative images. See also Figure S4G for similar characterizations on *Thy1-Cre*-traced neuronal preparations.

(D) Immunofluorescence for the synaptic markers Synaptophysin and SV2A in cortical neurons derived from *Syn1-Cre; R26-CAG-LSL-tdTomato* mice. Scale bars, 11 μm (top) and 5.7 μm in magnifications (bottom).

(E) Table summarizing the results obtained in different lineage tracing experiments.

(F) Lineage tracing experiment showing that yNSCs originate from neurons. The images are bright-field and tdTomato-fluorescence pictures of primary cortical neurons from *Syn1-Cre; R26-CAG-LSL-tdTomato* mice and yNSCs derived from them (at the indicated passages). Neurospheres from *Syn1-Cre; R26-CAG-LSL-tdTomato* NSCs are presented as a negative control.

(G) Panels are X-gal stainings of yNSCs derived from hippocampal neurons from *Thy1-Cre; R26-LSL-LacZ* mice (scale bars, 210 μm) at the indicated passages. Neurospheres from *Thy1-Cre; R26-LSL-LacZ* NSCs (scale bar, 210 μm) are presented as a negative control.

(H) Lineage tracing experiment with the *Syn1-Cre* driver showing that yNSCs originate from Syn1-traced neurons. Left: immunostaining for GFP and TUJ1 in cortical neurons obtained from *Syn1-Cre; R26-LSL-rtTA-IRES-EGFP* mice. Right: bright-field and GFP fluorescence pictures of neurospheres derived from yNSCs obtained from the same neurons after transduction with doxycycline-inducible YAP. See also Figures S4G–S4I for similar results with the alternative lineage tracing strategy using *Thy1-Cre; R26-LSL-rtTA-IRES-EGFP* mice.

(I–K) Representative images of yNSCs neurospheres (second passage, P2) derived from hippocampal (I) or cortical (J) neurons. Images from negative control transduced neurons are shown as a reference (I and J). Neurospheres from native NSCs are presented for comparison (K). Scale bars, 210 μm .

(L) P1 yNSCs were dissociated to single cells and replated at clonal density for neurosphere formation in the absence of doxycycline for further passages (P2, P3, and P4). Native NSCs are presented for comparison. Graphs are quantifications of neurospheres formed by the indicated cells. Results are representative of at least eight (P2), six (P3), and three (P4) independent experiments performed in six replicates. Data are presented as mean + SD.

See also Figure S4.

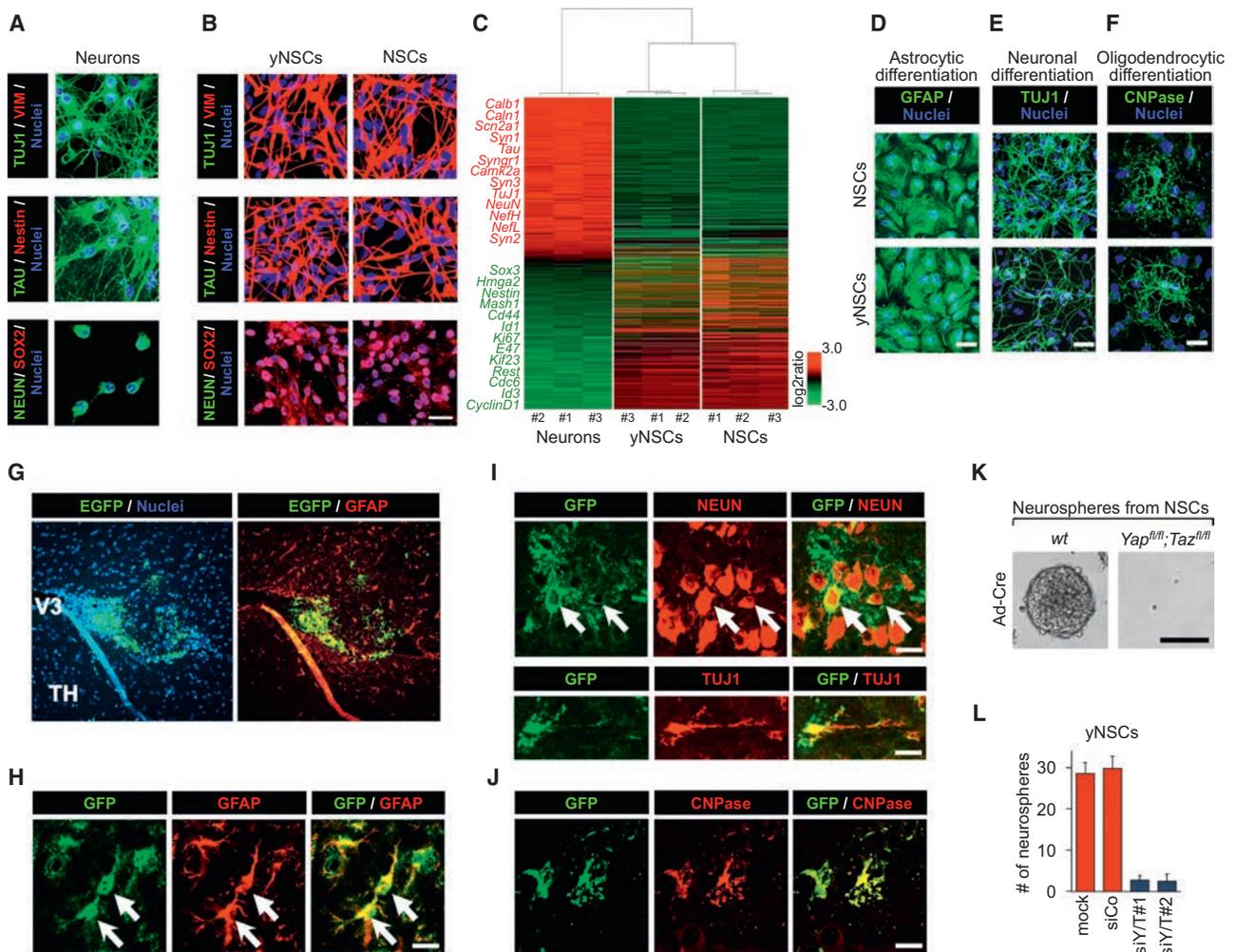


Figure 5. Characterization of yNSCs

(A and B) Immunofluorescence for the indicated markers in hippocampal neurons (A) and P3 yNSCs plated as a monolayer (B). Endogenous NSCs served as a positive control (scale bar, 23 μ m). As shown in Figure S5B, Nestin-, SOX2-, and Vimentin-positive yNSCs derived from *Thy1-Cre*; *R26-LSL-rtTA-IRES-EGFP* neuronal preparations are co-stained with EGFP, confirming their origin from differentiated cells.

(C) Unsupervised hierarchical clustering of gene expression profiles in cortical neurons, yNSCs, and NSCs. Each column represents one separated biological sample. Genes are ordered according to the decreasing average expression level in neurons. Representative genes upregulated in neurons (red) or in NSCs and yNSCs (green) are shown on the left.

(D–F) YAP-induced yNSCs and endogenous NSCs as a positive control were plated and differentiated toward an astrocytic, neuronal, or oligodendrocytic fate (Experimental Procedures). Confocal images showing the astrocytic marker GFAP (D), the neuronal differentiation marker TUJ1 (E), and the oligodendrocytic marker CNPase (F) are displayed. The results are representative of three independent experiments performed in triplicate. Scale bars, 50 μ m. For neuronal differentiation, see also Figures S5C–S5E, attesting acquisition of TAU and loss of Nestin expression.

(G–J) yNSCs were transduced with a constitutive EGFP-expressing vector and injected into the brains of recipient mice. Four weeks later, the brains were fixed and processed for immunofluorescence analyses. (G and H) Representative confocal images showing that injected yNSCs (GFP-positive) integrate in the brain parenchyma and massively differentiate into an astrocytic fate, as indicated by GFAP expression (arrows). V3, third ventricle; TH, thalamus. (I and J) Representative confocal images showing injected yNSCs (GFP-positive) differentiated in NEUN- and TUJ1-positive neurons (I, arrows) or CNPase-positive oligodendrocytes (J). Scale bars, 19 μ m.

(K) Representative images of neurospheres from WT or *Yap^{fl/fl}*; *Taz^{fl/fl}* NSCs transduced with Ad-Cre. Scale bar, 250 μ m.

(L) yNSCs (passage 4 as neurospheres) were dissociated, plated on fibronectin-coated dishes, and transfected with the indicated siRNAs. The graph represents the quantification (mean \pm SD) of neurospheres derived from the indicated cells.

See also Figure S5.

exogenous YAP is sufficient to induce self-renewal properties that are autonomously maintained. In line, as shown by experiments with a Cre-excisable tetO-YAP lentiviral vector, post-reprogramming deletion of the whole YAP-encoding viral cassette had no effects on yNSC maintenance (Figures S4L–S4N).

Characterization of yNSCs

Next we characterized yNSCs by marker gene expression using immunofluorescence, qRT-PCR, and gene profiling. As shown in Figures 5A and 5B and Figures S5A and S5B, yNSCs completely lost expression of the terminal differentiation markers present in

the original neurons (such as TUJ1, TAU, and NEUN) and, instead, expressed high levels of NSC markers (Nestin, SOX2, and Vimentin). Furthermore, we compared the transcriptome of parental neurons, yNSCs, and native NSCs and found that yNSCs completely lost their neuronal identity and acquired a gene expression profile closely similar to native NSCs (Figure 5C).

Neural SCs are defined as tripotent, as attested to by their ability to differentiate in astrocytes, neurons, and oligodendrocytes. We therefore examined the differentiation potential of yNSCs by placing them under appropriate culture conditions (Experimental Procedures). yNSCs could differentiate into astrocytes, neurons, and oligodendrocytes as defined by markers and morphology (Figures 5D–5F and S5C–S5E) and thus are indeed tripotent. To investigate the *in vivo* differentiation potential of yNSCs, we transplanted EGFP-labeled yNSCs in the brain of newborn mice ($n = 5$). Four weeks after transplantation, grafted yNSCs invariably lost Nestin positivity (Figure S5F) and mainly remained close to the injection site, where they primarily acquired expression of GFAP, indicative of astrocyte differentiation (Figures 5G and 5H). Injected yNSCs also differentiated into NEUN- and TUJ1-positive neurons or CNPase-positive oligodendrocytes (Figures 5I and 5J). Importantly, no tumor formation was ever observed after histological examination of the yNSCs-injected brain parenchyma. Thus, YAP induces conversion of neurons into cells that have functional properties similar to those of normal NSCs.

As for MaSCs, endogenous YAP/TAZ are essential to sustain the expansion of native NSCs *in vitro* because *ex-vivo* Adeno-Cre-mediated deletion of YAP/TAZ from *Yap^{fl/fl}*; *Taz^{fl/fl}* NSCs blunted neurosphere formation (Figure 5K). We further established that the self-renewal properties of yNSCs are also sustained by reactivation of endogenous YAP/TAZ. Two lines of evidence support this conclusion. First, endogenous TAZ is induced in yNSCs and remains as such after doxycycline withdrawal and Cre-mediated excision of the tetO-YAP cassette (Figure S5G). Second, YAP/TAZ depletion in yNSCs by transfecting independent pairs of small interfering RNAs (siRNAs) greatly impairs their self-renewal properties (Figure 5L). These results raise an interesting parallel between the requirement of YAP/TAZ in native NSCs and induced yNSCs.

Ex Vivo Generation of Pancreatic Progenitors from Exocrine Cells

Pancreatic progenitors are rare in the normal pancreas but can be regenerated by differentiated acinar exocrine cells upon injury (Pan et al., 2013). Pancreatic progenitors are expandable *in vitro* as ductal organoids (Huch et al., 2013; Figures 6A and 6B), and, like MaSCs and NSCs, display nuclear and transcriptionally active YAP/TAZ and genetically require YAP/TAZ for their propagation (Figures S6A–S6C). Considering the intrinsic plasticity of acinar cells, we used them as a third reprogramming paradigm, asking whether a transient pulse of YAP expression could be sufficient to convert them into progenitors. To this end, we isolated pancreatic acini from *R26-rtTA*; *tetOYAP^{S127A}* adult mice and dissociated them to obtain a single cell preparation. Cells were plated in 100% Matrigel and cultured in the presence of doxycycline in pancreas organoid medium. In just a few days, acinar cells induced to express YAP, but not cells left without doxycy-

cline, expanded as cyst-like organoids (Figures 6C and S6D). Acinar cells derived from control *R26-rtTA* mice remained as single cells or, more rarely, formed small cysts but never expandable organoids (Figure S6D). After initial derivation, YAP-induced organoids (or “yDucts”) could be passaged for several months even in the absence of doxycycline and, thus, in the absence of exogenous YAP/TAZ (for at least 10 passages, 6 months). Individual organoids could be manually picked and expanded as clonal lines. By morphology, size, and growth pattern, organoids derived from converted acinar cells were comparable with those obtained from handpicked pancreatic duct fragments after whole pancreas dissociation (Huch et al., 2013; Figures 6B and 6D).

As an alternative strategy, we embedded whole pancreatic acini explanted from *R26-rtTA*; *tetOYAP^{S127A}* in collagen and cultured them under low-serum conditions known to preserve acinar cell identity *ex vivo* (Means et al., 2005). When treated with doxycycline to induce YAP expression (Figure S6E), pancreatic acini converted within a few days to ductal organoid structures and with high efficiency (>70%) (Movies S1 and S2; Figures S6F–S6H). As a control, acini lacking exogenous YAP expression (e.g., left without doxycycline; Figure S6H) remained as such and never converted to organoids. After transfer to 100% Matrigel-pancreatic organoid medium, the YAP-induced ducts, but not control acini, regrew into organoids and could be maintained for several passages after single-cell dissociation even in the absence of doxycycline (Figure S6G).

To validate that yDucts were indeed derived from differentiated exocrine acinar cells, we carried out genetic lineage tracing experiments using *Ptf1a-CreERTM*; *R26-LSL-rtTA-IRES-EGFP*; *tetO-YAP^{S127A}* mice (Figure S6I). In this genetic background, tamoxifen treatment of adult mice causes irreversible genetic tracing exclusively of pancreatic acinar cells (but not of endocrine, ductal, or centroacinar cells), as reported previously by others (Pan et al., 2013) and revalidated here (Figure S6J). After treatment, mice were kept without tamoxifen for 1 week, and then pancreata were explanted to prepare whole acini or single acinar cells that were cultured as above (see experimental outline in Figure S6I). These EGFP-positive cells never formed any organoid in the absence of doxycycline (Figures S6K and S6L). Instead, doxycycline-induced YAP expression caused the formation of expandable yDucts, that retained EGFP positivity over passaging, formally demonstrating their derivation from terminally differentiated exocrine cells (Figures 6E, 6F, S6M, and S6N; Movie S3). As a control of driver specificity in our culture conditions, organoids derived from endogenous ductal progenitors explanted from *Ptf1a-CreERTM*; *R26-LSL-rtTA-IRES-EGFP*; *tetO-YAP^{S127A}* pancreata were never labeled by EGFP (Figure S6O). This indicates that EGFP-traced yDucts emerge from exocrine cells and not from pre-existing ductal progenitors. In line with this conclusion, YAP overexpression is inconsequential in endogenous ductal progenitors for their clonogenic and organoid forming ability (Figure S6P), making unlikely the possibility that YAP expression might expand rare pre-existing contaminants.

We also performed a time course analysis of gene expression dynamics occurring at the single-cell level during reprogramming induced by tetO-Yap. As shown in Figure 6G, at the beginning of the experiment, cells expressed exocrine markers (*Amy*

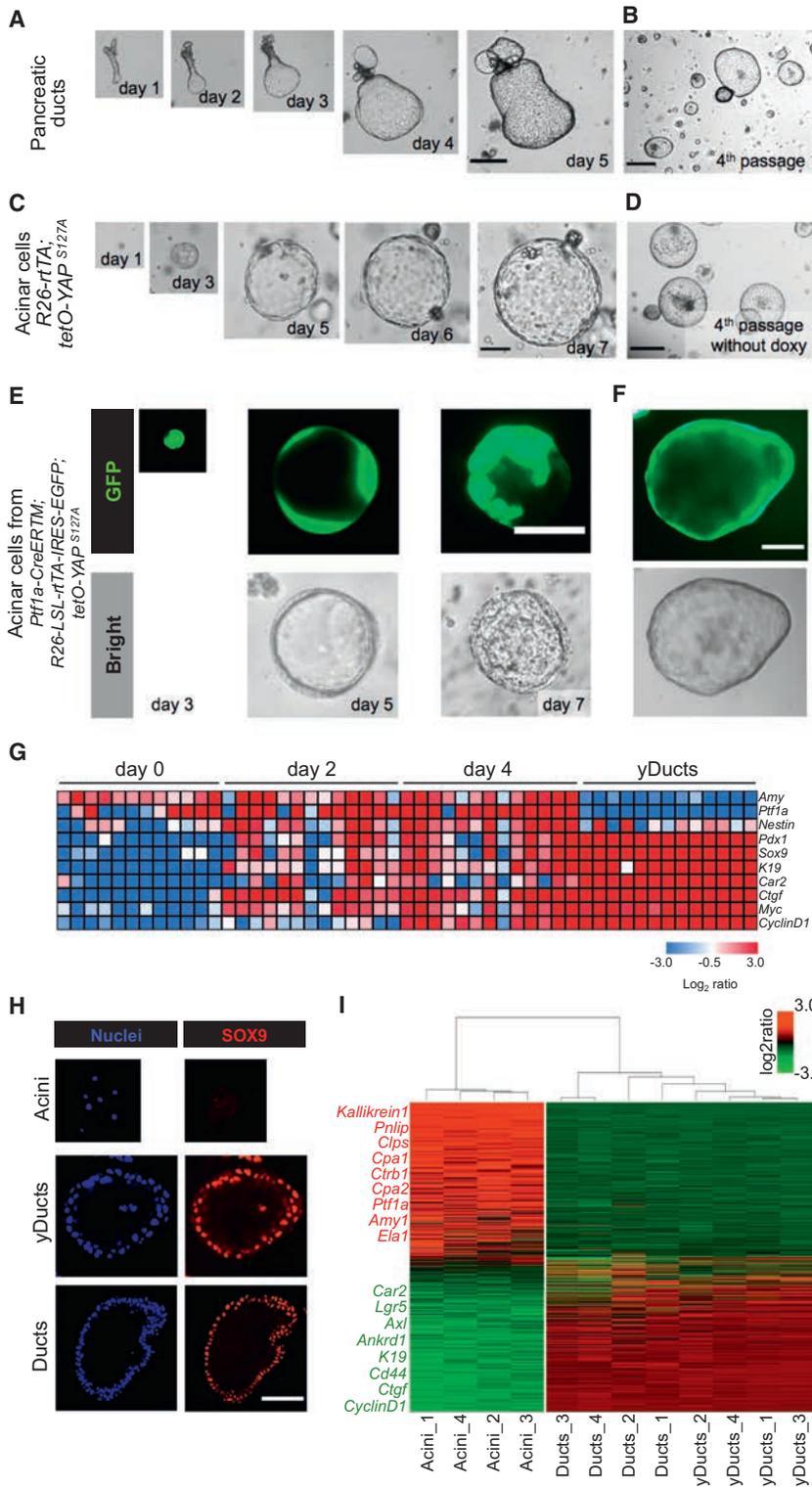


Figure 6. YAP Converts Pancreatic Acinar Cells to Duct-like Organoids

(A and B) Representative images of a pancreatic duct fragment growing in pancreatic organoid medium at the indicated times (A) and after four passages in fresh Matrigel (B). Pictures are representative of three independent experiments performed with four technical replicates. Scale bars, 290 μ m.

(C and D) Serial images of a single acinar cell derived from *R26-rtTA; tetO-YAP^{S127A}* growing as cyst-like organoids at the indicated time points after Doxy addition (C) and after four passages in fresh Matrigel in the absence of Doxy (D). Pictures are representative of five independent experiments performed with four technical replicates. Scale bars, 70 μ m in (C) and 290 μ m in (D).

(E and F) Lineage tracing experiments using the *Ptf1a-CreERTM* driver. Images are bright-field and GFP fluorescence pictures of transgenic YAP-expressing exocrine cells at the indicated time points of Doxy treatment (E) and after passaging in absence of Doxy (F). See also Figure S6f for a schematic of the experiment. Scale bars, 70 μ m in (E) and 130 μ m in (F).

(G) Single-cell gene expression profile of pancreatic cultures of the *R26-rtTA; tetO-YAP^{S127A}* genotype during the YAP-induced conversion of acinar cells to yDucts. Rows are evaluated genes, and columns are individual cells. Day 0, starting acini (without doxycycline); day 2, cultures that experienced 48 hr of doxycycline; day 4, cultures that experienced 96 hr of doxycycline. The heatmap represents expression levels as log₂ ratio normalized to 18S rRNA.

(H) Organoids from duct fragments (Ducts, bottom, as in B) and YAP-induced organoids (yDucts, center) expressed the ductal marker SOX9 and were negative for the exocrine marker Amylase (data not shown) by immunofluorescence. Acinar cells (top) are shown as a control. Scale bar, 80 μ m.

(I) Unsupervised hierarchical clustering of gene expression profiles in acini, yDucts, and Ducts. Each column represents one separated biological sample. Genes are ordered according to the decreasing average expression level in acini. Representative genes upregulated in acinar cells (red) or in Ducts and yDucts (green) are shown on the left.

See also Figure S6.

and *Ptf1a*) but not markers of pancreatic progenitors or ductal/centroacinar cells (*Pdx1*, *Sox9*, *K19*, and *Car2*), cell proliferation (*CyclinD1*), or the YAP targets *Ctgf* and *Myc* (Zanconato et al., 2015). After 2 days of doxycycline treatment, pancreatic progenitor markers were turned on in most cells and did so more

robustly on day 4 and then in yDuct cells (in line with what is shown in Movies S1, S2, and S3). On days 2 and 4 (but not in yDucts), most acinar cells retained concomitant expression of acinar markers and could thus be considered cells caught in transition. We noticed that progenitor markers were already present on day 2,

before cells acquired expression of the proliferation marker *CyclinD1*, indicating that phenotypic conversion into a ductal progenitor state can be initially uncoupled from proliferation.

In section, yDuct-derived organoids appeared as epithelial monolayers surrounding a central cavity (Figure 6H). By

qRT-PCR and immunofluorescence, organoids lost markers of exocrine differentiation (*Ptf1a*, α -*amylase*, *Elastase*, and *CPA1*) and acquired expression of ductal markers (*K19*, *Sox9*, *Hes1*, and *Cd44*) and proliferative markers (*cMyc* and *CyclinD1*), all to levels comparable with those of native ductal organoids (Figures 6H and S6Q). To determine the extent of YAP-induced conversion of acinar cells and their molecular overlap with native ductal progenitors, we carried out transcriptomic analyses. As shown in Figure 6I, yDucts diverged from parental acinar cells to become ostensibly similar to bona fide pancreatic progenitors. Under differentiating conditions, yDuct-derived cells could be induced to re-express the differentiated exocrine marker CPA1 and to downregulate K19 (Figure S6R). When transplanted into the pancreas of NOD-SCID mice, yDucts remained as such and never formed any tumor ($n = 6$, data not shown), indicating that yDucts are indeed non-transformed and non-tumorigenic. Together, the results indicate that exocrine cells with a history of exposure to YAP acquired key molecular and biological features of ductal pancreatic progenitors.

DISCUSSION

Here we report that expression of a single factor into differentiated cells explanted from distinct tissues induces cells with functional and molecular attributes of their corresponding tissue-specific SCs that can be expanded *ex vivo*. The ySC state can be transmitted through cell generations without the need for continuous expression of ectopic YAP/TAZ, indicating that a transient activation of ectopic YAP or TAZ is sufficient to induce a heritable self-renewing state. Differently from induced pluripotent stem cells (iPSCs) or other reprogramming efforts, ySCs preserve a memory of the tissue of origin, expanding the current reprogramming paradigms by focusing on somatic stem cell generation from related cells of the same lineage.

Several lines of evidence support the notion that ySCs originate from conversion of differentiated cells rather than from amplification of rare pre-existing endogenous SCs. In particular, we used lineage tracing strategies employing established Cre drivers to label differentiated cells and follow their fate after YAP-induction. We found that ySCs indeed retained the genetic label specific of the original differentiated cells. Conversely, colonies and organoids emerging from native SCs were invariably unlabeled by the same genetic tracers. In the case of the mammary gland, we also used FACS to sort terminally differentiated cells from luminal progenitors and show that YAP can effectively operate on LD cells. Beyond lineage tracing, ySCs were induced with relatively high frequency after YAP expression, whereas no outgrowths emerged from differentiated cells expressing control vectors or transcriptionally inactive YAP, a scenario that argues against ySCs emergence from rare, pre-existing SCs. We also entertained the possibility that YAP overexpression may selectively expand rare native stem/progenitor cells. We consider this possibility also unlikely because YAP expression has no effect on the colony-forming capacity and expandability of native SCs. The latter finding is perhaps consistent with the fact that native SCs cultured *ex vivo* already contain transcriptionally active endogenous YAP/TAZ.

It is worth noting that, in mammary gland and pancreatic acinar cells, we were able to obtain transition to the correspond-

ing tissue SCs starting from mature, adult differentiated cells, highlighting how YAP can imbue these lineages with a remarkable plasticity. However, we carried out our reprogramming experiments on fetal neurons because primary adult neurons cannot be effectively cultured *ex vivo*. Although they are post-mitotic, these early neurons may be particularly competent for YAP-induced reprogramming. Future work and technological advancements will be required to determine whether adult neurons can be reprogrammed by YAP similarly to fetal neurons.

Our procedure generates cells with normal SC traits as suggested by several lines of evidence: ySCs can be expanded over multiple passages as self-expanding organoids or neurospheres; ySCs readily generate a multilineage progeny reminiscent of the corresponding native SCs (for example, in the case of yMaSCs, reprogrammed SCs generate minigland-like organoids *in vitro* that, when transplanted *in vivo*, regenerate a normal ductal tree in the cleared fat pad); ySCs are not transformed and non-tumorigenic; and, at the transcriptional level, ySCs display remarkable overlaps with their native counterparts.

Lineage plasticity and reversion to an SC-like status rarely occur in normal tissues but are associated with tissue repair or oncogenic activation (Blanpain and Fuchs, 2014). Of note, genetic depletion of YAP and/or TAZ in several adult epithelia is inconsequential for normal homeostasis but, in fact, essential for regeneration, tumor growth (Zanconato et al., 2016), and, as shown here, for expansion of somatic SCs *in vitro*. It is thus tempting to propose that the path for lineage-specific reprogramming outlined here may be activated whenever natural, pathological, or *ex vivo* conditions demand generation and expansion of new SCs using differentiated cells as facultative SCs without losing tissue memory. Further work is required to validate that YAP reprogramming can occur *in vivo* and particularly in humans. If it does, it may be worth exploring means to exploit this path to facilitate tissue repair and regeneration in distinct tissues. From this perspective, the inability to cross lineage boundaries might represent a potential limitation for *in vitro* applications because some differentiated cells may not be readily accessible. However, if the same strategy could be applied directly *in vivo*, then the ability of a single factor to generate proliferative and multipotent SCs while retaining tissue memory would provide an advantage over other reprogramming and transdifferentiation strategies that employ either complex cocktails of transcription factors or the passing through an embryonic-like pluripotent state (Bar-Nur et al., 2015; Xu et al., 2015).

In conclusion, finding that YAP/TAZ, as a single factor, can reprogram distinct cell types into their corresponding tissue-specific SCs may have implications for regenerative medicine, for discovering still unknown determinants of somatic stemness, and, more broadly, for ad hoc expansion of somatic cells.

EXPERIMENTAL PROCEDURES

Primary Mammary Epithelial Cell Isolation and Induction of yMaSCs

Primary mammary epithelial cells (MECs) were isolated from the mammary glands of 8- to 12-week-old virgin C57BL/6J mice (unless otherwise specified) according to standard procedures (Stingl et al., 2006; see Supplemental Experimental Procedures for details). Animal experiments were performed adhering to our institutional guidelines, and approved by OPBA and the Ministry of Health. To separate various MEC subpopulations, cells were stained for

30 min at 4°C with antibodies against CD49f (phycoerythrin [PE]-Cy5, catalog no. 551129, BD Biosciences), CD29 (PE-Cy7, catalog no. 102222, BioLegend), CD61 (PE, catalog no. 553347, BD Biosciences), EpCAM (fluorescein isothiocyanate [FITC], catalog no. 118208, BioLegend), and lineage markers (allophycocyanin [APC] mouse lineage antibody cocktail, catalog no. 51-9003632, BD Biosciences) in DMEM/F12.

The stained cells were then resuspended in PBS/BSA 0.1% and sorted on a BD FACS Aria sorter (BD Biosciences) into LD cells, LP cells, and MaSCs.

Primary sorted subpopulations from FACS were plated on collagen I-coated supports and cultured in two dimensions in mammary gland (MG) medium (DMEM/F12 supplemented with glutamine, antibiotics, 10 ng/ml murine epidermal growth factor [EGF], 10 ng/ml murine basic Fibroblast growth factor (bFGF), and 4 µg/ml heparin with 2% fetal bovine serum [FBS]).

For induction of yMaSCs, LD cells were transduced for 48 hr with FUW-tetO-YAP or FUW-tetO-TAZ in combination with rTA-encoding lentiviruses. As a (negative) control, LD cells were transduced with either FUW-tetO-EGFP (Figures 1E, 1F and S1F) in combination with rTA-encoding lentiviruses. After infection, adherent cells were washed and treated with 2 µg/ml doxycycline for 7 days in MG medium for activating tetracycline-inducible gene expression (see scheme in Figure 1D) to obtain yMaSCs. After doxycycline treatment for 7 days in two-dimensional culture, yMaSCs were processed for further assays or analysis. Unless otherwise specified, yMaSCs were generated from wild-type YAP (FUW-tetO-WTYAP). For the experiment depicted in Figures 2B, 2C and S2A, we first FACS-purified LD cells and MaSC-enriched populations (using CD61 and CD49f as described previously for Figure 1A) from *K8-CreERT2; R26-LSL-YFP/+* or *K14-CreERT2; R26-LSL-YFP/+* virgin female mice. These cells were plated, and, after attachment, they were treated with 1 µM 4-hydroxy (4OH)-tamoxifen for 24 hr. Cells were then transduced for 48 hr with FUW-tetO-WTYAP in combination with stable rTA-encoding lentiviral supernatant. Negative control cells were provided by LD cells transduced with FUW-tetO-MCS (empty vector) in combination with rTA-encoding lentiviral supernatants. After infection, cells were washed and treated with doxycycline in MG medium as above.

Primary Neuron Isolation and Induction of yNSCs

Neurons were prepared from hippocampi or cortices of late (E18–19) embryos or post-natal day 1 (P1) pups as described previously (Han et al., 2008). Briefly, hippocampi and cortices were dissected under the microscope in ice-cold Hank's balanced salt solution (HBSS) as quickly as possible, incubated with 0.05% trypsin (Life Technologies) for 15 min at 37°C, and, after trypsin blocking, resuspended in DMEM/10% FBS supplemented with 0.1 mg/ml DNase I (Roche) and mechanically dissociated by extensive pipetting. Cells were then plated on poly-L-lysine-coated wells in DMEM supplemented with 10% FBS, glutamine, and antibiotics for hippocampal neurons or in DMEM/Neurobasal (1:1) supplemented with 5% FBS, 1 × B27, glutamine, and antibiotics for cortical neurons (day 1). After 24 hr (day 2), the medium of both hippocampal and cortical preparations was changed to fresh DMEM/Neurobasal (1:1) supplemented with 5% FBS, 1 × B27, glutamine, and antibiotics. For reprogramming experiments, neurons were infected on the following day (day 3) with FUW-tetO-WTYAP and FUDeltaGW-rTA viral supernatants. Negative controls were provided by neurons transduced with FUDeltaGW-rTA alone or in combination with FUW-tetO-EGFP or FUW-tetO-MCS (empty vector). After 24 hr (day 4), the medium was changed, and cells were incubated in Neurobasal medium supplemented with 1 × B27, glutamine, antibiotics, and 5 µM AraC (Sigma) for an additional 7 days, at the end of which well differentiated, complex network-forming neurons were visible. To induce yNSCs formation, treated neurons were switched to NSC medium (DMEM/F12 supplemented with 1 × N2, 20 ng/ml murine EGF, 20 ng/ml murine bFGF, glutamine, and antibiotics) and doxycycline for activating tetracycline-inducible gene expression. After 7 days, fresh doxycycline was added. Sphere formation was evident upon YAP induction after 14 days of doxycycline treatment. Spheres (P0 spheres) were mechanically detached from the plates by tapping and flushing (not by trypsinizing), transferred into a 15-ml plastic tube, and left to sediment (usually 5 min). After discarding the supernatant, spheres were dissociated with 1–2 ml of TrypLE Express (Life Technologies) and mechanical pipetting. TrypLE Express was then diluted 1:5 in NSC medium, and cells were centrifuged and resuspended in NSC medium without doxycycline. For the successive passages, spheres were harvested and dissociated, and yNSCs were

routinely cultured and passaged without doxycycline in NSC medium as for normal NSCs.

Pancreatic Acinar Cell Isolation and Induction of yDucts

Primary pancreatic acini were isolated from the pancreata of 6- to 9-week-old mice according to standard procedures (Means et al., 2005). Digested tissue was filtered through a 100-µm nylon cell strainer. The quality of isolated acinar tissue was checked under the microscope. For culture of entire acini, explants were seeded in neutralized rat tail collagen type I (Cultrex)/acinar culture medium (1:1) (Means et al., 2005), overlaid with acinar culture medium (Waymouth's medium [Life Technologies] supplemented with 0.1% FBS [Life Technologies], 0.1% BSA, 0.2 mg/ml soybean trypsin inhibitor [SBTI], 1 × insulin-transferrin-selenium-ethanolamine [ITS-X] [Life Technologies], 50 µg/ml bovine pituitary extract [BPE] [Life Technologies], 1 µg/ml dexamethasone [Sigma], and antibiotics) when collagen formed a gel. For culture of isolated acinar cells, acini were further digested in 0.05% trypsin for 30 min at 37°C to obtain a single-cell suspension. Single acinar cells were plated in 100% Matrigel. When Matrigel formed a gel, cells were supplemented with pancreatic organoid medium (advanced DMEM/F12 supplemented with 1 × B27, 1.25 mM N-acetylcysteine, 10 nM gastrin, 50 ng/ml murine EGF, 100 ng/ml human Noggin, 100 ng/ml human FGF10, 10 mM nicotinamide, 1 µg/ml R-Spondin1, and antibiotics) supplemented with 0.2 mg/ml SBTI. For induction of pancreatic organoids, entire acini or single acinar cells of the indicated genotypes cells were seeded in medium supplemented with 2 µg/ml doxycycline. Negative control cells were cultured under the same conditions in the absence of doxycycline. Cells were treated with 2 µg/ml doxycycline for 7 days, and organoid formation was morphologically followed.

ACCESSION NUMBERS

The accession number for the Affymetrix data reported in this paper is GEO: GSE70174.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and three movies and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.08.009>.

AUTHOR CONTRIBUTIONS

T.P. carried out the experiments on neurons and exocrine pancreas; A.F., on neurons; and L.A., on the mammary gland cells. C.F. and G.B. performed the sorting experiments. S. Bresolin, G.B., and S. Bicciato performed microarrays and biostatistics. D.D.B. performed molecular biology and IFS. A.R. performed mouse surgery and helped with transplantation. S.P. conceived the initial hypothesis and experimental design. M.C. and S.P. planned, discussed, and organized the work. S.P., M.C., L.A., T.P., and A.F. wrote the manuscript.

ACKNOWLEDGMENTS

We thank O. Wessely, G. Martello, S. Dupont, and F. Zanconato for comments; V. Guzzardo for histology; C. Mucignat for advice on brain stainings; A. Migliorati, P. Raffa, and V. Di Iorio for help with live cell imaging; D.J. Pan, F. Camargo, C. Blanpain, G. Carmignoto, P. Chambon, D. Merkler, I. De Curtis, and R. Brambilla for gifts of mice; and L. Naldini for plasmids. We are indebted to P.G. Pellici and A. Santoro for advice on mammary fat pad transplantation, M. Forcato for help with bioinformatic analyses, and V. Broccoli and A. Sessa for advice on NSC transplantation. *R26-rTA*, *R26-LSL-rTA-IRES-EGFP*, and *Ptf1a-CreERTM* mice were purchased from The Jackson Laboratory, where they were deposited by R. Jaenisch, A. Nagy, and C. Wright, respectively. L.A. is supported by a fellowship from the Italian Association for Cancer Research (AIRC). A.F. is the recipient of an Astellas Foundation for Research on Metabolic Disorders fellowship and Japan Society for the Promotion of Science fellowship. This work is supported by a Fondazione Città della Speranza grant and CARIPARO Bando Ricerca Pediatrica (to G.B.), by AIRC Special Program Molecular Clinical Oncology “5 per mille” and an AIRC PI grant (to S.P.),

and by Epigenetics Flagship project CNR-Miur grants (to S.P.). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (Grant Agreement No. 670126-DENOVOSTEM).

Received: March 15, 2016

Revised: July 13, 2016

Accepted: August 12, 2016

Published: September 15, 2016

REFERENCES

- Azzolin, L., Panciera, T., Soligo, S., Enzo, E., Bicciato, S., Dupont, S., Bresolin, S., Frasson, C., Basso, G., Guzzardo, V., et al. (2014). YAP/TAZ incorporation in the β -catenin destruction complex orchestrates the Wnt response. *Cell* **158**, 157–170.
- Bar-Nur, O., Verheul, C., Sommer, A.G., Brumbaugh, J., Schwarz, B.A., Lipchina, I., Huebner, A.J., Mostoslavsky, G., and Hochedlinger, K. (2015). Lineage conversion induced by pluripotency factors involves transient passage through an iPSC stage. *Nat. Biotechnol.* **33**, 761–768.
- Blanpain, C., and Fuchs, E. (2014). Stem cell plasticity. Plasticity of epithelial stem cells in tissue regeneration. *Science* **344**, 1242281.
- Chakrabarti, R., Wei, Y., Hwang, J., Hang, X., Andres Blanco, M., Choudhury, A., Tiede, B., Romano, R.A., DeCoste, C., Mercatali, L., et al. (2014). DeltaNp63 promotes stem cell activity in mammary gland development and basal-like breast cancer by enhancing Fzd7 expression and Wnt signalling. *Nat. Cell Biol.* **16**, 1004–1015.
- Cordenonsi, M., Zanconato, F., Azzolin, L., Forcato, M., Rosato, A., Frasson, C., Inui, M., Montagner, M., Parenti, A.R., Poletti, A., et al. (2011). The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell* **147**, 759–772.
- Dewachter, I., Reversé, D., Caluwaerts, N., Ris, L., Kuipéri, C., Van den Haute, C., Spittaels, K., Umans, L., Serneels, L., Thiry, E., et al. (2002). Neuronal deficiency of presenilin 1 inhibits amyloid plaque formation and corrects hippocampal long-term potentiation but not a cognitive defect of amyloid precursor protein [V717I] transgenic mice. *J. Neurosci.* **22**, 3445–3453.
- Fitamant, J., Kottakis, F., Benhamouche, S., Tian, H.S., Chuvin, N., Parachoniak, C.A., Nagle, J.M., Perera, R.M., Lapouge, M., Deshpande, V., et al. (2015). YAP Inhibition Restores Hepatocyte Differentiation in Advanced HCC, Leading to Tumor Regression. *Cell Rep.* **10**, 1692–1707.
- Guo, W., Keckesova, Z., Donaher, J.L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zürrer-Härdi, U., Bell, G., et al. (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* **148**, 1015–1028.
- Han, X.J., Lu, Y.F., Li, S.A., Kaitsuka, T., Sato, Y., Tomizawa, K., Nairn, A.C., Takei, K., Matsui, H., and Matsushita, M. (2008). CaM kinase I α -induced phosphorylation of Drp1 regulates mitochondrial morphology. *J. Cell Biol.* **182**, 573–585.
- 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.
- Lee, D.H., Park, J.O., Kim, T.S., Kim, S.K., Kim, T.H., Kim, M.C., Park, G.S., Kim, J.H., Kuninaka, S., Olson, E.N., et al. (2016). LATS-YAP/TAZ controls lineage specification by regulating TGF β signaling and Hnf4 α expression during liver development. *Nat. Commun.* **7**, 11961.
- Means, A.L., Meszoely, I.M., Suzuki, K., Miyamoto, Y., Rustgi, A.K., Coffey, R.J., Jr., Wright, C.V., Stoffers, D.A., and Leach, S.D. (2005). Pancreatic epithelial plasticity mediated by acinar cell transdifferentiation and generation of nestin-positive intermediates. *Development* **132**, 3767–3776.
- Pan, F.C., Bankaitis, E.D., Boyer, D., Xu, X., Van de Castelee, M., Magnuson, M.A., Heimberg, H., and Wright, C.V. (2013). Spatiotemporal patterns of multipotentiality in Ptf1a-expressing cells during pancreas organogenesis and injury-induced facultative restoration. *Development* **140**, 751–764.
- Piccolo, S., Dupont, S., and Cordenonsi, M. (2014). The biology of YAP/TAZ: hippo signaling and beyond. *Physiol. Rev.* **94**, 1287–1312.
- Plaks, V., Brenot, A., Lawson, D.A., Linnemann, J.R., Van Kappel, E.C., Wong, K.C., de Sauvage, F., Klein, O.D., and Werb, Z. (2013). Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. *Cell Rep.* **3**, 70–78.
- 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.
- Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. *Nature* **439**, 993–997.
- Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189–193.
- Wang, D., Cai, C., Dong, X., Yu, Q.C., Zhang, X.O., Yang, L., and Zeng, Y.A. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517**, 81–84.
- Xu, J., Du, Y., and Deng, H. (2015). Direct lineage reprogramming: strategies, mechanisms, and applications. *Cell Stem Cell* **16**, 119–134.
- Yimlamai, D., Christodoulou, C., Galli, G.G., Yanger, K., Pepe-Mooney, B., Gurung, B., Shrestha, K., Cahan, P., Stanger, B.Z., and Camargo, F.D. (2014). Hippo pathway activity influences liver cell fate. *Cell* **157**, 1324–1338.
- Zanconato, F., Forcato, M., Battilana, G., Azzolin, L., Quaranta, E., Bodega, B., Rosato, A., Bicciato, S., Cordenonsi, M., and Piccolo, S. (2015). Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat. Cell Biol.* **17**, 1218–1227.
- Zanconato, F., Cordenonsi, M., and Piccolo, S. (2016). YAP/TAZ at the Roots of Cancer. *Cancer Cell* **29**, 783–803.
- Zhu, Y., Romero, M.I., Ghosh, P., Ye, Z., Charnay, P., Rushing, E.J., Marth, J.D., and Parada, L.F. (2001). Ablation of NF1 function in neurons induces abnormal development of cerebral cortex and reactive gliosis in the brain. *Genes Dev.* **15**, 859–876.

Targeted Epigenetic Remodeling of Endogenous Loci by CRISPR/Cas9-Based Transcriptional Activators Directly Converts Fibroblasts to Neuronal Cells

Joshua B. Black,¹ Andrew F. Adler,^{1,9} Hong-Gang Wang,^{6,8,10} Anthony M. D'Ippolito,^{2,3} Hunter A. Hutchinson,¹ Timothy E. Reddy,^{2,4} Geoffrey S. Pitt,^{6,7,8} Kam W. Leong,^{1,11} and Charles A. Gersbach^{1,2,5,*}

¹Department of Biomedical Engineering, Duke University, Durham, NC 27708, USA

²Center for Genomic and Computational Biology, Duke University, Durham, NC 27708, USA

³University Program in Genetics and Genomics, Duke University Medical Center, Durham, NC 27710, USA

⁴Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC 27710, USA

⁵Department of Orthopaedic Surgery, Duke University Medical Center, Durham, NC 27710, USA

⁶Ion Channel Research Unit, Duke University Medical Center, Durham, NC 27710, USA

⁷Department of Neurobiology, Duke University Medical Center, Durham, NC 27710, USA

⁸Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, NC 27710, USA

⁹Present address: Department of Neurosciences, University of California, San Diego, La Jolla, CA 92093, USA

¹⁰Present address: Cardiovascular Research Institute, Weill Cornell Medicine, New York, NY 10021, USA

¹¹Present address: Department of Biomedical Engineering, Columbia University, New York, NY 10027, USA

*Correspondence: charles.gersbach@duke.edu

<http://dx.doi.org/10.1016/j.stem.2016.07.001>

SUMMARY

Overexpression of exogenous fate-specifying transcription factors can directly reprogram differentiated somatic cells to target cell types. Here, we show that similar reprogramming can also be achieved through the direct activation of endogenous genes using engineered CRISPR/Cas9-based transcriptional activators. We use this approach to induce activation of the endogenous *Brn2*, *Ascl1*, and *Myt1l* genes (BAM factors) to convert mouse embryonic fibroblasts to induced neuronal cells. This direct activation of endogenous genes rapidly remodeled the epigenetic state of the target loci and induced sustained endogenous gene expression during reprogramming. Thus, transcriptional activation and epigenetic remodeling of endogenous master transcription factors are sufficient for conversion between cell types. The rapid and sustained activation of endogenous genes in their native chromatin context by this approach may facilitate reprogramming with transient methods that avoid genomic integration and provides a new strategy for overcoming epigenetic barriers to cell fate specification.

INTRODUCTION

Direct reprogramming of somatic cells has tremendous potential to advance applications in disease modeling, drug discovery, and gene and cell therapies. Common approaches to achieve cellular reprogramming rely on the ectopic expression of transgenes encoding lineage-specific transcription factors (Davis et al., 1987; Takahashi and Yamanaka, 2006; Vierbuchen

et al., 2010). To demonstrate stable cellular reprogramming to an autonomous cell phenotype, the expression of exogenous transcription factors should be transient. Thus the establishment of positive feedback networks regulating endogenous genes is necessary to sustain a transgene-independent cellular identity (Vierbuchen and Wernig, 2011). In many cases, the endogenous genes are occluded by *cis*-acting repressive chromatin marks that are slow to remodel (Vierbuchen and Wernig, 2012). This slow remodeling process typically necessitates prolonged expression of the exogenous factors, limiting the efficacy of transient delivery methods, and poses a major bottleneck to improving the efficiency, speed, and robustness of reprogramming (Hanna et al., 2009).

The type II clustered regularly interspaced short palindromic repeat (CRISPR) system and the CRISPR-associated Cas9 nuclease have recently been repurposed from an adaptive immune system in bacteria and archaea to a gene editing tool (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013b) and transcriptional regulator (Cheng et al., 2013; Gilbert et al., 2013; Konermann et al., 2013; Maeder et al., 2013b; Mali et al., 2013a; Perez-Pinera et al., 2013; Qi et al., 2013) of endogenous genes in mammalian cells. The ability to program these transcription factors to target any genomic locus of interest through the simple exchange of the 20-nt targeting sequence of the guide RNA (gRNA) enables a simple, robust, and highly scalable method for control of complex transcriptional networks (Thakore et al., 2016). Furthermore, dCas9-based transcription factors can target stably silenced genes within compacted chromatin to initiate chromatin remodeling and transcriptional activation (Perez-Pinera et al., 2013; Polstein et al., 2015). Thus, this technology may provide a method to deterministically initiate expression of endogenous gene networks of alternate cell lineages.

The CRISPR/Cas9 system and other platforms for programmable transcriptional regulation have been incorporated into methods for cellular reprogramming in a few recent studies. Gao et al. used transcription activator-like effector (TALE)-based

transactivators targeting an enhancer of *Oct4* to generate mouse induced pluripotent stem cells. Notably, that study required co-delivery of vectors directly encoding ectopic *C-MYC*, *KLF4*, and *SOX2* to achieve pluripotency (Gao et al., 2013). More recently, we have demonstrated the direct conversion of primary mouse embryonic fibroblasts (PMEFs) to skeletal myocytes using a dCas9-based transactivator targeting the endogenous *Myod1* gene (Chakraborty et al., 2014). Several groups have also applied CRISPR/Cas9-based transcriptional regulation to direct the differentiation of human induced pluripotent and embryonic stem cells (Balboa et al., 2015; Chavez et al., 2015; Wei et al., 2016).

The above examples involve the targeted activation of a single transcription factor to guide reprogramming or differentiation, but many approaches require concurrent expression of multiple factors to efficiently establish a mature phenotype (Takahashi and Yamanaka, 2006; Vierbuchen et al., 2010). There have been no examples demonstrating multiplex endogenous gene activation to induce cellular reprogramming, and the versatility of that approach for direct conversion to other cell phenotypes is not known. Moreover, only the report of TALE transcription factors targeting *Oct4* evaluated changes to epigenetic marks at the target loci (Gao et al., 2013), and this group later reported that dCas9-based transcriptional activators were inefficient at endogenous gene activation and reprogramming (Gao et al., 2014). In this study, we tested the hypothesis that targeted epigenetic reprogramming of the regulatory elements controlling expression of lineage-specific transcription factors is sufficient for direct conversion between cell types by applying dCas9-based transactivators to the activation of endogenous genes that directly convert PMEFs to induced neuronal cells (iNs).

RESULTS

Multiplex Endogenous Gene Activation of Neurogenic Factors in PMEFs

Overexpression of transgenes encoding the transcription factors *Brn2*, *Ascl1*, and *Myt1l* (BAM factors) has been shown to directly convert cultured PMEFs to functional induced neuronal cells (Vierbuchen et al., 2010). We hypothesized that the targeted activation of the endogenous genes encoding these same factors in their native chromatin context via a dCas9-based transactivator could more rapidly and deterministically remodel the chromatin at the target loci and provide an alternate method to achieve the reprogramming of PMEFs to iNs (Figure 1A). To achieve targeted gene activation, we used a transactivator with both N-terminal and C-terminal VP64 transactivation domains (^{VP64}dCas9^{VP64}) (Chakraborty et al., 2014) that generated a ~10-fold improvement in activation of *ASCL1* in HEK293T cells at 3 days post-transfection compared to the first-generation dCas9 transcription factor with a single C-terminal VP64 domain (Maeder et al., 2013b; Perez-Pinera et al., 2013) (Figure 1B). We used ^{VP64}dCas9^{VP64} for the remainder of this study.

We used lentiviral delivery to constitutively express ^{VP64}dCas9^{VP64} in PMEFs. Initially, we delivered the gRNAs through transient transfection of plasmid DNA in order to assess stable reprogramming of cell phenotype following transient activity of transactivators. The induction of *Brn2* and *Ascl1* gene expression by ^{VP64}dCas9^{VP64} was attained by delivering

four gRNAs targeted to the putative promoter region directly upstream of the transcription start site (TSS). The decision to deliver four gRNAs for each gene was based on the reported synergistic effects of multiple gRNAs on gene activation (Maeder et al., 2013b; Mali et al., 2013a; Perez-Pinera et al., 2013). The optimal gRNAs were selected from a pool of eight gRNAs through elimination screening (Figure S1A). The gRNAs targeting regions proximal to the TSS of the *Myt1l* locus did not induce detectable levels of activation, but targeting an intronic region directly upstream of the first coding exon of *Myt1l* was sufficient to activate expression (Figure S1B).

Co-transfection of 12 gRNA expression plasmids (CR-BAM), targeting each of the three endogenous BAM factors with 4 gRNAs, into PMEFs stably expressing ^{VP64}dCas9^{VP64} was sufficient to induce transcriptional upregulation of all three endogenous genes when compared to the transfection of a plasmid encoding firefly luciferase (pLuc; Figure 1C). We also detected *Brn2* and *Ascl1* protein expression by western blot (Figure S1C), although we could not detect *Myt1l* protein using commercially available antibodies. In addition to gRNA transfections, we transfected three plasmids encoding the BAM factor transgenes under the control of the EF1 α /HTLV promoter (pBAM) into the same cells and observed a modest increase in the mRNA levels of the corresponding endogenous genes (Figure 1C).

To attain successful reprogramming, it is generally considered necessary to express the exogenous factors at high levels (Vierbuchen and Wernig, 2011). Therefore, we compared the total mRNA and protein levels of *Brn2*, *Ascl1*, and *Myt1l* produced 3 days after CR-BAM and pBAM plasmid transfections (Figures 1D–1F). Despite the higher levels of transcriptional activation from the endogenous loci by CR-BAM (Figure 1C), pBAM transfection generated significantly more total mRNA encoding each BAM factor than induction by CR-BAM, as determined by qRT-PCR (Figure 1D). Quantitation of single-cell protein levels from immunofluorescence staining also revealed significantly higher single-cell levels of *Brn2* and *Ascl1* in cells transfected with pBAM compared to those transfected with CR-BAM (Figures 1E and 1F).

Induction of Neuronal Cells from PMEFs via ^{VP64}dCas9^{VP64}-Mediated Gene Activation

Treated PMEFs were assayed for neuronal phenotypes as detailed schematically in Figure 2A. We observed an increase in mRNA of the early pan-neuronal marker β III tubulin (*Tuj1*) 3 days after transfection with either pBAM or CR-BAM when compared to a pLuc control (Figure 2B). We cultured the cells for 2 weeks in neurogenic medium and analyzed expression of pan-neuronal markers by immunofluorescence staining. We identified cells with neuronal morphologies that expressed *Tuj1* in populations transfected with CR-BAM (Figure 2C). A subset of *Tuj1*⁺ cells also expressed the more mature pan-neuronal marker *Map2* (Figure 2C). The generation of *Tuj1*⁺ *Map2*⁺ cells with neuronal morphologies following treatment with ^{VP64}dCas9^{VP64} and gRNAs was contingent on the addition of a small-molecule cocktail to the medium that has been used previously for neural differentiation of embryonic stem cells and has been shown to improve the efficiency of the direct conversion of human fibroblasts to neurons when used in parallel

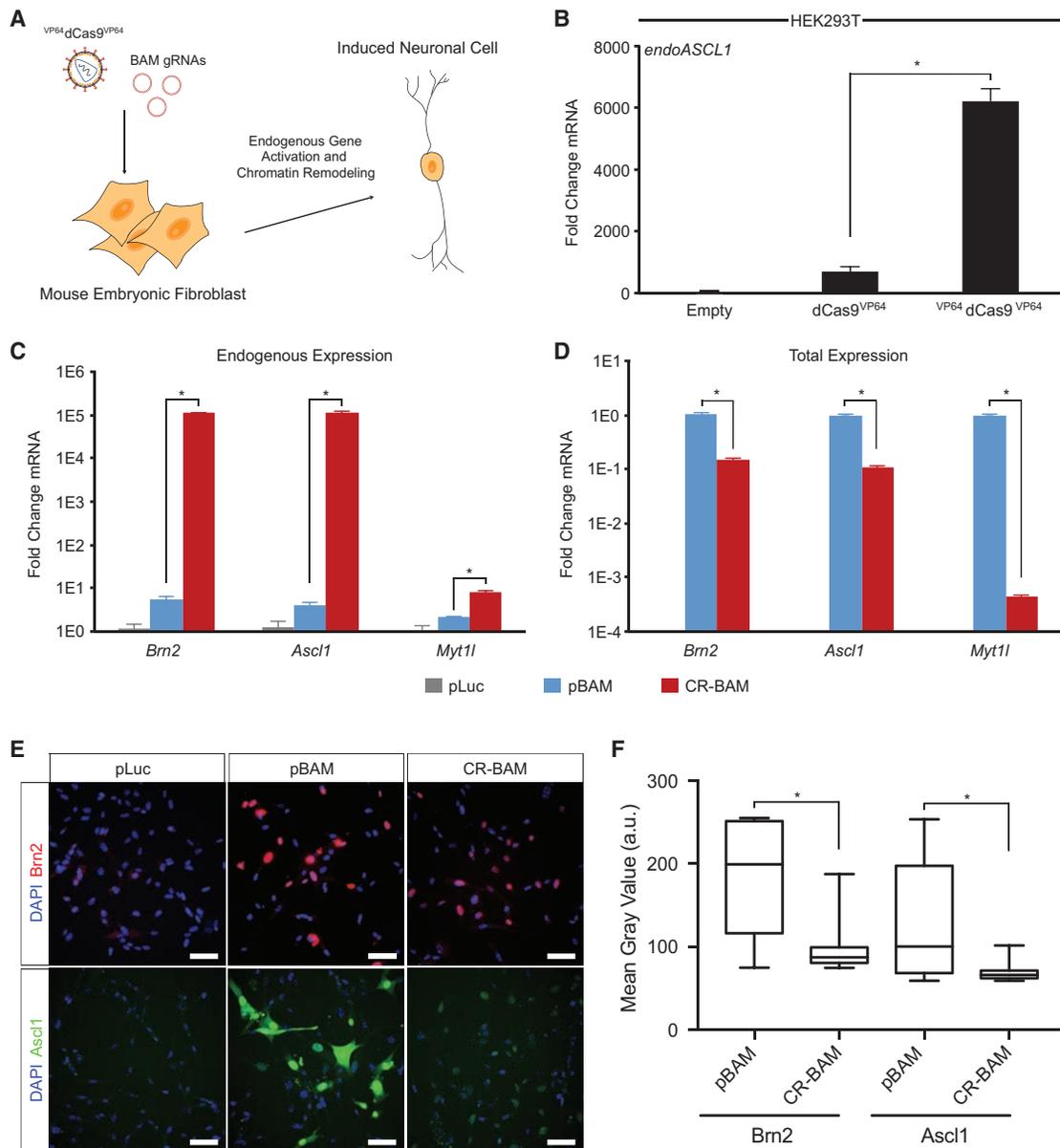


Figure 1. Endogenous Gene Activation of Neuronal Transcription Factors in PMEFs

(A) Reprogramming of PMEFs to neuronal cells via transduction of *VP64-dCas9^{VP64}* and transfection of gRNA expression plasmids targeting the endogenous BAM factors.

(B) Transcriptional activation of *ASCL1* in HEK293T cells with *dCas9^{VP64}* or *VP64-dCas9^{VP64}* (**p* < 0.05).

(C and D) Endogenous expression (C) and total expression (D) of the BAM factors in PMEFs with targeted activation (CR-BAM) or ectopic overexpression (pBAM; **p* < 0.05).

(E) Immunofluorescence staining of *Brn2* and *Ascl1* in PMEFs demonstrated protein expression through targeted activation of the endogenous loci or expression from ectopic plasmids (scale bar, 50 μ m).

(F) Automated image analysis of fluorescence intensity revealed significantly more single-cell *Brn2* and *Ascl1* protein with pBAM transfection compared to CR-BAM (**p* < 0.05 between distributions of single-cell mean fluorescence; Z-test).

All gRNAs used are listed in Table S1. All assays were performed on day 3 post-transfection. qRT-PCR data are presented as mean \pm SEM for *n* = 3 biological replicates. *p* values for qRT-PCR data were determined by global one-way ANOVA with Holm-Bonferroni post hoc tests (α = 0.05). See also Figure S1.

with ectopic expression of neural transcription factors (Ladewig et al., 2012).

We used a lentiviral fluorescent reporter encoding dsRed-Express under the control of the synapsin I promoter (Syn-RFP) as a proxy to define the most functionally mature iNs in

the heterogeneous population of reprogrammed cells (Adler et al., 2012). We readily identified RFP⁺ cells with elaborate arborizations in CR-BAM-transfected PMEFs (Figure 2C). We also identified rare cells with fibroblastic morphologies reactive to the Tuj1 antibody in PMEFs following pLuc transfection

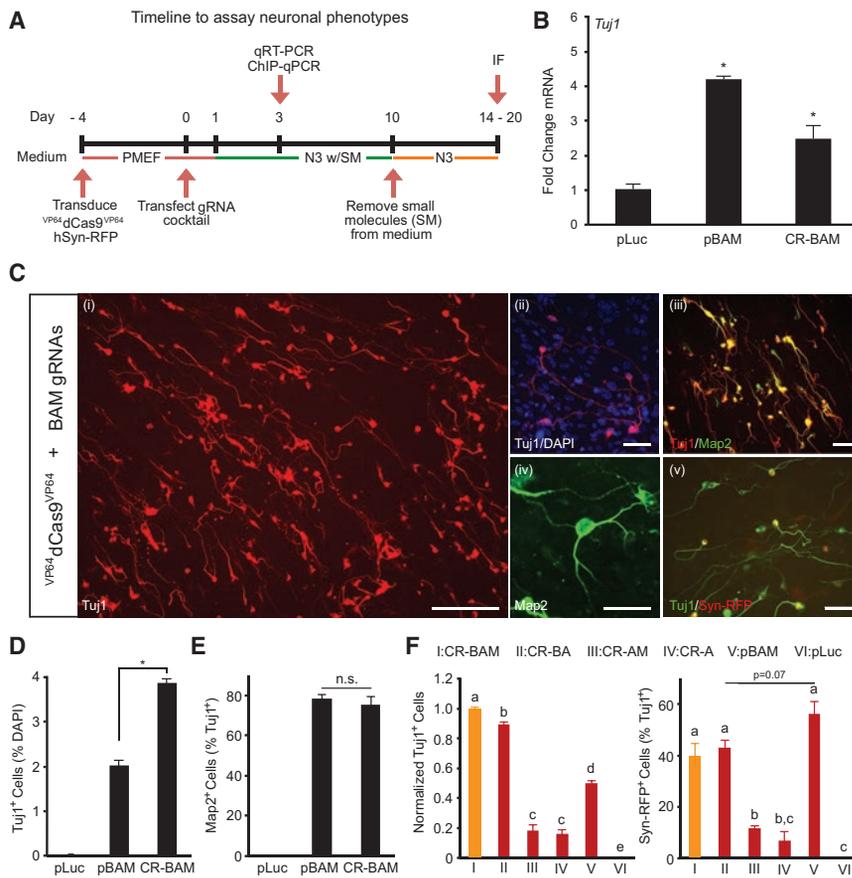


Figure 2. Induction of Neuronal Cells from PMEFs via $VP64dCas9^{VP64}$ -Mediated Gene Activation

(A) PMEFs were transduced with a lentivirus encoding the $VP64dCas9^{VP64}$ transactivator and subsequently transfected with gRNAs targeting *Brn2*, *Ascl1*, and *Myt1l*. Neuronal phenotypes were assayed as indicated.

(B) Transcriptional activation of *TuJ1* was detected in PMEFs at day 3 post-transfection of pBAM or CR-BAM (* $p < 0.05$ relative to transfection of a plasmid encoding firefly luciferase [pLuc]).

(C) Immunofluorescence staining revealed numerous *TuJ1*⁺ cells with neuronal morphologies co-expressing *Map2* at day 14 post-transfection of CR-BAM. The cells with the most elaborate neuronal morphologies activated the synapsin promoter in a Syn-RFP lentiviral reporter (scale bars, 100 μ m [i], 50 μ m [ii–v]).

(D) Quantitation of *TuJ1*⁺ cells as percent nuclei at day 14 post-transfection of pLuc, pBAM, or CR-BAM (* $p < 0.05$).

(E) Quantitation of *Map2*⁺ cells as percent *TuJ1*⁺ cells at day 14 post-transfection of pLuc, pBAM, or CR-BAM (n.s., not significant).

(F) Quantitation of *TuJ1*⁺ and RFP⁺ cells with transfection of different combinations of gRNAs. *TuJ1*⁺ cells are normalized to CR-BAM transfection. Conditions that share the same letter (a–e) are not significantly different.

p values were determined by global one-way ANOVA with Holm-Bonferroni post hoc tests ($\alpha = 0.05$). See also Figure S2.

(Figure S2A), but these cells were never reactive to the *Map2* antibody. Consistent with previous studies, direct overexpression of the ectopic BAM factors via transfection of constitutive expression plasmids generated *TuJ1*⁺*Map2*⁺ cells with neuronal morphologies (Figure S2B) (Adler et al., 2012; Vierbuchen et al., 2010).

Image analysis revealed that CR-BAM transfection generated a modest, but statistically significant and reproducible, increase in the number of *TuJ1*⁺ cells compared to pBAM transfection after 14 days in culture post-transfection (Figure 2D), despite much lower overall expression of the BAM factors (Figures 1D–1F). There was no difference in the percentage of *TuJ1*⁺ cells that also expressed *Map2* (Figure 2E). To evaluate the contribution of each neurogenic factor to the generation of *TuJ1*⁺ cells and to the level of neuronal maturation, we transfected gRNAs targeting different combinations of the endogenous factors. Removal of gRNAs targeting the *Brn2* locus attenuated iN production ~5-fold when compared to that generated with targeted activation of all three endogenous factors (Figure 2F). We detected a slight reduction in *TuJ1*⁺ cell production with the removal of *Myt1l* gRNAs (Figure 2F). Neuronal maturity was assessed as the percentage of *TuJ1*⁺ cells co-positive for the Syn-RFP reporter. Removal of *Brn2* gRNAs reduced the percentage of RFP⁺ cells >2-fold, but no change was detected with removal of *Myt1l* gRNAs (Figure 2F). pBAM transfection generated a higher percentage of RFP⁺ cells than CR-BAM transfection, though it was not statistically significant (Figure 2F).

Induction of Endogenous Gene Expression Is Rapid and Sustained

For any reprogramming strategy, activation of the endogenous genes encoding the master fate-specifying transcription factors is an important step to the successful reprogramming and stability of the new cellular phenotype (Vierbuchen and Wernig, 2011). Consequently, we compared the kinetics of endogenous gene expression through late stages of reprogramming with pBAM or CR-BAM transfection. We observed activation of all three endogenous genes as early as 1 day post-transfection with CR-BAM that remained at high levels through day 18 in culture (Figure S3A). Expression of the BAM factors from the endogenous loci was significantly higher with targeted activation via CR-BAM compared to ectopic overexpression via pBAM transfection throughout the time course of the experiment. Activation of the endogenous genes by pBAM transfection was delayed, and a significant and sustained increase over baseline levels was only detected for endogenous *Ascl1* and *Myt1l* (Figure S3A).

We next assessed the kinetics of expression of the downstream pan-neuronal marker *TuJ1*. Both pBAM and CR-BAM treatment generated a significant increase in *TuJ1* expression throughout the time course of the experiment (Figure S3B). At early time points, *TuJ1* levels were higher with pBAM treatment than CR-BAM. However, *TuJ1* levels with pBAM treatment peaked 7 days post-transfection and declined thereafter, whereas expression following CR-BAM treatment remained

stable through day 18 in culture (Figure S3B). Importantly, the exogenous BAM factors and gRNAs were significantly depleted by day 18 in culture after transient transfection (Figure S3C), though levels of activation from the endogenous genes remained high in cells treated with CR-BAM (Figure S3A).

Direct Activation via $VP64$ dCas9 $VP64$ Rapidly Remodels Chromatin at Target Loci

The kinetics of gene activation led us to speculate whether the rapid and sustained elevated levels of endogenous gene expression achieved with CR-BAM corresponded to an altered epigenetic program at the target loci. We used chromatin immunoprecipitation followed by next-generation sequencing (ChIP-seq) data generated as part of the Encyclopedia of DNA Elements (ENCODE) Project (Mouse ENCODE Consortium, 2012) to identify histone modifications enriched at the transcriptionally active BAM factor loci in mouse embryonic brain tissue, including H3K27ac and H3K4me3 (Figures 3A, 3C, and S4A). We hypothesized that targeting the endogenous BAM factors for activation with $VP64$ dCas9 $VP64$ in PMEFs could recapitulate the chromatin signatures found at these loci in developing brain tissue.

To investigate the effects of BAM-factor induction on the epigenetic programming at the target loci, we performed chromatin immunoprecipitation (ChIP) qPCR in PMEFs transduced with $VP64$ dCas9 $VP64$ and transfected with pLuc, pBAM, or CR-BAM plasmids (Figures 3 and S4). We used qPCR primers tiled along intragenic and regulatory regions of the *Brn2*, *Ascl1*, and *Myt1l* loci. We detected a significant enrichment in H3K27ac and H3K4me3 at the *Brn2* and *Ascl1* loci on day 3 post-transfection of CR-BAM (Figures 3B and 3D). H3K4me3 was enriched along the gene bodies of *Brn2* and *Ascl1*. H3K27ac was enriched along the gene bodies and regions surrounding the putative promoter sequences of both genes. In contrast, targeted activation of *Myt1l* only induced modest detectable enrichment in H3K27ac at the gRNA target sites directly upstream of the first coding exon (Figure S4B). No significant change in H3K27ac or H3K4me3 was measured within the putative *Myt1l* promoter. Though overexpression of the BAM factors induced modest levels of expression of the endogenous genes by day 3 post-transfection (Figures 1C and S3A), we did not detect corresponding enrichment in H3K27ac and H3K4me3 at the endogenous loci (Figures 3B, 3D, and S4B).

Generation of Induced Neuronal Cells with Multiplex gRNA Lentiviral Vectors

To explore a strategy for stable expression of the CRISPR/Cas9 transcription factors, and to see if the same outcomes observed with transient expression held true with constitutive expression, we used a single lentiviral vector capable of expressing four gRNAs from four independent RNA polymerase III promoters (Kabadi et al., 2014) (Figure 4A). Co-transduction of lentiviruses encoding $VP64$ dCas9 $VP64$ and a set of four gRNAs targeting each of the three BAM factors (lentiCR-BAM) permitted concurrent activation of the endogenous BAM factors in PMEFs by day 6 post-transduction (Figure 4B). For comparison, we used lentiviral vectors directly encoding the BAM factors (lentiBAM), and demonstrated activation of the corresponding endogenous genes by day 6 post-transduction (Figure 4B). Similar to the results we obtained

with transient transfection of expression plasmids, targeted activation of the endogenous genes via lentiviral delivery generated significantly more endogenous transcript from the *Brn2* and *Ascl1* loci than that induced through ectopic expression of the BAM factors. However, unlike the transfection experiments, endogenous *Myt1l* expression was significantly higher with transduction of lentiBAM compared to lentiCR-BAM (Figure 4B).

Following extended culture for 2 weeks in neurogenic medium, we readily identified Tuj1⁺Map2⁺ cells with complex neuronal morphologies (Figure 4C). All Tuj1⁺ cells identified also co-expressed Map2. To promote further neuronal maturation and for electrophysiological assessments, PMEFs were replated onto a previously established monolayer of primary rat astrocytes following transduction of $VP64$ dCas9 $VP64$ and gRNAs (Vierbuchen et al., 2010). Synapsin-RFP expression and cell morphology were used to select the most mature neuronal cells for patch-clamp analysis after 21 days in culture. In current-clamp mode, single or multiple action potentials were readily elicited in response to depolarizing current injections (six out of seven cells analyzed; Figure 4D). The same cells displayed voltage-dependent inward and outward currents. The transient inward currents were abolished in the presence of the voltage-gated Na⁺ channel blocker tetrodotoxin (TTX; Figure 4E). The average resting membrane potential, action potential (AP) threshold and AP amplitude were -41 ± 3.8 mV, -33 ± 2.6 mV, and 49 ± 9.7 mV, respectively (mean \pm SEMs, n = 7 cells).

In contrast to what we observed by transient transfection of the reprogramming factors, constitutive expression of the BAM factor transgenes via lentiviral vectors generated significantly more Tuj1⁺Map2⁺ cells than that detected with $VP64$ dCas9 $VP64$ (Figure 4F). We hypothesized that the prolonged and high levels of expression of the BAM factor transgenes enabled by lentiviral delivery permitted further epigenetic and transcriptional reprogramming that improved the efficiency of iN generation when compared to transient transfection methods. Consequently, we revisited the analysis of chromatin remodeling at the endogenous BAM factor loci in the context of lentiviral delivery of the reprogramming factors. We found that, as shown with transient transfection, targeted activation of the endogenous genes via lentiCR-BAM transduction led to the rapid deposition of H3K27ac at the *Brn2* and *Ascl1* loci as early as day 3 post-transduction that persisted at day 6 (Figure 4G). Also, as seen with transient transfection, we did not detect enrichment of H3K27ac at the *Myt1l* locus with lentiCR-BAM transduction, although we did measure an increase in *Myt1l* mRNA (Figures 4B and 4G). In contrast to what we observed with transient transfection of the BAM factors, we detected enrichment of H3K27ac along regions of all three endogenous genes with lentiBAM transduction (Figure 4G). Furthermore, we only detected minor enrichment in H3K27ac at all three genes at day 3 post-transduction of lentiBAM; however, both *Ascl1* and *Myt1l* showed a substantial increase in H3K27ac deposition by day 6 post-transduction (Figure 4G).

DISCUSSION

In this study, we demonstrate direct cellular reprogramming to induced neuronal cells through targeted activation of endogenous genes. We utilized the CRISPR/Cas9 system as a programmable, locus-specific transcriptional regulator for the

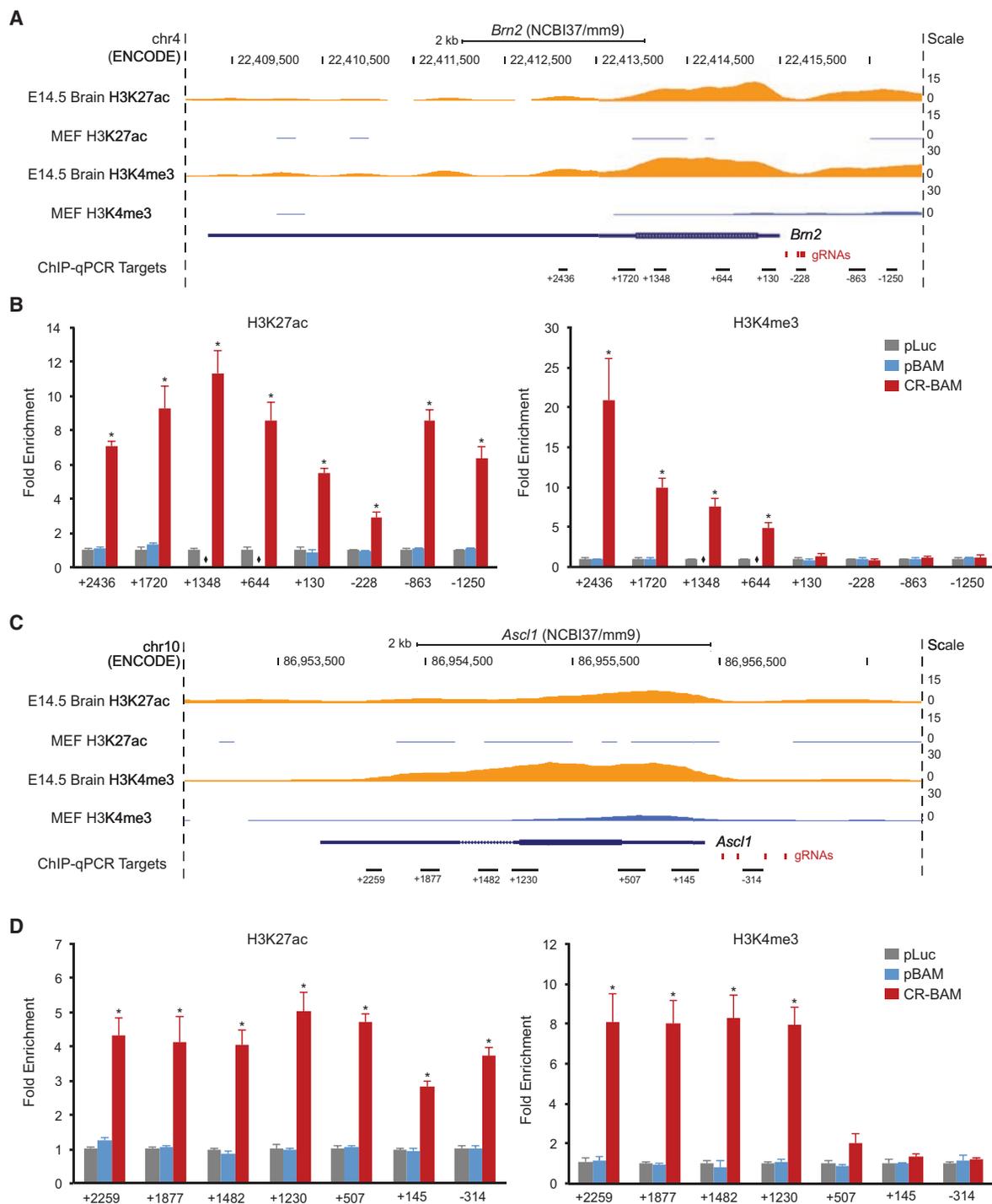


Figure 3. ^{VP64}dCas9^{VP64} Rapidly Remodels Epigenetic Marks at Target Loci

(A and C) Mouse genomic tracks depicting histone H3 modifications H3K27ac and H3K4me3 at the *Brn2* and *Ascl1* loci in embryonic brain tissue and fibroblasts (data from Mouse ENCODE; GEO: GSE31039). Red bars indicate gRNA target sites near the transcription start site, and black bars indicate the location of ChIP-qPCR amplicons along the gene locus.

(B and D) Targeted activation of endogenous *Brn2* and *Ascl1* in PMEFs induced significant enrichment of H3K27ac and H3K4me3 at multiple sites along the genomic loci at day 3 post-transfection ($p < 0.05$, one-way ANOVA with Holm-Bonferroni post hoc tests, $n = 3$ biological replicates). Overexpression of the BAM factors via transfection of expression plasmids encoding BAM factor transgenes did not induce a significant change in these chromatin marks. qPCR primers targeting coding regions of the genes are not included for the pBAM transfection condition, as contaminating plasmid DNA biased enrichment values, and are marked with diamonds in (B). All fold enrichments are relative to transfection of a plasmid encoding firefly luciferase and normalized to a region of the *Gapdh* locus. See also Figure S3 and S4.

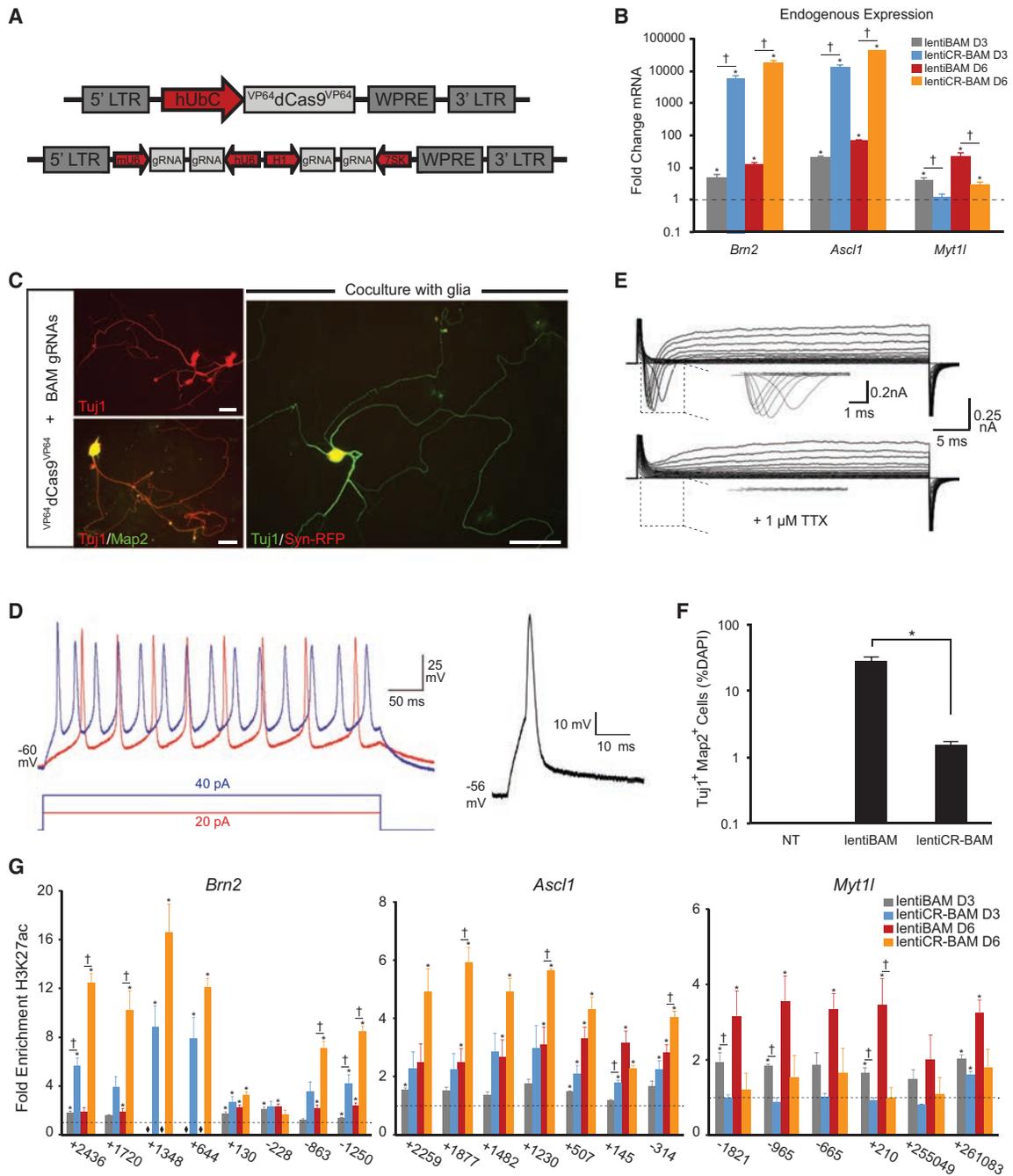


Figure 4. Generation of Functionally Mature iNs with Multiplex gRNA Vectors

(A) Schematic of $VP64dCas9^{VP64}$ and multiplex gRNA lentiviral constructs used to enable stable integration and constitutive expression.

(B) Relative mRNA expression of the endogenous BAM factors following transduction of transgenes encoding the BAM factors (lentiBAM) or $VP64dCas9^{VP64}$ and gRNAs targeting the endogenous BAM factors (lentiCR-BAM; * $p < 0.05$ relative to non-treated PMEFs; † $p < 0.05$ between lentiBAM versus lentiCR-BAM transduction).

(C) Immunofluorescence staining of PMEFs following transduction of lentiCR-BAM. Cells were co-positive for Tuj1 and Map2 and exhibited complex neuronal morphologies (scale bar, 50 μ m).

(D) Action potentials were evoked from $VP64dCas9^{VP64}$ -induced neuronal cells in response to 5-ms (right) or 500-ms (left) step depolarizing current injection (six out of seven cells analyzed) after empiric hyperpolarizing current injection to hold membrane potential at ~ -60 mV.

(E) Representative whole-cell currents recorded with or without perfusion of 1 μ M tetrodotoxin (TTX).

(F) Quantitation of Tuj1⁺Map2⁺ cells as percent nuclei (* $p < 0.05$ between lentiBAM versus lentiCR-BAM transduction; NT, non-treated PMEFs).

(G) Time course of H3K27ac enrichment along the *Brn2*, *Ascl1*, and *Myt11* loci (* $p < 0.05$ relative to non-treated PMEFs; † $p < 0.05$ between lentiBAM versus lentiCR-BAM transduction).

All p values calculated by global ANOVA with Holm-Bonferroni post hoc tests ($\alpha = 0.05$).

multiplex activation of the neurogenic factors *Brn2*, *Ascl1*, and *Myt1l* (BAM factors). We hypothesized that targeted activation of the endogenous genes in PMEFs, as opposed to the forced overexpression of the corresponding transgenes, could more directly access the endogenous loci and rapidly remodel their epigenetic signatures, thus potentially reflecting a more natural mechanism of action and serving as an alternate method to achieve cell lineage conversion.

In PMEFs, the *cis*-repressive chromatin landscape at neuronal loci may preclude binding of regulatory factors, in turn impeding transcriptional activation. As a result, expression of the BAM factors in PMEFs from exogenous vectors likely relies on stochastic processes for reactivation of the corresponding endogenous genes. Furthermore, transient delivery of the BAM factors, as done in our initial experiments (Figures 1, 2, and 3), limits the time window within which the endogenous networks and positive feedback loops can be established. We demonstrated that targeting the endogenous genes directly induced the enrichment of histone H3 modifications H3K27ac and H3K4me3 at the *Brn2* and *Ascl1* loci at 3 days post-transfection, whereas transgene overexpression via transfection of plasmids encoding the reprogramming factors did not alter these chromatin marks (Figures 3 and S4). Additionally, we observed sustained high levels of expression from the endogenous genes at later stages of reprogramming despite the transient delivery of the gRNA plasmids (Figure S3).

In contrast, we found that stable integration and constitutive expression of the exogenous reprogramming factors via lentiviral delivery led to the eventual deposition of H3K27ac at their endogenous loci with a concomitant improvement in reprogramming capacity (Figures 4F and 4G). We did not observe a similar improvement with constitutive expression of VP64 dCas9 VP64 and gRNAs, which is possibly attributable to the lower levels of overall expression of the neuronal transcription factors achieved by transactivation of the endogenous genes compared to ectopic overexpression. Consequently, the direct activation of the endogenous genes via VP64 dCas9 VP64 may be more amenable to transient delivery approaches that avoid undesired consequences of vector integration into the genome. Such transient methods, including the direct delivery of ribonucleoprotein Cas9-gRNA complexes, may be a more clinically translatable method of generating reprogrammed cells that are genetically unmodified.

Achieving robust and well-defined reprogrammed cell populations is still a central challenge. Reprogrammed cells often fail to acquire completely mature phenotypes and can retain epigenetic remnants of the native cell type (Kim et al., 2010). Moreover, a recent study demonstrated that reprogramming efficiency can be limited by divergence to a competing cell identity (Treutlein et al., 2016). The molecular mechanisms and practical consequences of these limitations are largely unknown. As the toolkit of designer transcription factors expands to precisely modify the epigenome (Hilton et al., 2015; Kearns et al., 2015; Maeder et al., 2013a; Mendenhall et al., 2013; Thakore et al., 2016), these tools may be used to prime specific genomic loci in diverse cell types, promote endogenous transcription factor binding, and directly correct regions of epigenetic remnants that prove to be problematic for a given application. This may lead to improved reprogramming fidelity and extension of the breadth of donor cells amenable to reprogramming.

EXPERIMENTAL PROCEDURES

Cell Culture, Transfections, and Viral Transductions

PMEFs were maintained in high serum media during transduction and transfection of expression plasmids and subsequently cultured in neurogenic serum-free medium for the duration of the experiments to promote neuronal survival and maturation. Lentivirus was produced in HEK293T cells using the calcium phosphate precipitation method. All transfections were performed using Lipofectamine 3000 (Invitrogen) in accordance with the manufacturer's protocol. All expression plasmids used in this study can be found in Table S2.

Immunofluorescence Staining and qRT-PCR

All sequences for qRT-PCR primers can be found in Table S3. Total RNA was isolated using the QIAGEN RNeasy and QIAshredder kits, reverse transcribed using the SuperScript VILO Reverse Transcription Kit (Invitrogen), and analyzed using Perfecta SYBR Green Fastmix (Quanta BioSciences). All qRT-PCR data are presented as fold change in RNA normalized to *Gapdh* expression. For immunofluorescence staining, cells were fixed with 4% paraformaldehyde, permeabilized with 0.2% Triton X-100, and incubated with primary and secondary antibodies.

Electrophysiology

A synapsin-RFP lentiviral reporter was used to identify cells in co-culture with primary rat astroglia for patch-clamp analysis at indicated time points. Action potentials and inward and outward currents were recorded in whole-cell configuration. Data were analyzed and prepared for publication using pCLAMP and MATLAB.

Chromatin Immunoprecipitation qPCR

Chromatin was immunoprecipitated using antibodies against H3K27ac and H3K4me3, and gDNA was purified for qPCR analysis. All sequences for ChIP-qPCR primers can be found in Table S3. qPCR was performed using SYBR green Fastmix (Quanta BioSciences), and the data are presented as fold change gDNA relative to negative control and normalized to a region of the *Gapdh* locus.

Mouse ENCODE ChIP-Sequencing Datasets

H3K4me3 and H3K27ac ChIP-sequencing data from C57BL/6 E14.5 whole brain and mouse embryonic fibroblasts (GSE31039) were acquired from the Mouse ENCODE Consortium generated in Bing Ren's laboratory at the Ludwig Institute for Cancer Research.

Statistical Methods

Statistical analysis was done using GraphPad Prism 7. All data were analyzed with at least three biological replicates and presented as mean \pm SEM. See figure legends for details on specific statistical tests run and p values calculated for each experiment.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.07.001>.

AUTHOR CONTRIBUTIONS

J.B.B., A.F.A., K.W.L., and C.A.G. designed experiments. J.B.B., A.F.A., H.-G.W., A.M.D., and H.A.H. performed the experiments. All authors analyzed the data. J.B.B. and C.A.G. wrote the manuscript with contributions by all authors.

CONFLICTS OF INTEREST

C.A.G. and J.B.B. are inventors on filed patent applications related to this work.

ACKNOWLEDGMENTS

The authors thank Ami Kabadi, Pablo Perez-Pinera, and Syandan Chakraborty for providing plasmid constructs; Christopher Grigsby for protocols for PMEF

transfections and providing the pUNO-Brn2 plasmid; Rui Dai for technical assistance; and Jeff Stogsdill and Cagla Eroglu for supplying astrocytes and providing protocols for co-culture experiments. This work was supported by US NIH grants to C.A.G. and T.E.R. (R01DA036865 and U01HG007900), an NIH Director's New Innovator Award (DP2OD008586) and National Science Foundation (NSF) Faculty Early Career Development (CAREER) Award (CBET-1151035) to C.A.G., and NIH grant P30AR066527. J.B.B. was supported by an NIH Biotechnology Training Grant (T32GM008555).

Received: May 13, 2015

Revised: May 11, 2016

Accepted: June 30, 2016

Published: August 11, 2016

REFERENCES

- Adler, A.F., Grigsby, C.L., Kulangara, K., Wang, H., Yasuda, R., and Leong, K.W. (2012). Nonviral direct conversion of primary mouse embryonic fibroblasts to neuronal cells. *Mol. Ther. Nucleic Acids* *1*, e32.
- Balboa, D., Weltner, J., Euroala, S., Trokovic, R., Wartiovaara, K., and Otonkoski, T. (2015). Conditionally stabilized dCas9 activator for controlling gene expression in human cell reprogramming and differentiation. *Stem Cell Reports* *5*, 448–459.
- Chakraborty, S., Ji, H., Kabadi, A.M., Gersbach, C.A., Christoforou, N., and Leong, K.W. (2014). A CRISPR/Cas9-based system for reprogramming cell lineage specification. *Stem Cell Reports* *3*, 940–947.
- Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., P R Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* *12*, 326–328.
- Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* *23*, 1163–1171.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.
- Mouse ENCODE Consortium (2012). An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome Biol.* *13*, 418.
- Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* *51*, 987–1000.
- Gao, X., Yang, J., Tsang, J.C., Ooi, J., Wu, D., and Liu, P. (2013). Reprogramming to pluripotency using designer TALE transcription factors targeting enhancers. *Stem Cell Reports* *1*, 183–197.
- Gao, X., Tsang, J.C., Gaba, F., Wu, D., Lu, L., and Liu, P. (2014). Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Res.* *42*, e155.
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* *154*, 442–451.
- Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creighton, M.P., van Oudenaarden, A., and Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* *462*, 595–601.
- Hilton, I.B., D'Ipollito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* *33*, 510–517.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.
- Kabadi, A.M., Ousterout, D.G., Hilton, I.B., and Gersbach, C.A. (2014). Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res.* *42*, e147.
- Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat. Methods* *12*, 401–413.
- Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I., et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* *467*, 285–290.
- Konermann, S., Brigham, M.D., Trevino, A.E., Hsu, P.D., Heidenreich, M., Cong, L., Platt, R.J., Scott, D.A., Church, G.M., and Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. *Nature* *500*, 472–476.
- Ladewig, J., Mertens, J., Kesavan, J., Doerr, J., Poppe, D., Glaue, F., Herms, S., Wernet, P., Kögler, G., Müller, F.J., et al. (2012). Small molecules enable highly efficient neuronal conversion of human fibroblasts. *Nat. Methods* *9*, 575–578.
- Maeder, M.L., Angstman, J.F., Richardson, M.E., Linder, S.J., Cascio, V.M., Tsai, S.Q., Ho, Q.H., Sander, J.D., Reyon, D., Bernstein, B.E., et al. (2013a). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat. Biotechnol.* *31*, 1137–1142.
- Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013b). CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods* *10*, 977–979.
- Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* *31*, 833–838.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013b). RNA-guided human genome engineering via Cas9. *Science* *339*, 823–826.
- Mendenhall, E.M., Williamson, K.E., Reyon, D., Zou, J.Y., Ram, O., Joung, J.K., and Bernstein, B.E. (2013). Locus-specific editing of histone modifications at endogenous enhancers. *Nat. Biotechnol.* *31*, 1133–1136.
- Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* *10*, 973–976.
- Polstein, L.R., Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Bledsoe, P., Song, L., Safi, A., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Res.* *25*, 1158–1169.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* *152*, 1173–1183.
- 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.
- Thakore, P.I., Black, J.B., Hilton, I.B., and Gersbach, C.A. (2016). Editing the epigenome: technologies for programmable transcription and epigenetic modulation. *Nat. Methods* *13*, 127–137.
- Treutlein, B., Lee, Q.Y., Camp, J.G., Mall, M., Koh, W., Shariati, S.A., Sim, S., Neff, N.F., Skotheim, J.M., Wernig, M., and Quake, S.R. (2016). Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* *534*, 391–395.
- Vierbuchen, T., and Wernig, M. (2011). Direct lineage conversions: unnatural but useful? *Nat. Biotechnol.* *29*, 892–907.
- Vierbuchen, T., and Wernig, M. (2012). Molecular roadblocks for cellular reprogramming. *Mol. Cell* *47*, 827–838.
- Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* *463*, 1035–1041.
- Wei, S., Zou, Q., Lai, S., Zhang, Q., Li, L., Yan, Q., Zhou, X., Zhong, H., and Lai, L. (2016). Conversion of embryonic stem cells into extraembryonic lineages by CRISPR-mediated activators. *Sci. Rep.* *6*, 19648.

A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis

Masayuki Fujii,^{1,3,5} Mariko Shimokawa,^{1,5} Shoichi Date,^{1,4} Ai Takano,¹ Mami Matano,¹ Kosaku Nanki,¹ Yuki Ohta,¹ Kohta Toshimitsu,¹ Yoshihiro Nakazato,¹ Kenta Kawasaki,¹ Toshio Uraoka,² Toshiaki Watanabe,³ Takanori Kanai,¹ and Toshiro Sato^{1,*}

¹Department of Gastroenterology

²Division for Research and Development of Minor Invasive Treatment

Cancer Center, Keio University School of Medicine, Tokyo 160-8582, Japan

³Department of Surgical Oncology, The University of Tokyo, Tokyo 113-8654, Japan

⁴Fujii Memorial Research Institute, Otsuka Pharmaceutical Company, Limited, Shiga 520-0106, Japan

⁵Co-first author

*Correspondence: t.sato@keio.jp

<http://dx.doi.org/10.1016/j.stem.2016.04.003>

SUMMARY

Colorectal tumor is a heterogeneous disease, with varying clinical presentation and prognosis in patients. To establish a platform encompassing this diversity, we generated 55 colorectal tumor organoid lines from a range of histological subtypes and clinical stages, including rare subtypes. Each line was defined by gene expression signatures and optimized for organoid culture according to niche factor requirements. *In vitro* and *in xenografts*, the organoids reproduced the histopathological grade and differentiation capacity of their parental tumors. Notably, we found that niche-independent growth is predominantly associated with the adenoma-carcinoma transition reflecting accumulation of multiple mutations. For matched pairs of primary and metastatic organoids, which had similar genetic profiles and niche factor requirements, the metastasis-derived organoids exhibited higher metastatic capacity. These observations underscore the importance of genotype-phenotype analyses at a single-patient level and the value of our resource to provide insights into colorectal tumorigenesis and patient-centered therapeutic development.

INTRODUCTION

Colorectal cancer (CRC) pathogenesis has been classically portrayed as the stepwise progression of cancerous lesions from potentially malignant precursors, predominantly tubular adenomas (Fearon and Vogelstein, 1990). Intensive studies have uncovered genetic alterations involved in CRC carcinogenesis, including recurrent genetic mutations, copy number variations, chromosomal changes, and epigenetic alterations (Cancer Genome Atlas Network, 2012; Seshagiri et al., 2012; Wood et al., 2007). The spectra of these genetic modifications are

highly variable across the tumors. However, a recent large-scale deep sequencing study revealed that the majority of the recurrently mutated genes in sporadic CRCs are ascribed to five common signaling pathways: WNT, RAS/mitogen-activated protein kinase (MAPK), phosphoinositide 3-kinase (PI3K), transforming growth factor (TGF), and TP53 signals (Cancer Genome Atlas Network, 2012). Interestingly, most of these driver pathways are coincidentally essential to regulate the self-renewal of the human intestinal stem cells (ISCs), in which the surrounding niche environments either activate or inactivate these pathways *in vivo* (Medema and Vermeulen, 2011). In a cancer tissue, aberrant driver pathway signals delimit the niche-restricted growth of the cancer cells and permit their dominant overgrowth in the hostile environments of the remote tissues they invade or metastasize. However, because of the lack of a biological analysis platform, functional niche dependency of patient-derived CRCs is largely uncertain.

Recently, by exploiting the understanding of stem cell niche factors, a novel 3D ISC culture system has been developed (Sato et al., 2009). This methodology enabled single ISCs to reconstitute stereotypic structures resembling intestinal epithelial crypts, referred to as organoids. The human colonic organoid culture uses a laminin-rich extracellular matrix (Matrigel) and the specific niche factors epithelial growth factor (EGF), Noggin, R-spondin1, and Wnt3A, together with a TGF- β inhibitor (A83-01) and a p38 inhibitor (SB202190) (Sato et al., 2011). The niche factor-based culture system has also been successfully applied to patient-derived CRCs, and the majority of the CRCs successfully underwent long-term culture with the less stringent niche factor combination of EGF, Noggin, A83-01, and SB202190 (ENAS) (Sato et al., 2011). This niche dependency of the CRC organoids suggests that cancer tissues inherently embrace the stem cell basis similarly to normal intestinal epithelium, supporting the caricature models (Pierce and Speers, 1988). Although the robust establishment efficiency of the colonic organoids has been confirmed in several studies, we found that the propagation of some samples, mainly tissues from rare subtypes of colonic tumors, were not amenable to the published culture protocol. This limitation of the original culture protocol might be attributed to the altered niche requirements resulting from adaptation to the

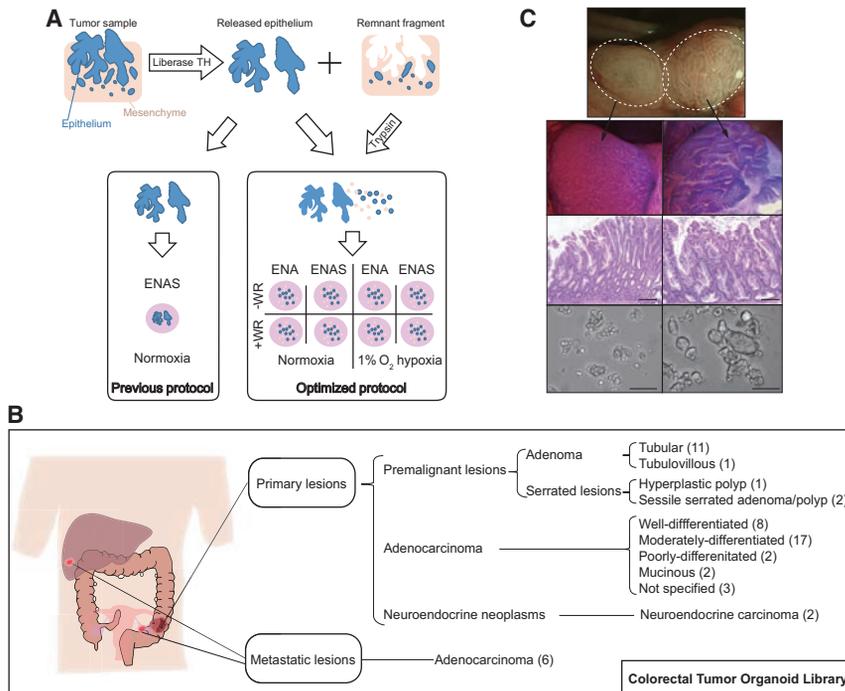


Figure 1. Establishment of Organoid Lines from Colorectal Tumors with Refined Protocol

(A) Diagram of optimized establishment methods. A, A83-01; E, EGF; N, Noggin; S, SB202190; W, Wnt3A; R, R-spondin1.

(B) Composition of the CTOL. Each figure represents the number of organoids derived from each tumor.

(C) Endoscopic images and histology of a continuum lesion consisting of HP (left) and SSAP (right). Narrow band imaging (top) and chromoendoscopic view (the second from top) with magnification. Bottom pictures show organoids established from each component. The scale bars represent 200 μm. See also Tables S1 and S2.

hostile tumor environments. These findings prompted us to investigate the niche factor alterations during colorectal tumorigenesis in detail and to derive colorectal tumor organoids with broader ranges of histological subtypes and clinical stages.

In this study, we first refined the culture protocol on the basis of the niche factor requirements and subsequently established a colorectal tumor organoid library (CTOL). Detailed molecular characterization and xenotransplantation analyses of the CTOL validated faithful ex situ modeling of diverse patient-derived colorectal tumors. Furthermore, we elucidated both functional linkages and gaps between the genetic alterations, niche requirements, and biological phenotypes of the tumors. The CTOL will offer a versatile platform for biomedical research including drug screening, pharmacogenetics, targeted therapy and personalized medicine.

RESULTS

Establishment of a CTOL with Improved Culture Protocol

With the previously published protocol (Jung et al., 2011; Sato et al., 2011), we successfully established organoids from human normal colorectal epithelium and from colorectal adenomas. In contrast, organoid propagation efficiency from CRC specimens diverged drastically among the samples. Tissues derived from rectal cancer and advanced-stage colon cancer were prone to insufficient isolation or attenuated tumor cell growth. We attributed these failures primarily to the infiltration of the cancer cells into the deep layers of the intestinal wall, which is the hallmark of invasive CRCs. Contrary to intestinal adenoma cells, which were readily detached from the basal membrane by enzymatic treatment, invasive CRC cells often remained with connective tissues even after digestion. We found that infiltrated cancer cells could

be released by additional dissociation of the remaining fragments with trypsin digestion medium (data not shown). Substantial numbers of epithelial cells were isolated using this method. However, some CRC cells still failed to expand in the original ENAS condition. Further testing of various culture conditions demonstrated the profound impact of three ambivalent factors, Wnt activators (Wnt3A/R-spondin1), oxygen concentration, and a p38 inhibitor, on CRC organoid proliferation. These factors were essential for a subset of CRC organoids, whereas others showed adverse or no sensitivity to these factors. These effects were unpredictable until the establishment of the organoids. Hence, eight combinatorial culture conditions were routinely allocated for isolated CRC cells at culture initiation (Figure 1A). With these improvements, we were able to propagate organoids from CRC tissues with 100% efficiency, except for cases with bacterial contaminations or poor-quality samples such as tissues with massive necrosis.

This result facilitated the systematic establishment of organoids from endoscopic biopsy specimens or from surgically resected samples of colorectal neoplasms. Consequently, a CTOL consisting of 55 organoids derived from 52 tumors and 43 patients was established (Figure 1B; Table S1). We also established 41 counterpart organoids from normal colorectal mucosa. For each sample, whole specimen was used for culture, and histological subtype was determined from H&E staining prepared from the adjacent region. Tumor samples with biomedical significance, such as metastatic tumors, or rare histological subtypes were preferentially collected. Thus, the CTOL included several colonic tumors that had rarely been available as cell lines. Hyperplastic polyps (HPs) and sessile serrated adenoma/polyps (SSAPs) are characterized by dysplastic-serrated histology and a high frequency of *BRAF* mutation, but the latter should be distinguished for their predisposition to carcinoma (Juspeert et al., 2015). Recent advances in high-resolution narrow band imaging and magnifying pit pattern enabled endoscopic diagnosis and demarcation of the two lesions prior to histological examination (Figure 1C) (Uraoka et al., 2015). We obtained biopsy samples from a continuum lesion consisting of HP and SSAP and independently established pathologically different organoids

from each component. An SSAP organoid line from a different patient was additionally established, totaling 3 HP/SSAP organoids in the CTOL. We also established organoids from 12 adenomas with different etiologies; 9 were derived from patients with familial adenomatous polyposis (FAP), and 3 were from sporadic adenomas. Ten adenomas were established from conventional tubular adenomas, whereas 1 was from a tubulovillous adenoma (TVA), a less frequent adenoma histological subtype. Derivation of cell lines from these benign colonic tumors had been difficult because of their strict niche factor dependency. Considering their potential as precursor lesions, these benign tumor organoids provide useful biological resources in studying stepwise tumorigenesis (Prasetyanti et al., 2013).

We successfully established 40 CRC organoids, of which 3 required Wnt activators, 5 favored hypoxia, and 9 exhibited significant growth reduction with SB202190 treatment. Collectively, 32.5% (13 of 40) of the established CRC organoids were unable to propagate in the original ENAS condition with atmospheric oxygen concentrations, underscoring the improved establishment efficiency in the new culture conditions. Thirteen organoids were derived from right-sided colon cancers, 7 were from left-sided colon cancers, and 11 were from rectal cancers, indicating that CRC organoids could be established regardless of their original anatomical sites. The library included 6 organoids from metastatic sites. Notably, organoids from rare histological subtypes of CRCs, such as poorly differentiated adenocarcinoma, mucinous adenocarcinoma, and neuroendocrine carcinoma (NEC), were successfully established.

Detailed patient clinical information was available for most tumor organoids in the library (Table S1). Sustained culture for longer than 3 months was confirmed for all organoids, and the organoids were stored as frozen stocks for on-demand use.

Colorectal Tumor Organoids Preserve the Histopathological Structures of Parental Tumors

Taking advantage of the diverse histological subtypes in CTOL, we questioned the extent to which CRC organoids recapitulate their original tumor histology *in vitro* and *in vivo*. For reconstruction of the CRCs *in vivo*, representative organoids were xenotransplanted into the kidney subcapsules of immunodeficient (NOD/Shi-scid, IL-2R γ null; NOG) mice as previously described (Matano et al., 2015) (Figure 2A). Notably, the histological grading and the differentiation patterns of the parental sites were highly recapitulated both *in vitro* and in the xenografted tumors. For example, CRC organoids derived from well-differentiated adenocarcinomas formed a mono-layered, cystic structure *in vitro* and reconstituted histology recapitulating the original parental tumors upon transplantation (Figures 2B–2D, S1A, and S1B). Likewise, moderately (Figures 2E–2G, S1C, and S1D) and poorly differentiated adenocarcinoma (Figures 2H–2J, S1E, and S1F) organoids formed similar morphologies both *in vitro* and *in vivo*. The organoids established from two rare clinical subtypes of CRC, NEC (Figures 2K–2M, S1G, and S1H) and mucinous adenocarcinoma (Figures 2N–2P, S1I, and S1J), also reproduced the histopathological features of the clinical specimens. NEC organoids exhibited positive staining of neuroendocrine markers, whereas mucinous adenocarcinoma organoids developed tumors with massive production of mucin.

For three CRC samples (CRC7, CRC20, and CRC28), patients' tumors, organoids, and xenografted tumors were additionally stained for CD44 and CK20 to evaluate whether differentiation hierarchies in parental tumors were preserved in the subsequent organoid culture and the xenograft assays (Figure S2). Except for one CRC sample (CRC7) with weak CD44 staining in the xenograft, the organoids and the xenografted tumors exhibited reciprocal expression patterns of CD44 and CK20 similar to the parental tumors.

Taken together, these data demonstrated that the histological subtypes as well as the differentiation hierarchies of CRCs were cell-intrinsically conserved regardless of their environmental context (i.e., in patients, *in vitro*, and in xenografts).

Gene Expression Profiling of Colorectal Tumor Organoids

To characterize the molecular features of the CTOL, 52 tumor organoids in the CTOL and 7 organoids established from healthy colon epithelium were subjected to gene expression microarray analysis (Table S2). In parallel, the microsatellite instability status was analyzed (Suraweera et al., 2002). Seven organoids were classified as microsatellite unstable, all of which were derived from right-sided colon cancers. Unbiased transcriptome analysis with principal-component analysis revealed discrete populations of normal, adenoma, serrated, microsatellite-instable (MSI) CRC, microsatellite-stable (MSS) CRC, and NEC organoids (Figure 3A). Unsupervised hierarchical consensus-based clustering of the CTOL gene signature provided the basis of a robust five-subtype classification associated with the histological features of colorectal tumors (Figure 3B; Table S3). Cluster 1 and cluster 4 were enriched with MSS CRC and adenoma organoids, respectively. HP/SSAP organoids were categorized into cluster 2, though 2 MSS CRC organoids were also classified into cluster 2. All MSI CRC and NEC organoids were distinctively categorized into clusters 3 and 5, respectively. To validate whether the organoid classifier could be applied to the molecular signatures of diseased tissue samples, we performed gene set enrichment analysis (GSEA) using the previous gene expression data sets from patient tissues (Valcz et al., 2014; Watanabe et al., 2012). GSEA confirmed the association of CTOL subtypes with the corresponding tumor subtypes in clinical samples (Figure 3C). The organoids were purely structured by the epithelial cells and were maintained in homogeneous culture environments. Thus, their transcriptome properties reflect the tumor cells' intrinsic gene signatures, contrary to the molecular signatures of fresh tissue samples that are profoundly influenced by non-epithelial cells and tissue conditions.

Using the CTOL gene signatures, we performed subtype analysis using the recently proposed consensus molecular CRC classifier, which underscored a poor prognosis subtype characterized by the enriched expression of genes related to epithelial-to-mesenchymal transition (EMT) or stromal infiltration (Guinney et al., 2015). On the basis of this classification system, no CTOL organoid was categorized into the EMT/stromal subtype (Figure 3D), whereas the MSI CRC organoids were properly assigned to the MSI subtype with 100% accuracy. This result was consistent with the other recent reports demonstrating that fibroblast contamination rather than the enrichment of the EMT phenotype in CRC was responsible for the EMT/stromal signature (Calon

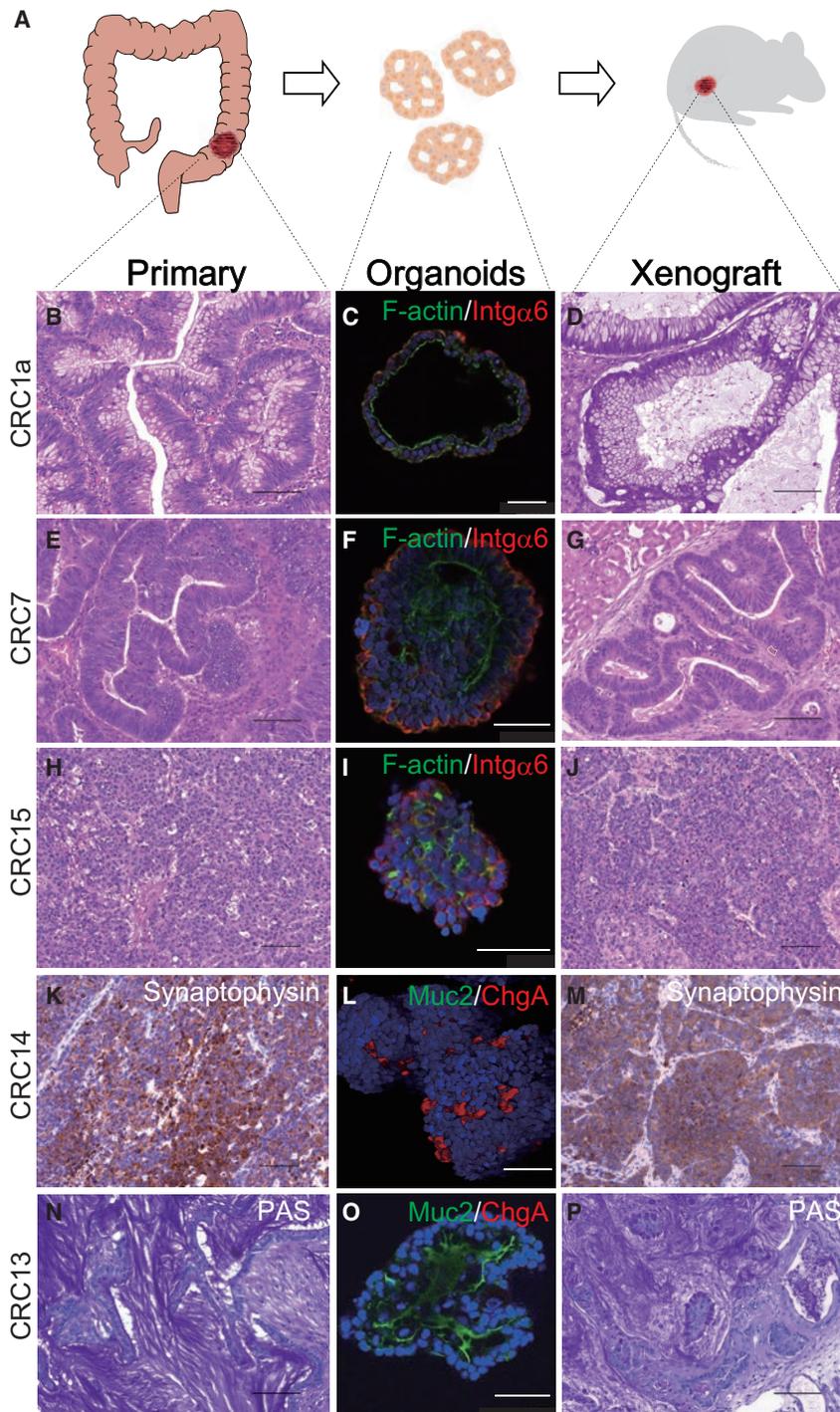


Figure 2. Organoids Reconstitute the Histology and Morphology of Primary Tumors In Vitro and In Vivo

(A) Schematic depicting the isolation of primary CRC samples, derivation of organoids, and xenotransplanted tumors.

(B–P) Representative images of parental tumors (B, E, H, K, and N), organoids (C, F, I, L, and O) and xenografts (D, G, J, M, and P) derived from well-differentiated adenocarcinoma (B–D), moderately differentiated adenocarcinoma (E–G), poorly differentiated adenocarcinoma (H–J), NEC (K–M), and mucinous adenocarcinoma (N–P). Epithelial polarity was visualized using phalloidin (green, F-actin) and integrin $\alpha 6$ (red, Intg $\alpha 6$). Differentiation markers chromogranin A (red, ChgA) and mucin2 (green, Muc2) were immunostained. Nuclear counterstaining (blue, Hoechst33342). The black scale bar represents 100 μ m; the white scale bar represents 50 μ m.

See also Figures S1 and S2.

Genetic Profiling of Colonic Tumor Organoids

We subsequently assessed the genomic alterations in the CTOL. Deep sequencing (targeted resequencing of 485 reported recurrently mutated genes in CRC [Table S4] or whole-exome sequencing) provided a detailed profile of genetic mutations in the CTOL (Figure 4A; Tables S2 and S5). *APC* mutations were identified in all adenoma organoids and in the majority of the CRC organoids. Nine of 12 adenoma organoids lacked other driver pathway mutations, whereas the other 3 adenoma organoids harbored well-known mutations, such as *KRAS* and *FBXW7*. *BRAF* mutations were present in all HP/SSAP organoids and in four of seven MSI CRCs. *TP53* mutations were found throughout the CRC organoids but were uncommon in the organoids from benign lesions. Overall, CRC organoids exhibited diverse patterns of genetic mutations, yet the mutation spectrum was consistent with that of The Cancer Genome Atlas (TCGA) database. Concordance of the mutation profiles between the CRC organoids and the original tumors has been

documented previously (van de Wetering et al., 2015), and these results suggested that the colorectal tumors could derive organoids irrespective of specific mutations.

Comparative genomic hybridization/SNP microarray analysis revealed karyotypes and somatic copy number aberrations (Figure 4B). All FAP-adenoma and HP/SSAP organoids were diploid, whereas a few chromosomal alterations were observed in sporadic adenoma organoids. We found well-defined arm-level changes, represented as gains of 13q and 20p and losses of

et al., 2015; Isella et al., 2015). Because the organoid culture system is devoid of mesenchymal cells, there remains a possibility that fibroblasts render EMT phenotype in the epithelial cells. To address this question, new methodologies such as organoid co-culture with fibroblasts may prospectively be essential.

Collectively, these results suggest that the innate molecular signatures of the colorectal tumors are ruled predominantly by the genetic program of the diseased epithelium, independent of the tissue environments.

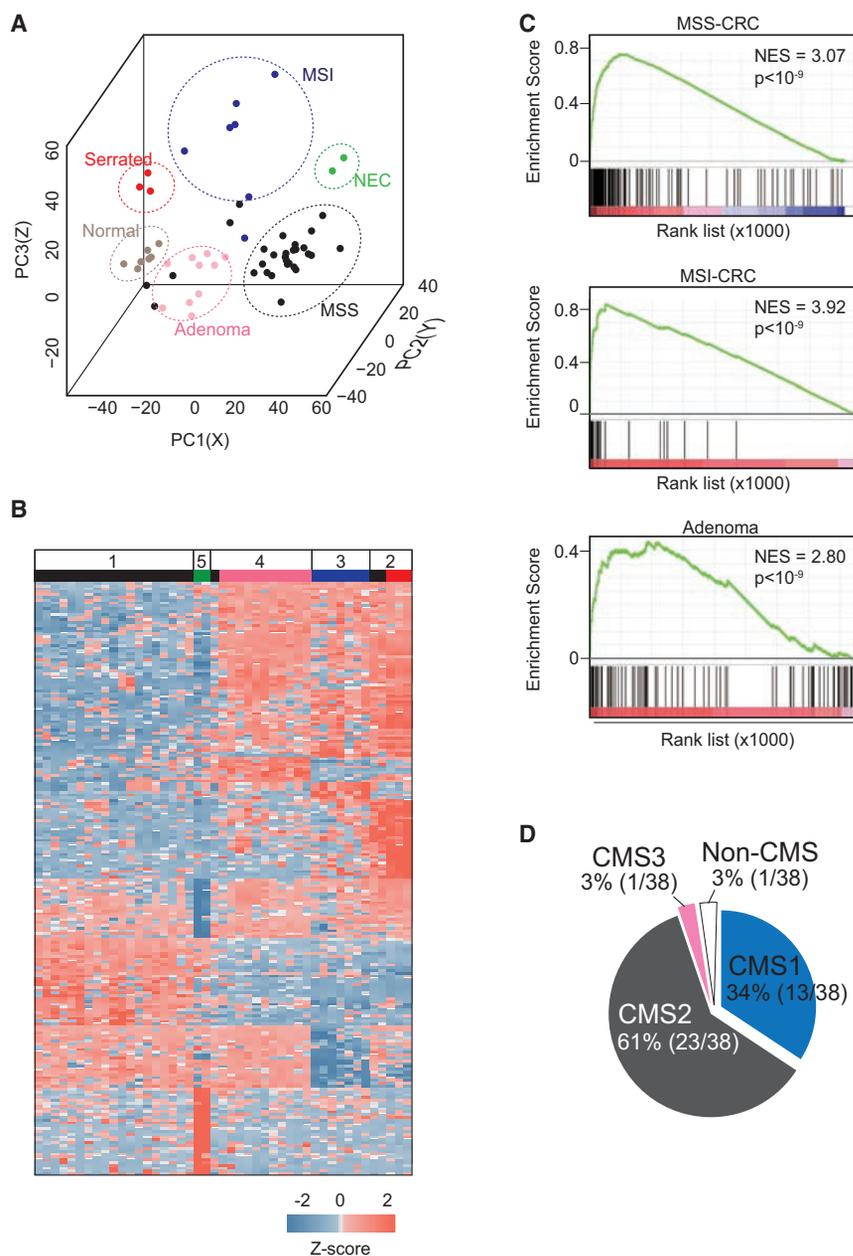


Figure 3. Gene Expression Analysis of the CTOL

(A) Principal-component analysis of the CTOL demarcates normal, adenoma, serrated lesions, MSI CRCs, MSS CRCs, and NECs.

(B and C) The CTOL tumor gene set is classified into five subtypes according to the classifier. GSEA confirmed that gene sets in the classifier represent the gene signatures of tissue samples derived from the corresponding subtypes.

(D) The consensus molecular subtypes in CTOL. Color code in (A) and (B): beige, normal; black, MSS; green, NEC; pink, adenoma; blue, MSI; red, serrated lesions. See also Tables S2 and S3.

in counterpart normal organoids. Therefore, these data suggested that the inherent CpG island methylation patterns of the tumors were sustained during in vitro organoid culture.

Genetic Characterization of Tumor Heterogeneity

The detailed genetic mutation data of the CTOL and the histopathological reproducibility of the organoids prompted us to define the genetic variance shaping intra-tumor heterogeneity. To elaborate the extent of intra-tumor genetic heterogeneity, genetic mutation properties were compared between the organoids derived from independent lesions of the same individuals. The exome sequence data of HP and SSAP organoids from a continuum lesion revealed somatic mutations acquired during neoplastic tumor progression. Respectively, 100 and 145 non-synonymous mutations were detected in HP and SSAP organoids, 31 of which were common (Table S5). The shared mutations included the *BRAF* hot spot mutation (V600E), providing evidence that these tumors were derived from a common origin.

18p and 17p, in 22 of 27 MSS CRC organoids, whereas the other 5 MSS CRC organoids were diploid (Figure 4B). MSI CRC organoids exhibited near-diploid karyotypes and harbored 713–2,634 non-synonymous mutations. The *hMLH1* mutation was identified in 1 of 7 MSI CRC organoids (Figure 4A), which was derived from a patient diagnosed with Lynch syndrome. The reference organoid established from the normal counterpart also had a heterozygous mutation in *hMLH1*, confirming the diagnosis. The remaining 6 MSI CRC organoids showed reduced expression of *hMLH1*, presumably resulting from promoter methylation, as all these organoids presented the CpG island methylator phenotype (CIMP) (Weisenberger et al., 2006). Of note, CIMP was also detected in all HP/SSAP organoids. We never observed alteration of the CpG island methylation profile

Additionally, we established a line of MSI CRC organoid from a clinical specimen with histopathological heterogeneity consisting of moderately and poorly differentiated adenocarcinoma (Figures 4C and 4D). Distinctively different morphologies of the organoids enabled us to isolate histologically unique subclones. After subcloning, subclones CRC17a (moderately differentiated adenocarcinoma) and CRC17b (poorly differentiated adenocarcinoma) stably maintained their structures in vitro (Figure 4D). The morphological phenotypes were still maintained even after xenotransplantation, reinforcing the inheritance of the histological subtype within each subclone (Figure 4D). Exome sequencing of these subclones identified 718 shared non-synonymous mutations, whereas 1,400 and 1,916 unique mutations were observed in CRC17a and CRC17b organoids,

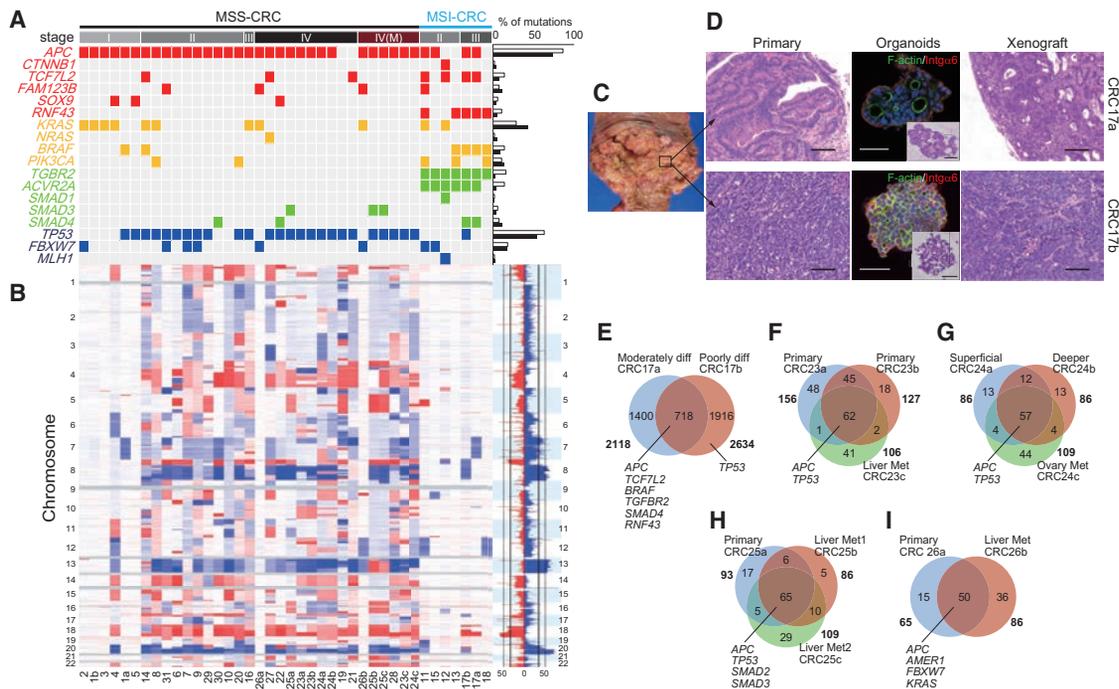


Figure 4. Genomic Alterations of CRC Organoids in the CTOL

(A) Gene mutations of CRC organoids. The frequency of each gene mutation in the CTOL (white bar) and in TCGA database (black bar) is presented on the right. (B) Chromosomal alterations of CRC organoids. Blue represents gains, and losses are shown in red. (C and D) A heterogeneous tumor consisting of moderately and poorly differentiated adenocarcinoma yielded morphologically and phenotypically distinct organoids and xenografts reflecting each component. Macroscopic (C) and microscopic (D) images of the tumors are shown. The scale bars represent 100 μ m for tissue H&E sections and 50 μ m for organoids. (E) Venn diagram of mutations in each organoid. (F–I) Comparison of genetic mutations between primary and metastatic organoids from the same patients. Numbers of non-synonymous mutations are indicated. See also Figure S3 and Tables S4 and S5.

respectively (Figure 4E; Table S5). Interestingly, the majority of the driver gene mutations converged into the shared pool of mutations except for *TP53* mutation unique in CRC17b, which indicated that these subclones derived from a common precursor tumor, and their histological phenotypes diverged along with independent molecular evolutions. Our result first provided the evidence that genetic heterogeneity shaped the histological heterogeneity of a CRC, though further studies are necessary to identify the molecular changes responsible for histological heterogeneity.

We also established CRC organoid lineages from different tumor regions within the same individuals. CRC24a and CRC24b were established from a superficial lesion and from a deeper layer of the primary tumor, respectively. CRC24c was derived from a synchronous ovarian metastasis. These three subclones shared 57 non-synonymous mutations, including *APC* and *TP53* mutations, but COSMIC-tagged mutations were absent in the single-nucleotide variants (SNVs) uniquely found in each subclone (Figure 4F). Similarly, we observed shared driver gene mutations in the CRC23 triplet, the CRC25 triplet, and the CRC26 pair, each consisting of the primary and liver-metastasized lesion(s). Similar to CRC24, major driver gene mutations were common among the subclones. We found 41, 44, and 36 metastatic organoid-unique SNVs in CRC23, CRC25, and CRC26, respectively, none of which were reported in the

COSMIC database (Figures 4G–4I; Table S5). This result was consistent with the findings that the driver gene mutations were common between the sets of primary and metastatic CRCs, implying the driver gene mutations precede metastasis development (Brannon et al., 2014; Vakiani et al., 2012).

Variation of Genetic Mutation Profile and Cellular Heterogeneity during Organoid Culture

The provided deep sequencing data (Table S5) reflect genetic mutation distribution in the expanded organoid populations, and the culture duration at the time of genomic DNA extraction differs between the lines depending on several factors, such as initial sample size or growth speed of the organoids. Genetic stability in mouse normal intestinal organoids has been reported previously (Behjati et al., 2014), though the extent of mutational susceptibility of tumor organoids has not been investigated. To address this issue, three lines of CRC organoids (two MSI CRC organoids and one MSS CRC organoid) were expanded from single organoids to obtain isogenic populations and were subjected to exome sequencing before and after long-term culture (at least 6 months). The comparison of genetic mutations demonstrated each MSI CRC organoid line acquired 75 and 82 de novo non-synonymous mutations, respectively, whereas the MSS CRC organoids were distinctively stable (Figure S3A). This mutation burden in the MSI CRC organoids presumably

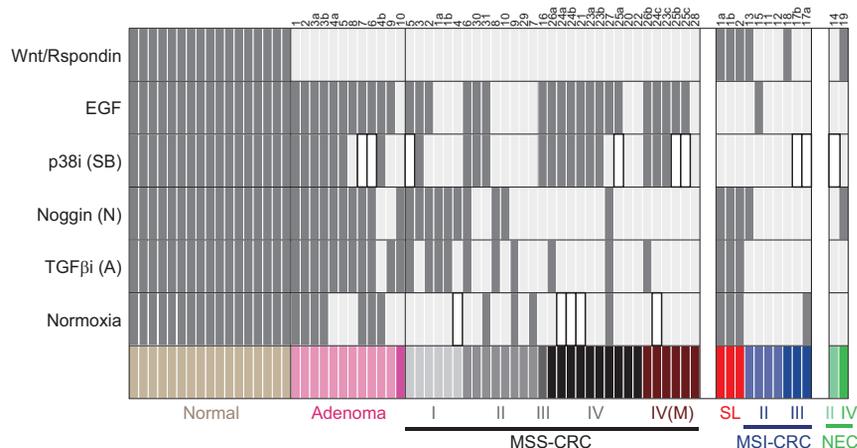


Figure 5. Each Organoid Displays a Unique Dependency on Niche Factors

Niche factor dependency of each organoid. Number codes of the organoids are demonstrated on top. Each bar represents the essential (dark gray), dispensable (light gray), and detrimental (white) factors in organoid growth. Organoids were sorted according to their clinical profiles: normal, adenoma, MSS CRC, SL (serrated lesion), MSI CRC, and NEC. The roman numerals represent the stages of the CRC organoids.

reflects their deficiency in the mismatch repair machinery and the propensity toward accumulating replication errors.

At the outset of culture, tumor organoids potentially include diverse genetic heterogeneity, yet it remains unclear to what extent organoids retains the genetic heterogeneity over the course of culture. To visualize the dynamics of cellular heterogeneity during the culture, we differentially labeled the paired CRC organoids (CRC17a/b, CRC23a/b, and CRC25b/c) with fluorescent markers and mixed equal numbers of cells to mimic the heterogeneous cell population in the parental tumors. The organoid admixtures were passaged every 10 days with equal cell density, and the proportion of each fluorescence area was monitored. Within 30–40 days, all three pairs were progressively dominated by single populations, indicating dynamic alteration of cellular occupancy during the organoid culture.

These results imply that genetic and cellular heterogeneity are chronologically modified, especially in the MSI CRC organoids, and mutational or functional analyses data must carefully be interpreted for the organoids expanded from bulk populations. For specific assays, generating clonal organoids from single cells at culture initiation or at early passages may be essential to derive minor organoid clones that are potentially lost along with culture.

Elucidation of the Niche Dependency of Tumor Organoids

The niche culture conditions for CRC organoid initiation included ENAS for all organoids, whereas some showed sensitivity to Wnt3A/R-spondin1 (WR) or SB202190. For further profiling of niche factor dependency, each factor was separately withdrawn from the culture media, and its impact on colony formation efficiency was evaluated. These steps were repeated until the minimal essential niche factor combination was extracted. As expected, niche factor requirement varied substantially among the organoids (Figure 5). All adenoma organoids propagated robustly in the absence of WR, whereas they all required EGF except for one TVA organoid that carried a *KRAS*^{G12V} mutation. In contrast, WR was necessary for the growth of HP/SSAP organoids, whereas EGF was dispensable.

Niche factor dependency was more diverse in the CRC organoids. Generally, MSS CRC organoids exhibited decreased dependency on Noggin, A83-01, and normoxic culture conditions, in line with clinical progression. Such deregulated niche factor

requirements implied the growth advantage of the tumor cells in hostile niche conditions, such as invasive fronts or metastatic environments that are often abundant in TGF- β /bone morphogenetic protein (BMP) ligands and hypoxic condition (Hanahan and Weinberg, 2011; Hardwick et al., 2008; Massagué, 2008). Although most of CRC organoids derived from liver metastasis (LM) grew with few niche factors, one LM CRC organoid displayed high niche dependency (EAS dependent). Furthermore, each pair of primary-metastatic CRC organoids in the CTOL exhibited concordant niche factor requirements, indicating that enhanced niche independence during clinical progression might not be an essential driver for invasive/metastatic process. In our previous report, the artificial introduction of pathway mutations in naive organoids conferred niche-independent growth, although the invasive or metastatic phenotype was not observed in engineered organoids (Matano et al., 2015). Taken together, altered niche requirements strongly contribute to the selective growth advantage within a hostile cancer environment. However, niche independence per se may not be essential for the tumor cells to invade or metastasize.

Wnt Niche Factor Requirements in Tumor Organoids

We subsequently investigated the extent to which each niche factor requirement was influenced by the molecular alterations. As mentioned previously, all adenoma organoids and the majority of the CRC organoids propagated in the absence of exogenous WR. These WR-independent organoids invariably carried mutations within the Wnt signaling pathway, including *APC*, *CTNNB1*, and *TCF7L2* (Figure 6A). We established three lines of WR-dependent CRC organoids. Interestingly, all of these organoids were derived from rare subtypes of CRC, mucinous adenocarcinoma, and NEC. Exome sequencing of these organoids revealed the absence of Wnt pathway mutation but detected other driver mutations by which the possibility of contaminated growth of normal organoids was excluded.

These data indicate that Wnt niche requirement is strictly regulated by genetic mutations within the Wnt pathway in colonic tumors. Approximately 95% of currently available CRC cell lines carry mutations in *APC*, *AXIN2*, *CTNNB1*, or *TCF7L2* (Mouradov et al., 2014), whereas the frequency in clinical CRC samples was reported to be approximately 80% (Cancer Genome Atlas Network, 2012). This non-negligible gap in frequency suggests the existence of Wnt-dependent CRCs, which are rarely established with the standard culture condition. Importantly, regardless of epithelial or non-epithelial Wnt, sensitivity to Wnt is a

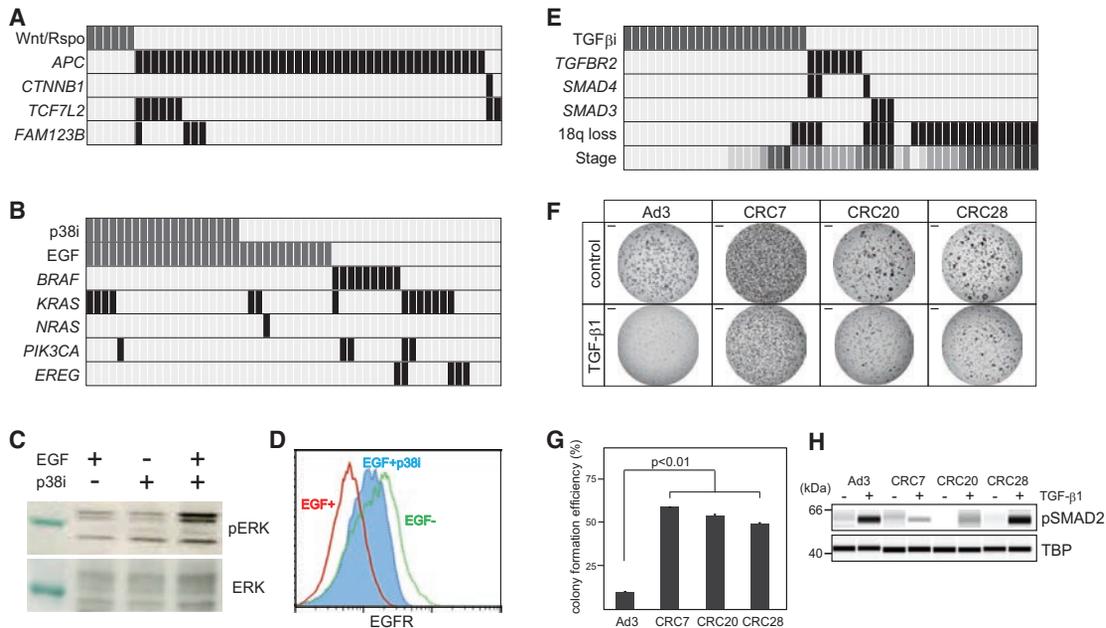


Figure 6. Dependency on Each Niche Factor Is Regulated by Diverse Mechanisms

(A) Comparison of dependency on Wnt (gray, dependent) and Wnt pathway mutations (black, mutated).

(B) Comparison of dependency on p38 inhibitor, EGF (gray, essential), and RAS/MAPK pathway mutations (black, mutated) and EREG gene expression (black, overexpression). Organoids with RAS/MAPK pathway mutations or EREG overexpression (Z score > 3.0) are independent of SB and EGF.

(C and D) p38 inhibitor modulates MAPK signaling by inhibition of EGFR negative feedback downregulation.

(E) Dependency on TGF- β inhibitor (gray, essential), TGF- β pathway gene mutations (black, mutated), chromosome 18q loss (black, deleted), and clinical stages (gray density represents clinical stage) of organoids are plotted.

(F) Images of TGF- β -treated organoids. The scale bar represents 500 μ m.

(G) Colony formation efficiency of A-independent organoids after TGF- β 1 treatment (Ad3, A-dependent; CRC7, CRC20, CRC28, A-independent). Values are means of technical triplicates. Error bars indicate SEMs.

(H) Elevated expression of nuclear phospho-SMAD2 in TGF- β 1-treated A-independent CRC organoids.

See also Figure S4.

potential but critical vulnerability of these Wnt-dependent CRCs, as they drastically ceased their proliferation upon withdrawal of exogenous WR in vitro (Figure S4A). This result strongly indicated that Wnt signaling could be a potential target in a small subset of CRCs, and further studies are required to validate the Wnt targeting therapy in CRC patients.

EGF Niche Requirements in Tumor Organoids

In the CTOL, 22 tumor organoids propagated in the absence of exogenous EGF, and MAPK pathway mutations were identified in 16 of these EGF-independent lines. In contrast, the other 6 EGF-independent CRC organoids were devoid of RAS/MAPK pathway mutations, and 2 of these lines were associated with upregulated *Epiregulin* (*EREG*) expression (Figure 6B). Notably, treatment with EGFR inhibitor dampened proliferation of the CRC organoids with *EREG* upregulation (Figure S4B), indicating that the autocrine/paracrine production of growth factors activates their own niche signaling. Mutations in the RAS/MAPK pathway, including *KRAS*, *NRAS*, and *BRAF*, were expected to signify EGF independency. However, 6 lines of 14 *KRAS*-mutant CRC organoid lines remained dependent on EGF. This result was consistent with our previous finding that clustered regularly interspaced short palindromic repeats (CRISPR)-engineered *KRAS*^{G12V} organoids gained independence from EGF but

required EGF for their optimal growth (Matano et al., 2015). These data indicated that RAS/MAPK pathway mutations alone were insufficient to compensate for the absence of EGF stimulus, and other molecular alterations, such as mutations in the PI3K pathway, were required.

p38 MAPK is activated by exogenous or endogenous stresses, such as reactive oxygen, DNA damage, or inflammation (Ono and Han, 2000; Wagner and Nebreda, 2009). Although p38 signal inhibition is essential for the maintenance of normal human ISCs, the role of p38 signaling in colorectal tumorigenesis remains a subject of controversy (Dolado et al., 2007; Urosevic et al., 2014). In the CTOL, 20 organoid lines required a p38 inhibitor (SB) for stable culture, but the inhibitor negatively affected the growth of 9 lines (Figure 5). SB-dependent organoids invariably required exogenous EGF, whereas most of the SB-independent organoids were also EGF independent (Figure 6B), indicating the indispensable role of SB as a counterpart of exogenous EGF stimulus. Consistent with this finding, *KRAS* or *BRAF* mutant organoids were mostly independent of p38 inhibition (Figure 6B). Furthermore, CRISPR-engineered *KRAS*-mutant organoids grew in the absence of SB (data not shown). Interestingly, in SB-dependent organoids, a combination of EGF and SB efficiently enhanced ERK phosphorylation in comparison with the organoids treated with EGF alone (Figure 6C). To further

investigate the molecular mechanism by which SB modulates EGF signaling, we assessed surface EGFR expression. In keeping with the previous study (Goh et al., 2010), EGFR was internalized after EGF ligand stimulation through negative feedback regulation, whereas this downregulation was significantly rescued by the addition of SB (Figure 6D). These results indicated that the requirement for SB is modulated through EGFR activation in organoids and p38 signal inhibition functions as an EGF receptor-specific stabilizer in the context of colonic epithelium.

Taken together, the requirement for exogenous EGF and p38 inhibition was predominated by pathway mutations or the overexpression of a cognate growth factor. Insufficient compensation of EGF signaling by *KRAS* mutation suggested that a subset of *KRAS*-mutant CRCs, particularly with intact PI3K pathway, might receive benefit from EGFR targeting therapy. Dependency on autocrine/paracrine secretion of cognate growth factors may also render the tumor cells sensitive to targeted inhibition of relevant molecules. Considering that EGF signaling is a key target in CRCs at present, drug sensitivity testing in vitro using organoids provides an affordable but bona fide tool for the validation and prediction of EGF targeted therapy.

TGF- β Inhibitory Niche Requirements in Tumor Organoids

TGF- β pathway inhibition is recognized to be involved in the malignant progression of CRC. We observed that 29 CRC organoid lines grew robustly in the absence of TGF- β inhibitor (A83-01). Of these, only 11 CRC organoids harbored *SMAD3*, *SMAD4*, or *TGFBR2* mutations, whereas the known TGF- β pathway mutations were not identified in the remaining 18 organoids (Figure 6E). TGF- β tolerance related to pathway mutations was particularly evident in MSI CRC organoids, among which *TGFBR2* mutations were highly recurrent. In contrast, TGF- β tolerance in most MSS CRC organoids was irrelevant to pathway mutations, but was significantly associated with loss of chromosome 18q ($p < 0.01$, Fisher's exact test) (Figure 6E). Similarly, requirement for BMP inhibition (Noggin) was acquired partially through *SMAD4* mutation, but the majority of Noggin-dependent organoids lacked relevant pathway mutations (Figure S4C).

We next interrogated for the difference of sensitivity to TGF- β between A83-01 (A)-dependent and A-independent MSS CRC organoids that were devoid of TGF- β pathway mutations. To test TGF- β resistance of these A-independent organoids in vitro, colony formation efficiency after TGF- β 1 treatment was quantified for three A-independent organoid lines without pathway mutations (CRC7, CRC20, and CRC28). As expected, the control A-dependent organoids showed significant reduction of colony formation after TGF- β 1 treatment, whereas A-independent organoids exhibited moderate recovery after TGF- β 1 treatment (Figures 6F and 6G). These A-independent organoids displayed upregulation of nuclear phospho-SMAD2 after TGF- β 1 treatment (Figure 6H), indicating TGF- β resistance of these organoids results from aberration of downstream targets.

The detailed molecular mechanism of TGF- β tolerance in MSS CRC organoids remains unclear. However, dispensability of TGF- β inhibition was highly associated with advanced clinical stages (stages III and IV versus others; $p < 0.01$, Fisher's exact test) (Figure 6E), suggesting that the TGF- β independence might be a hallmark of clinical progression.

Tumorigenic Capacity of Colorectal Tumor Organoids In Vivo

Establishment of in vivo preclinical models is imperative to translate organoid-based research. However, the application of the CRC organoids to relevant in vivo mouse models has not been comprehensively investigated. To address this issue, 5 benign lesion and 21 CRC organoid lines were xenotransplanted into the kidney subcapsules of NOG mice (Figure 7A). We observed successful tumor formation in all CRC cases, whereas benign tumor organoids exhibited no or substantially minimal engraftment. Interestingly, the size of the engrafted subrenal CRC tumors correlated with niche factor requirement (Figures 7B and 7C). This result was consistent with our previous finding that the number of CRISPR-introduced niche signal mutations correlated with tumorigenic potential (Matano et al., 2015). As mentioned previously, clinical progression does not necessarily parallel niche factor dependencies. Our results suggest that niche signal alterations confer a localized selective growth advantage but do not reflect clinical progression.

Next, to gain insight into the metastatic capacity of patient-derived CRC organoids, we performed splenic injection of the organoids derived from primary or metastatic lesions of stage IV MSS CRCs, and the development of liver metastases was evaluated (Figure 7D). Unlike the robust engraftment efficiency in subrenal capsules, metastatic capacity of the spleen-injected organoids was substantially diverse. Of nine transplanted organoids, seven organoids formed macrometastatic colonies (≥ 1 mm in size) in the livers, whereas the others failed to colonize or formed only micrometastatic lesions. Histological features of the parental tumors were well preserved in these macrometastatic lesions (Figure S5A). Interestingly, we found that the organoids derived from metastatic lesions (mCRC organoids) exhibited significantly higher metastatic potential in the xenograft model than the organoids from primary tumors, a trait that was even more evident between the pairs of primary and metastatic organoids (Figures 7E and 7F). As previously described, driver pathway mutations, chromosomal alterations, and niche dependency were indistinguishable between the primary CRC and mCRC organoids (Figures 4 and 5). Hierarchical clustering in a gene expression microarray underscored the alliance of gene signatures between primary and matched metastatic CRC organoids (Figure S5B). The primary-metastatic CRC pairs of organoids were cultured in identical microenvironments, whereas tissue samples used in previous studies could be influenced by tissue microenvironments, such as colon versus liver (Koh et al., 2008; Ramaswamy et al., 2003).

Taken together, these data suggest that metastatic CRC organoids are innately competent to colonize in the liver, and this metastatic capacity might be governed by biological programs independent of genetic mutations or niche factor requirements. The gene expression patterns of mCRC organoids remained similar to counterpart primary CRC organoids. However, further studies are required to determine whether mCRC organoids elicit a different response during the metastatic process.

DISCUSSION

In the present study, we have established a CTOL using an improved culture method with eight niche factor conditions.

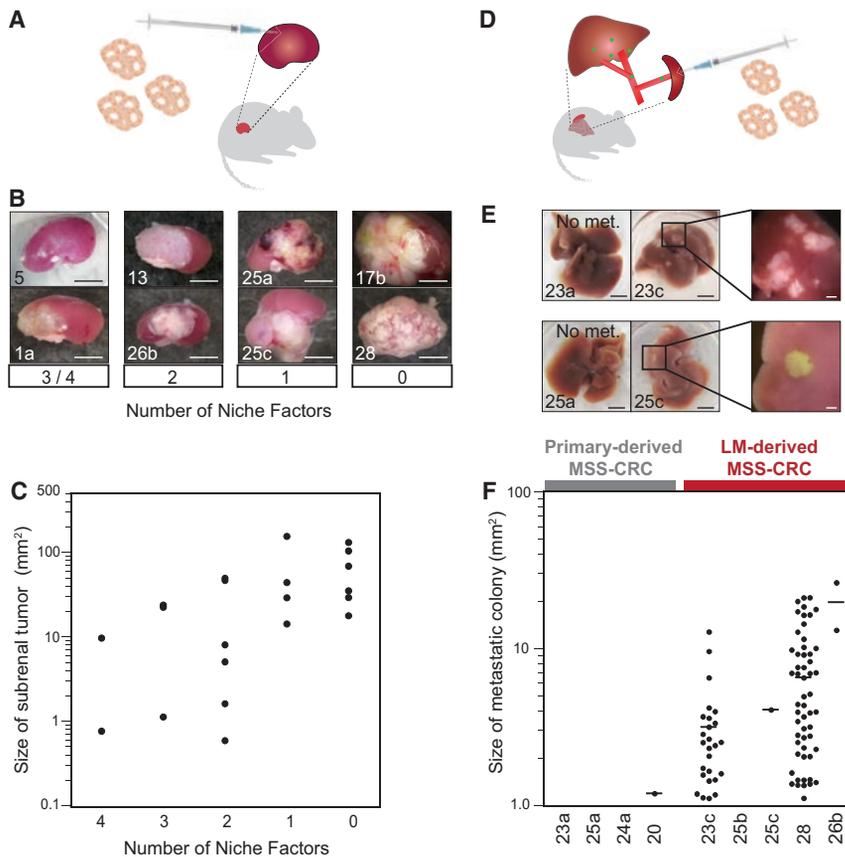


Figure 7. Xenotransplantation Assay of Organoids

(A) Schematic depicting subrenal organoid transplantation.

(B) Representative images of xenografted tumors in the subrenal capsule. The scale bar represents 5 mm.

(C) Size of xenografted subrenal tumors. Each dot represents the average size of engrafted tumors. Tumorigenicity in subrenal capsules correlates with the number of essential niche factors (Spearman's correlation coefficient = -0.62 , $p < 0.01$).

(D) Schematic of the hepatic metastasis assay by splenic organoid injection.

(E) Representative images of liver metastases from indicated tumor organoids. The black scale bar represents 5 mm; the white scale bar represents 1 mm.

(F) Size of liver metastases. Names of transplanted organoids are shown at the bottom. Each organoid line was transplanted into three mice, and the size of each metastatic colony among three livers is represented as a dot. Average sizes of metastases are shown as black bars. See also Figure S5.

Establishment of a CRC organoid-biobank was recently reported (van de Wetering et al., 2015). However, the contents of the biobank were limited to 22 conventional primary colon cancer organoid lines and lacked clinical information. Rectal cancers were also absent in the biobank for clinical reasons (rectal cancer patients typically undergo radiotherapy). Nevertheless, rectal cancers comprise approximately 40% of CRCs, vary substantially in either etiology or treatment from colon cancers, and are associated with worse prognosis. Thus, the inclusion of rectal cancers as a biomedical resource is a definite advantage of the CTOL. Furthermore, clinical information pertinent to the CTOL refined the quality of the resource. With improved culture conditions, we successfully established tumor organoids that had been impossible or difficult to derive, such as NEC, TVA, HP, SSAP, mucinous adenocarcinoma, poorly differentiated adenocarcinoma, and metastatic CRC. These tumor organoids often displayed Wnt or hypoxia dependency, or SB sensitivity, which complicated their propagation in the previously reported conditions. Development of new therapies for cancers has been restricted by the difficulty involved in conducting clinical trials. The collection of rare diseases in the CTOL and their xenograft modeling might contribute to efficient screening of novel treatments prior to clinical trials.

We also observed decreased niche factor requirements in colorectal tumors during the normal-adenoma and adenoma-carcinoma transition. This change in the niche factor requirements was predominantly determined by genetic mutations and contributed to local tumorigenicity. In contrast, during devel-

opment into advanced CRCs, the conversion of biological behaviors, such as acquisition of BMP/TGF- β resistance or metastatic progression, was observed independently of genetic alterations. It is possible that advanced genetic analysis will discover novel genetic alterations accounting for the malignant progression, and the CTOL is a useful resource platform to connect genetics and biology.

EXPERIMENTAL PROCEDURES

Human Tissues

All samples used for organoid establishment and biological analyses were obtained from Keio University Hospital or The University of Tokyo Hospital with written informed consent. The study was approved by the ethical committees of both institutes. Healthy and neoplastic colonic tissues were obtained from either endoscopic biopsy or surgically resected specimens. The samples were kept in ice-cold PBS until processing.

Organoid Culture

Fresh tumor tissue samples were processed as previously described (Sato et al., 2011), with several modifications. The tissue was cut into small pieces, washed with ice-cold PBS at least ten times, and subsequently digested with Liberase (TH grade; Roche Life Science) for 60 min at 37°C with vigorous pipetting every 15 min. The remaining fragments were additionally treated with TrypLE Express (Invitrogen) at 37°C for 20 min. The supernatant was collected and centrifuged at 200 \times g for 3 min at 4°C. The cell pellet was suspended with Matrigel (growth factor reduced; BD Biosciences) and dispensed into two 48-well culture plates (25 μ l Matrigel/well). The basal culture medium for human intestinal organoids was prepared as recently described (Fujii et al., 2015). Advanced DMEM/F12 was supplemented with penicillin/streptomycin, 10 mM HEPES, 2 mM GlutaMAX, 1 \times B27 (Life Technologies), 10 nM gastrin I (Sigma), and 1 mM N-acetylcysteine (Wako). The following niche factors were used: 50 ng/ml mouse recombinant EGF (Life Technologies), 100 ng/ml mouse recombinant Noggin (Peprotech), 10% R-spondin-1 conditioned medium (Ootani et al., 2009), 50% Wnt-3A conditioned medium (Sato et al.,

2011), 500 nM A83-01 (Tocris), and 10 μ M SB202190 (Sigma). After Matrigel polymerization, the cells were overlaid with culture media representing four different niche factor conditions (WENRA, WENRAS, ENA, and ENAS). Each plate was separately incubated at 20% O₂ and at 1% O₂. Media were changed every 2 or 3 days, and the condition that allowed the most efficient growth of the organoids was determined as the primary optimal condition. Upon expansion, the organoids were passaged and cultured with media lacking each of the growth factors to determine the minimal essential niche requirement. The organoids were maintained in the minimal essential condition and were used for further analyses.

Organoid Xenotransplantation

All animal procedures were approved by the Keio University School of Medicine Animal Care Committee. NOD/Shi-scid, IL-2R γ null (NOG) mice (7–12 weeks old, male) were obtained from the Central Institute for Experimental Animals (Ito et al., 2002). The CRC organoids were labeled with GFP by piggyBAC transposon, as previously described (Fujii et al., 2015). The expanded organoids were harvested with Cell Recovery Solution (BD Biosciences) and suspended in Matrigel at a concentration of 1×10^4 cells/ml. A total 10 μ l (1×10^5 cells) of Matrigel-organoid suspension was injected into the left subrenal capsules or the spleens of NOG mice. The mice were euthanized 2 months after transplantation, and the left kidneys or the livers were isolated. The sizes of the engrafted or metastasized tumors were determined on the basis of the area of GFP fluorescence (Nikon Multi-zoom microscopy, LuminaVision software). Each CRC organoid line was transplanted into three or four independent NOG mice. Some samples were then formalin-fixed and embedded in paraffin for subsequent immunohistochemical analysis.

ACCESSION NUMBERS

The accession number for the expression microarray data reported in this paper is GEO: GSE74843.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and five tables can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.04.003>.

AUTHOR CONTRIBUTIONS

M.F., A.T., M.M., K.N., Y.N., and K.K. performed the establishment and characterization of the human colorectal tumor organoids. M.F. and S.D. performed xenotransplantation assays. M.S. and K.T. analyzed integrated molecular data. Y.O. performed immunohistochemistry. T.U., T.K., and T.W. provided tissue specimens. T.S. conceived and conducted the project. M.F. and T.S. wrote the manuscript.

CONFLICTS OF INTEREST

S.D. is an employee of Ohtsuka Pharmaceuticals.

ACKNOWLEDGMENTS

This work was supported by the Project for Development of Innovative Research on Cancer Therapeutics (P-Direct) and by the Research Center Network for Realization of Regenerative Medicine project from Japan Agency for Medical Research and Development (AMED), and by a Grant-in-Aid for Scientific Research on Innovative Areas "Stem Cell Aging and Disease" from the Ministry of Education, Culture, Sports, Science and Technology of Japan (2611507). We thank Xin Wang (City University of Hong Kong) for assistance with subtype classification. We also thank Y. Tanada for animal care and the Collaborative Research Resources, School of Medicine, Keio University, for technical assistance. The Wnt3A- and R-spondin-producing cell lines were kind gifts from H. Clevers (Hubrecht Institute) and C. Kuo (Stanford University), respectively.

Received: November 11, 2015

Revised: March 2, 2016

Accepted: April 14, 2016

Published: May 19, 2016

REFERENCES

- Behjati, S., Huch, M., van Boxtel, R., Karthaus, W., Wedge, D.C., Tamuri, A.U., Martincorena, I., Petljak, M., Alexandrov, L.B., Gudem, G., et al. (2014). Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422–425.
- Brannon, A.R., Vakiani, E., Sylvester, B.E., Scott, S.N., McDermott, G., Shah, R.H., Kania, K., Viale, A., Oswald, D.M., Vacic, V., et al. (2014). Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. *Genome Biol.* **15**, 454.
- Calon, A., Lonardo, E., Berenguer-Llargo, A., Espinet, E., Hernando-Mombalona, X., Iglesias, M., Sevillano, M., Palomo-Ponce, S., Tauriello, D.V., Byrom, D., et al. (2015). Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47**, 320–329.
- Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337.
- Dolado, I., Swat, A., Ajenjo, N., De Vita, G., Cuadrado, A., and Nebreda, A.R. (2007). p38 α MAP kinase as a sensor of reactive oxygen species in tumorigenesis. *Cancer Cell* **11**, 191–205.
- Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767.
- Fujii, M., Matano, M., Nanki, K., and Sato, T. (2015). Efficient genetic engineering of human intestinal organoids using electroporation. *Nat. Protoc.* **10**, 1474–1485.
- Goh, L.K., Huang, F., Kim, W., Gygi, S., and Sorkin, A. (2010). Multiple mechanisms collectively regulate clathrin-mediated endocytosis of the epidermal growth factor receptor. *J. Cell Biol.* **189**, 871–883.
- Guinney, J., Dienstmann, R., Wang, X., de Reyniès, A., Schlicker, A., Soneson, C., Marisa, L., Roepman, P., Nyamundanda, G., Angelino, P., et al. (2015). The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* **144**, 646–674.
- Hardwick, J.C., Kodach, L.L., Offerhaus, G.J., and van den Brink, G.R. (2008). Bone morphogenetic protein signalling in colorectal cancer. *Nat. Rev. Cancer* **8**, 806–812.
- Ijspeert, J.E., Vermeulen, L., Meijer, G.A., and Dekker, E. (2015). Serrated neoplasia-role in colorectal carcinogenesis and clinical implications. *Nat. Rev. Gastroenterol. Hepatol.* **12**, 401–409.
- Isella, C., Terrasi, A., Bellomo, S.E., Petti, C., Galatola, G., Muratore, A., Mellano, A., Senetta, R., Cassenti, A., Sonetto, C., et al. (2015). Stromal contribution to the colorectal cancer transcriptome. *Nat. Genet.* **47**, 312–319.
- Ito, M., Hiramatsu, H., Kobayashi, K., Suzue, K., Kawahata, M., Hioki, K., Ueyama, Y., Koyanagi, Y., Sugamura, K., Tsuji, K., et al. (2002). NOD/SCID/ γ mouse: an excellent recipient mouse model for engraftment of human cells. *Blood* **100**, 3175–3182.
- Jung, P., Sato, T., Merlos-Suárez, A., Barriga, F.M., Iglesias, M., Rossell, D., Auer, H., Gallardo, M., Blasco, M.A., Sancho, E., et al. (2011). Isolation and in vitro expansion of human colonic stem cells. *Nat. Med.* **17**, 1225–1227.
- Koh, K.H., Rhee, H., Kang, H.J., Yang, E., You, K.T., Lee, H., Min, B.S., Kim, N.K., Nam, S.W., and Kim, H. (2008). Differential gene expression profiles of metastases in paired primary and metastatic colorectal carcinomas. *Oncology* **75**, 92–101.
- Massagué, J. (2008). TGF β in Cancer. *Cell* **134**, 215–230.
- Matano, M., Date, S., Shimokawa, M., Takano, A., Fujii, M., Ohta, Y., Watanabe, T., Kanai, T., and Sato, T. (2015). Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat. Med.* **21**, 256–262.
- Medema, J.P., and Vermeulen, L. (2011). Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nature* **474**, 318–326.

- Mouradov, D., Sloggett, C., Jorissen, R.N., Love, C.G., Li, S., Burgess, A.W., Arango, D., Strausberg, R.L., Buchanan, D., Wormald, S., et al. (2014). Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. *Cancer Res.* **74**, 3238–3247.
- Ono, K., and Han, J. (2000). The p38 signal transduction pathway: activation and function. *Cell. Signal.* **12**, 1–13.
- Ootani, A., Li, X., Sangiorgi, E., Ho, Q.T., Ueno, H., Toda, S., Sugihara, H., Fujimoto, K., Weissman, I.L., Capecchi, M.R., and Kuo, C.J. (2009). Sustained in vitro intestinal epithelial culture within a Wnt-dependent stem cell niche. *Nat. Med.* **15**, 701–706.
- Pierce, G.B., and Speers, W.C. (1988). Tumors as caricatures of the process of tissue renewal: prospects for therapy by directing differentiation. *Cancer Res.* **48**, 1996–2004.
- Prasetyanti, P.R., Zimmerlin, C.D., Bots, M., Vermeulen, L., Melo, Fde.S., and Medema, J.P. (2013). Regulation of stem cell self-renewal and differentiation by Wnt and Notch are conserved throughout the adenoma-carcinoma sequence in the colon. *Mol. Cancer* **12**, 126.
- Ramaswamy, S., Ross, K.N., Lander, E.S., and Golub, T.R. (2003). A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* **33**, 49–54.
- Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., and Clevers, H. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265.
- Sato, T., Stange, D.E., Ferrante, M., Vries, R.G., Van Es, J.H., Van den Brink, S., Van Houdt, W.J., Pronk, A., Van Gorp, J., Siersema, P.D., and Clevers, H. (2011). Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762–1772.
- Seshagiri, S., Stawiski, E.W., Durinck, S., Modrusan, Z., Storm, E.E., Conboy, C.B., Chaudhuri, S., Guan, Y., Janakiraman, V., Jaiswal, B.S., et al. (2012). Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660–664.
- Suraweera, N., Duval, A., Reperant, M., Vaury, C., Furlan, D., Leroy, K., Seruca, R., Iacopetta, B., and Hamelin, R. (2002). Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* **123**, 1804–1811.
- Uraoka, T., Higashi, R., Horii, J., Harada, K., Hori, K., Okada, H., Mizuno, M., Tomoda, J., Ohara, N., Tanaka, T., et al. (2015). Prospective evaluation of endoscopic criteria characteristic of sessile serrated adenomas/polyps. *J. Gastroenterol.* **50**, 555–563.
- Urosevic, J., Garcia-Albéniz, X., Planet, E., Real, S., Céspedes, M.V., Guiu, M., Fernandez, E., Bellmunt, A., Gawrzak, S., Pavlovic, M., et al. (2014). Colon cancer cells colonize the lung from established liver metastases through p38 MAPK signalling and PTHLH. *Nat. Cell Biol.* **16**, 685–694.
- Vakiani, E., Janakiraman, M., Shen, R., Sinha, R., Zeng, Z., Shia, J., Cercek, A., Kemeny, N., D'Angelica, M., Viale, A., et al. (2012). Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J. Clin. Oncol.* **30**, 2956–2962.
- Valcz, G., Patai, A.V., Kalmár, A., Péterfia, B., Fúri, I., Wichmann, B., Múzes, G., Sipos, F., Krenács, T., Mihály, E., et al. (2014). Myofibroblast-derived SFRP1 as potential inhibitor of colorectal carcinoma field effect. *PLoS ONE* **9**, e106143.
- van de Wetering, M., Francies, H.E., Francis, J.M., Bounova, G., Iorio, F., Pronk, A., van Houdt, W., van Gorp, J., Taylor-Weiner, A., Kester, L., et al. (2015). Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933–945.
- Wagner, E.F., and Nebreda, A.R. (2009). Signal integration by JNK and p38 MAPK pathways in cancer development. *Nat. Rev. Cancer* **9**, 537–549.
- Watanabe, T., Kobunai, T., Yamamoto, Y., Matsuda, K., Ishihara, S., Nozawa, K., Yamada, H., Hayama, T., Inoue, E., Tamura, J., et al. (2012). Chromosomal instability (CIN) phenotype, CIN high or CIN low, predicts survival for colorectal cancer. *J. Clin. Oncol.* **30**, 2256–2264.
- Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M.A., Kang, G.H., Widschwendter, M., Weener, D., Buchanan, D., et al. (2006). CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat. Genet.* **38**, 787–793.
- Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J., et al. (2007). The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108–1113.

CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs

Mohammad A. Mandegar,^{1,*} Nathaniel Huebsch,^{1,2} Ekaterina B. Frolov,¹ Edward Shin,¹ Annie Truong,¹ Michael P. Olvera,¹ Amanda H. Chan,¹ Yuichiro Miyaoaka,^{1,12} Kristin Holmes,¹ C. Ian Spencer,¹ Luke M. Judge,^{1,2} David E. Gordon,^{3,4,5} Tilde V. Eskildsen,^{6,7} Jacqueline E. Villalta,^{3,4,8,9} Max A. Horlbeck,^{3,4,8,9} Luke A. Gilbert,^{3,4,8,9} Nevan J. Krogan,^{3,4,5} Søren P. Sheikh,^{6,7} Jonathan S. Weissman,^{3,4,8,9} Lei S. Qi,¹⁰ Po-Lin So,¹ and Bruce R. Conklin^{1,3,4,11,*}

¹Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158, USA

²Department of Pediatrics, University of California, San Francisco, San Francisco, CA 94158, USA

³Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA

⁴California Institute for Quantitative Biosciences, QB3, University of California, San Francisco, San Francisco, CA 94158, USA

⁵Gladstone Institute of Virology and Immunology, San Francisco, CA 94158, USA

⁶Department of Cardiovascular and Renal Research, University of Southern Denmark, 5000 Odense C, Denmark

⁷Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, 5000 Odense C, Denmark

⁸Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA

⁹Center for RNA Systems Biology, University of California, San Francisco, San Francisco, CA 94158, USA

¹⁰Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

¹¹Department of Medicine and Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA

¹²Present address: Regenerative Medicine Project, Tokyo Metropolitan Institute of Medical Science, Tokyo, 156-8506, Japan

*Correspondence: mo.mandegar@gladstone.ucsf.edu (M.A.M.), bconklin@gladstone.ucsf.edu (B.R.C.)

<http://dx.doi.org/10.1016/j.stem.2016.01.022>

SUMMARY

Developing technologies for efficient and scalable disruption of gene expression will provide powerful tools for studying gene function, developmental pathways, and disease mechanisms. Here, we develop clustered regularly interspaced short palindromic repeat interference (CRISPRi) to repress gene expression in human induced pluripotent stem cells (iPSCs). CRISPRi, in which a doxycycline-inducible deactivated Cas9 is fused to a KRAB repression domain, can specifically and reversibly inhibit gene expression in iPSCs and iPSC-derived cardiac progenitors, cardiomyocytes, and T lymphocytes. This gene repression system is tunable and has the potential to silence single alleles. Compared with CRISPR nuclease (CRISPRn), CRISPRi gene repression is more efficient and homogenous across cell populations. The CRISPRi system in iPSCs provides a powerful platform to perform genome-scale screens in a wide range of iPSC-derived cell types, dissect developmental pathways, and model disease.

INTRODUCTION

To understand the biological roles of genes in development and disease, we must decipher the relationships between genotype and phenotype. Until recently, RNAi has been the most commonly used loss-of-function tool to study human biology (Boettcher and McManus, 2015). However, RNAi suffers from off-target effects and incomplete silencing of the desired gene (Jackson et al., 2003; Kim et al., 2013b; Krueger et al., 2007).

Alternatively, programmable nucleases, such as zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), allow more precise gene editing in model organisms, particularly in mammalian and human systems (Gaj et al., 2013; Kim and Kim, 2014). While ZFNs and TALENs are efficient tools for targeting single alleles, they cannot be easily used for library-scale loss-of-function studies.

In 2012, clustered regularly interspaced short palindromic repeat (CRISPR) technology emerged as a new tool for gene editing. This technology is a microbial adaptive-immune system that uses RNA-guided nucleases to recognize and cleave foreign genetic elements (Doudna and Charpentier, 2014; Wiedenheft et al., 2012). The recently engineered CRISPR/Cas9 system consists of two components: a single-chimeric guide RNA (gRNA) that provides target specificity and a CRISPR-associated protein (Cas9) that acts as a helicase and a nuclease to unwind and cut the target DNA (Cong et al., 2013; Mali et al., 2013). In this system, the only restriction for targeting a specific locus is the protospacer adjacent motif (PAM) sequence ("NGG" in the case of *SpCas9*) (Doudna and Charpentier, 2014).

CRISPR nuclease (CRISPRn) has been used for genome-scale screens to identify essential genes for cell viability in cancer and embryonic stem cells (Shalem et al., 2014) and human leukemic cell lines (Wang et al., 2014, 2015). However, CRISPRn may not be the most robust system for loss-of-function studies, because it is limited by the number of cells within a population that do not produce knockout phenotypes (González et al., 2014). In addition, partial loss- or gain-of-function phenotypes can be generated by Cas9-induced in-frame insertion/deletions (INDELS) and hypomorphic alleles (Shi et al., 2015), which can obscure the readout.

The nuclease deactivated version of Cas9 (dCas9) blocks transcription in prokaryotic and eukaryotic cells (known as CRISPR interference; CRISPRi) (Qi et al., 2013). More recently, dCas9 was fused to the Krüppel-associated box (KRAB) repression domain to generate dCas9-KRAB, producing a

more efficient transcriptional interference (Gilbert et al., 2013, 2014; Kearns et al., 2014). To further this effort, we aimed to use CRISPRi technology to efficiently repress genes to study early differentiation and model disease with human induced pluripotent stem cells (iPSCs) (Takahashi et al., 2007).

iPSCs are well suited to study early embryonic development and disease since they can produce different functional cell types *in vitro* (Sterneckert et al., 2014). Early embryonic development consists of a series of accurately timed events that affect gene activation and repression (Bolouri and Davidson, 2003). Therefore, precisely regulating the timing and dosage of transcription factors critically affects embryonic development (McFadden et al., 2005; Takeuchi et al., 2011), and dysregulation in the timing and dosage of transcripts can lead to disease development (Theodoris et al., 2015). In this study, we compared inducible CRISPR systems for gene knockout (using Cas9) or knockdown (using dCas9-KRAB) to enable temporal control of loss-of-function phenotypes in iPSCs and differentiated cell types.

RESULTS

Generation of CRISPRi and CRISPRn iPSC Lines

For loss-of-function studies, we independently derived multiple stable CRISPRi and CRISPRn human iPSC clones in two genetic backgrounds: wild-type B (WTB) and wild-type C (WTC) (Miyaoaka et al., 2014). In separate targeting events, the CRISPRi and CRISPRn constructs (see Supplemental Experimental Procedures) were integrated into the AAVS1 locus of WTB and WTC iPSCs using a TALEN-assisted gene-trap approach (Figures 1A, 1B, and S1). Transgenes integrated at the AAVS1 locus remain transcriptionally active in both iPSCs and differentiated cell types (Hockemeyer et al., 2011; Lombardo et al., 2011). We generated several different versions of the CRISPRi system that are either inducible or constitutive; the inducible CRISPRi (Gen1 and Gen2) clones express dCas9-KRAB (KRAB domain fused at the N terminus) from the inducible TetO promoter, while the constitutive CRISPRi clones (Gen3) express dCas9-KRAB under the constitutively active CAG promoter. The CRISPRn (Gen1) clones express Cas9 under the inducible TetO promoter (Figure S1).

The average efficiency of forming stable clones was ~350 colonies per million iPSCs transfected with AAVS1 TALENs and donor plasmid (data not shown). From each condition, multiple independent colonies were isolated and expanded. A subset of the stable colonies from each targeting vector was screened using junction PCR. Two putative colonies from each targeting event were further characterized by stably introducing an *OCT4*-specific gRNA and performing knockdown or knockout assays with immunofluorescence and western blot analysis. All putative CRISPRi clones containing an *OCT4*-specific gRNA showed efficient knockdown (>95%) of *OCT4* in bulk populations, while a significant fraction of the CRISPRn cells remained *OCT4* positive (~30%–40%) in bulk populations containing *OCT4*-specific gRNA (Figure S1). One clone each from CRISPRi and CRISPRn (Gen1 lines in the WTC genetic background) were subsequently used as lead clones for further studies.

To enable non-invasive and high-throughput phenotypic analysis in iPSC-derived cardiomyocytes (iPS-CMs), we performed

a second targeting event that introduced the green fluorescent calcium-modulated protein 6 fast type (GCaMP) calcium sensor (Chen et al., 2013) into the other AAVS1 locus of the CRISPRi cell line. The GCaMP transgene is driven off the strong, constitutive CAG promoter (Figure S1). We found that CRISPRi iPSCs could differentiate into iPS-CMs, so that we could measure calcium transients based on the GCaMP-fluorescent intensity (Movie S1) (Huebsch et al., 2015). Lead CRISPRi and CRISPRn iPSCs were karyotypically normal (Figures S2A and S2B) and expressed pluripotency markers, as expected (Figures S2C and S2D).

RNA-sequencing (RNA-seq) analysis indicated that expression of dCas9-KRAB or Cas9 was undetectable in the absence of doxycycline, and addition of doxycycline without any gRNA resulted in robust selective induction of dCas9-KRAB or Cas9, while the rest of the transcriptome remained virtually unchanged (Figures S2E and S2F). Furthermore, the RNA-seq data suggest that the addition of the KRAB domain has no detectable off-target effects when compared to expression of Cas9 alone. Remarkably the one gene that appeared to be upregulated upon doxycycline induction (without gRNA) was the same gene (*Vimentin*; *VIM*) for both CRISPRi and CRISPRn cells (Figures S2E and S2F). Since the same gene is upregulated for CRISPRi and CRISPRn cells, we suspect it may represent an off-target activity of the doxycycline-induced transactivator. Importantly, our experiments suggest that the expression of dCas9-KRAB alone has no additional effects on gene expression.

We also expressed dCas9-KRAB and Cas9 by continuously culturing CRISPRi and CRISPRn lines with doxycycline for 3 weeks (four passages). With this long-term treatment, we observed no cytotoxicity, decrease in proliferation, or change in morphology in these cells (Figures S2G and S2H). Using a droplet digital PCR (ddPCR)-based copy-number assay, we measured the number of integration events (Figure S2I). We further validated on-target integration sites on the lead CRISPRi and CRISPRn clones with junction PCR (Figure S2J) and verified their sequences (data not shown).

To further ensure there was no leaky expression of the single doxycycline-inducible vector, we measured the protein levels of dCas9-KRAB and Cas9 in iPSCs. With immunostaining, flow cytometry and western blots did not detect dCas9-KRAB or Cas9 protein without doxycycline in either CRISPRi or CRISPRn iPSCs, indicating that the TetO promoter has high fidelity in the AAVS1 locus. After doxycycline treatment, all cells in the CRISPRi and CRISPRn lines expressed dCas9-KRAB or Cas9 within 48 hr, respectively (Figures 1C–1H). dCas9-KRAB and Cas9 were expressed at similar levels after induction, and both proteins rapidly degraded after removing doxycycline (Figures 1F, 1H, and S2K). These data showed that dCas9-KRAB and Cas9 expression could be tightly regulated with the TetO promoter, which would support studies that rely on precisely timing gene knockdown or knockout.

Comparison of Loss of Function between CRISPRi and CRISPRn

To compare CRISPRi and CRISPRn for loss-of-function studies, we designed a gRNA that targets the first exon of *NANOG*, a transcription factor necessary for maintaining the pluripotency network. We selected *NANOG* as our first target gene because its deficiency is sufficient to give an immediate readout, as

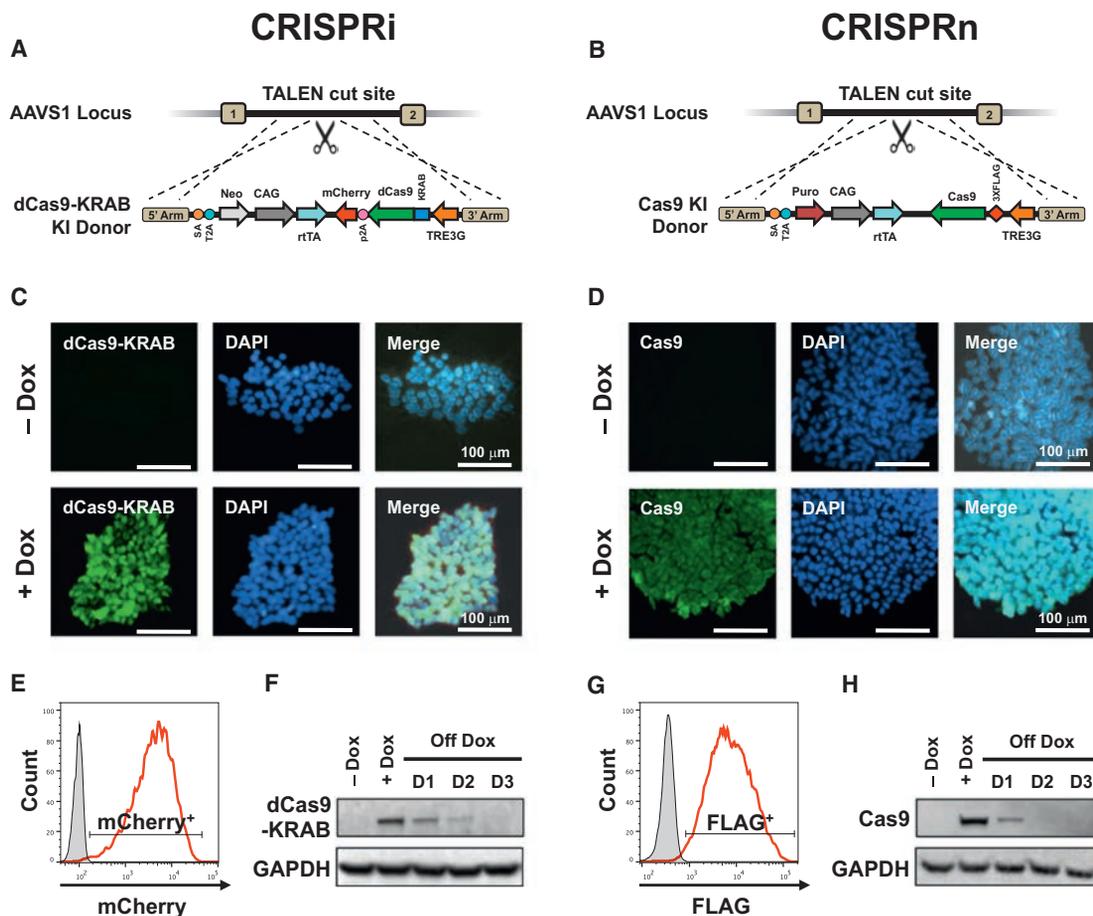


Figure 1. Generation and Characterization of Inducible CRISPRi and CRISPRn iPSCs

(A and B) Schematic overview of the strategy for TALEN-mediated targeting to the AAVS1 locus to generate the CRISPRi and CRISPRn iPSC lines. The doxycycline-controlled reverse transcriptional activator (rtTA) is driven by a strong constitutive promoter (CAG). The third-generation doxycycline-response element (TRE3G) drives transcription of either Cas9 (CRISPRn) or dCas9-KRAB-P2A-mCherry (CRISPRi) and is oriented in the opposite direction of the transactivator to ensure no leaky expression without doxycycline treatment.

(C and D) Immunostaining of CRISPRi and CRISPRn colonies before and after 48 hr of doxycycline treatment with an antibody against Cas9 (green). Nuclei are stained with DAPI (blue). All nuclei showed expression of dCas9-KRAB or Cas9 after adding doxycycline.

(E and G) Flow cytometry analysis of CRISPRi and CRISPRn iPSC lines before and after 48 hr of doxycycline treatment. Doxycycline treatment of CRISPRi and CRISPRn produced expression of mCherry and FLAG in all cells, respectively. The doxycycline-untreated sample is plotted in gray.

(F and H) CRISPRi and CRISPRn iPSC lines were treated with doxycycline (2 μ M) for 24 hr, which was then removed to measure the protein half-life of dCas9-KRAB and Cas9. Total protein was extracted from samples and analyzed by western blot with antibodies against Cas9 and GAPDH as a loading control. Both the CRISPRi and CRISPRn clones express dCas9-KRAB and Cas9 at similar levels after doxycycline treatment, and the half-life of both proteins was \sim 12 hr in iPSCs. Scale bars, 100 μ m.

indicated by a clear loss of pluripotent cell morphology (Hayashi et al., 2015). In general, Cas9 can disrupt gene function at any given exon (Doench et al., 2014), while dCas9-KRAB knocks down gene expression only when gRNAs are targeted to the transcription start site (TSS) (Gilbert et al., 2014). Hence, for this comparative study, we used the same gRNA sequence for both CRISPRi and CRISPRn. Here, we introduced a gRNA targeting 358 bp downstream of the *NANOG* TSS (142 bp into exon 1 of *NANOG*) into the CRISPRi and CRISPRn clones and selected subclones (as described in Experimental Procedures). We then treated multiple independent subclones of CRISPRi and CRISPRn iPSCs containing the *NANOG* gRNA-expression vector (as indicated by mKate2 expression) with doxycycline (Figure 2).

With CRISPRi, we found that *NANOG* expression was completely lost (>99%) in multiple independent iPSC subclones after doxycycline treatment (Figures 2A, 2C, 2E, S3A, and S3C). However, with CRISPRn, only 60%–70% of the cells lost *NANOG* expression in multiple independent subclones post-doxycycline induction (Figures 2B, 2D, 2G, S3B, and S3D). Next, we extracted genomic DNA from *NANOG* gRNA-containing CRISPRi and CRISPRn iPSCs and performed sequence analysis. As expected, we found that CRISPRi iPSCs did not harbor any mutations in the *NANOG* locus pre- or post-doxycycline treatment (Figure 2F). However, with CRISPRn, after 12–17 days of continuous doxycycline treatment, among the mutated alleles, 30%–50% of the sequences contained in-frame INDELs at the cut site (a total of 77 sequenced clones) (Figure 2H).

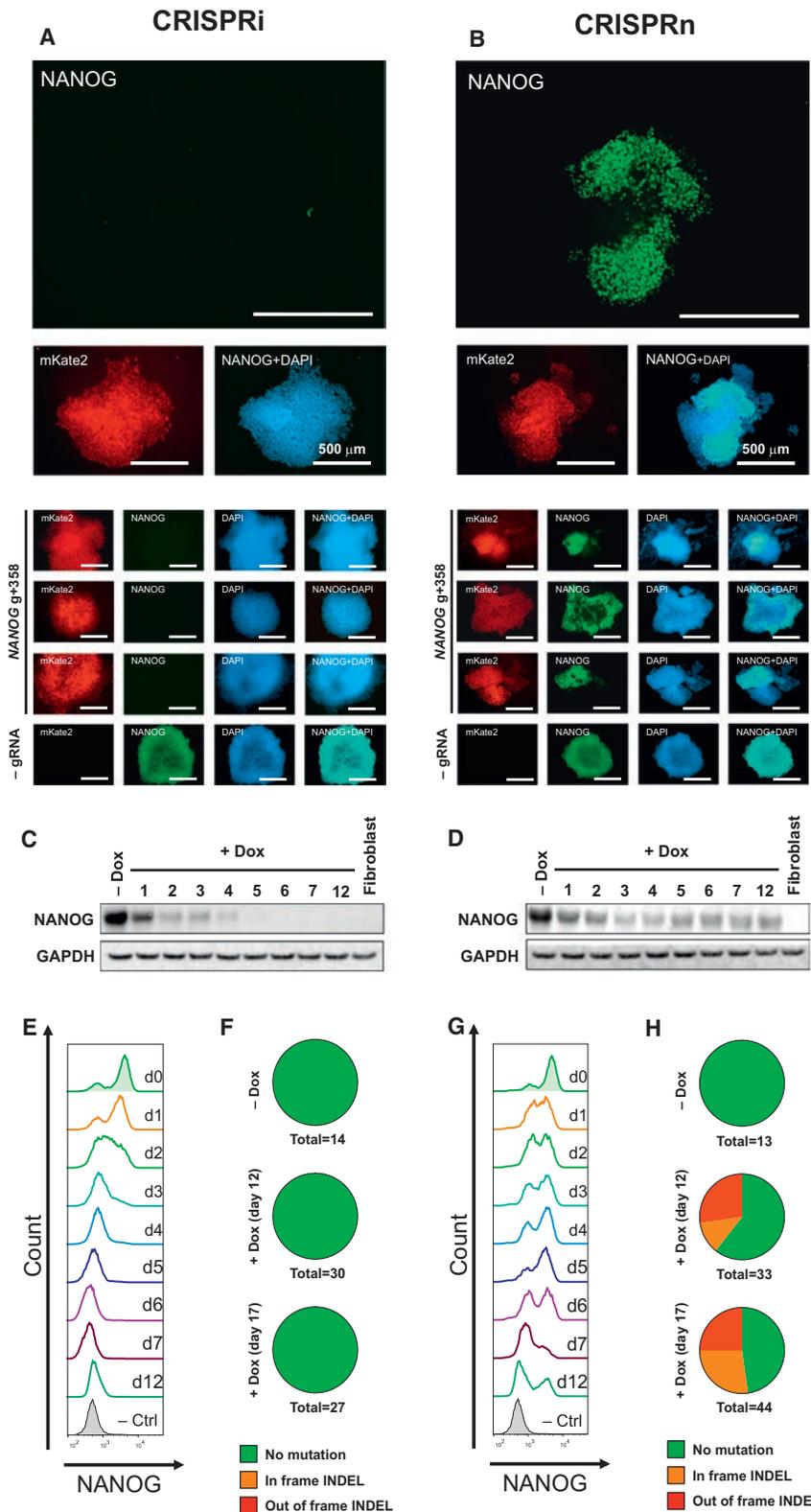


Figure 2. Comparison of the Efficiency of CRISPRi Knockdown and CRISPRn Knockout

(A and B) Immunostaining of representative (A) CRISPRi and (B) CRISPRn stable clones, each containing the same gRNA targeting the first exon of *NANOG* (*NANOG* g+358). After 7 days of doxycycline treatment, *NANOG* expression (green) was completely lost in all CRISPRi clones but showed a variegated pattern of knockout in multiple independent CRISPRn clones. The mKate2 signal indicates the presence of the gRNA-expression vector in all cells within the clone. Nuclei are counterstained with DAPI.

(C, D, E, and G) Western blot and flow cytometry analyses of (C and E) CRISPRi and (D and G) CRISPRn stable clones that contain the same gRNA against the first exon of *NANOG*. With CRISPRi, *NANOG* expression was uniformly decreased during doxycycline treatment and did not increase thereafter; however, with CRISPRn, the percentage of *NANOG*-positive cells fluctuated during doxycycline treatment. Even after 12 days of continuous doxycycline treatment, ~30% of the population stained positive for *NANOG*.

(F and H) Genomic DNA was extracted from (F) CRISPRi and (H) CRISPRn stable lines containing a gRNA against *NANOG* before and after continuous doxycycline treatment for up to 17 days and subjected to sequencing. Red, out-of-frame INDELS; orange, in-frame INDELS; green, non-mutated alleles. Even after 12–17 days of continuous doxycycline treatment, 50%–70% of sequenced alleles from CRISPRn contained no mutation, and 30%–50% of mutated alleles were in-frame INDELS. No mutations were observed in either CRISPRi or CRISPRn without doxycycline, and the CRISPRi clones did not contain any mutations after doxycycline treatment. The total number of sequenced colonies is listed below each pie graph.

Scale bars, 500 μ m.

was completely knocked down in independent CRISPRi clones expressing the gRNA vector after doxycycline treatment (Figure S3E). In contrast, the attempted knockout of *OCT4* with CRISPRn again yielded incomplete effects (Figure S3F). These findings were also replicated in a completely different iPSC line (WTB genetic background; CRISPRi Gen1B and CRISPRn Gen1B) (Figures S1D and S1F). We analyzed the genomic DNA of CRISPRn cells after 14 days of continuous doxycycline treatment and found 30%–40% of the mutated alleles had in-frame INDELS (a total of 91 sequenced clones) (Figure S3G). These results suggested that, in the context of targeting pluripotency factors, CRISPRi more rapidly generates loss-of-function phenotypes in bulk populations than CRISPRn. CRISPRi caused a complete

To further compare CRISPRi with CRISPRn, we targeted another pluripotency transcription factor, *OCT4*, with two independent gRNAs. Similar to our findings with *NANOG*, *OCT4*

was completely knocked down in independent CRISPRi clones expressing the gRNA vector after doxycycline treatment (Figure S3E). In contrast, the attempted knockout of *OCT4* with CRISPRn again yielded incomplete effects (Figure S3F). These findings were also replicated in a completely different iPSC line (WTB genetic background; CRISPRi Gen1B and CRISPRn Gen1B) (Figures S1D and S1F). We analyzed the genomic DNA of CRISPRn cells after 14 days of continuous doxycycline treatment and found 30%–40% of the mutated alleles had in-frame INDELS (a total of 91 sequenced clones) (Figure S3G). These results suggested that, in the context of targeting pluripotency factors, CRISPRi more rapidly generates loss-of-function phenotypes in bulk populations than CRISPRn. CRISPRi caused a complete

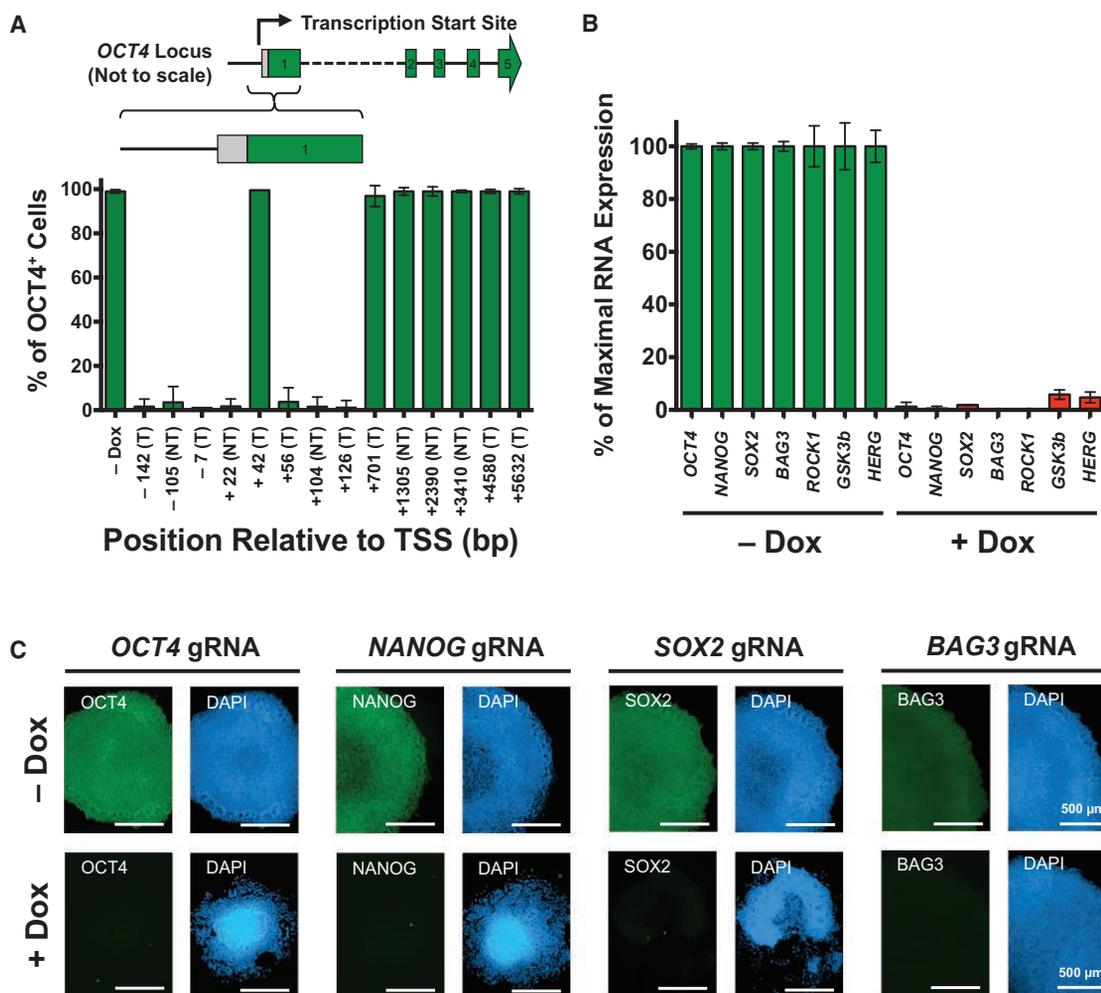


Figure 3. CRISPRi Knockdown Is Efficient in iPSCs

(A) Efficiency of gRNA knockdown based on proximity to the transcription start site (TSS). The binding location of each gRNA is indicated relative to the TSS of the *OCT4* locus and whether it targets the template (T) or non-template (NT) strand. Only gRNAs targeting near the TSS (approximately ± 150 bp) effectively knocked down *OCT4*.

(B) TaqMan qPCR analysis of stable iPSCs containing gRNA against the gene of interest showed greater than 90% knockdown efficiency after 7 days of doxycycline induction in different endogenous genetic loci.

(C) Immunostaining of stable clones containing a single gRNA against the gene of interest (*OCT4*, *SOX2*, *NANOG*, and *BAG3*). After 7 days of doxycycline treatment, there was a complete knockdown of the protein of interest (green). As expected, DAPI staining revealed that knocking down *OCT4*, *NANOG* and *SOX2* resulted in loss of pluripotency and clear morphological changes. Also, knocking down *BAG3* did not cause a loss of pluripotent morphology, as indicated by the distinct and round colony edges.

Error bar represents SD.

loss of transcript expression and rapid cell differentiation when targeting *NANOG* and *OCT4* within 5–7 days of knockdown initiation. With CRISPRn, even after ~ 2 weeks of doxycycline treatment, a significant fraction (30%–40%) of the cells remained *NANOG* and *OCT4* positive and maintained their pluripotency. Therefore, we focused on using CRISPRi as a loss-of-function tool in subsequent experiments.

CRISPRi Is Most Effective near the TSS

To further test the efficacy of gRNAs in CRISPRi, we designed multiple gRNAs that target near the TSS of *OCT4*. With flow cytometry assays for *OCT4* staining (Figure 3A), we found that most gRNAs targeting near the TSS (approximately -150 bp

to $+150$ bp around the TSS in this study) were highly effective at gene knockdown, but gRNAs targeting significantly (>700 bp) downstream of the TSS were not. This result agrees with previous data (Gilbert et al., 2014) and suggests that CRISPRi primarily blocks transcription at initiation, which reduces the likelihood of off-target effects from transcript interference elsewhere in the genome. Following these design criteria, for subsequent gene targets, we designed gRNAs to target near the TSS.

CRISPRi Efficiently Knocks Down a Broad Range of Genetic Loci

To test the efficiency of CRISPRi across a broad range of genetic loci in both iPSCs and differentiating/differentiated cell types, we

designed gRNAs against a total of nine genomic loci. The loci included core pluripotency transcription factors (*OCT4*, *NANOG*, and *SOX2*), kinases (*ROCK1* and *GSK3-β*), a cardiac mesoderm-transcription factor (*MESP1*), and cardiac disease-associated genes (*BAG3*, *MYBPC3*, and *HERG*). Except for *MESP1* (expressed only transiently in cardiac mesoderm cells) and *MYBPC3* (expressed only in cardiomyocytes), all other genes are expressed in iPSCs at different levels. We generated populations of CRISPRi iPSCs containing stably integrated gRNA-expression constructs. We then cultured these stable polyclones or clonal populations either with or without doxycycline for at least 7 days.

Three to five gRNAs were designed to target near the TSS of each gene and initially were tested individually in polyclonal populations. Approximately half of the tested gRNAs were active in polyclonal populations with a silencing activity of over 70% (Figure S4A). We did not observe a difference in the knockdown efficiency between gRNAs targeting either the template or non-template strands (Figures 3A, S4A, and S4B). The most active gRNA-containing polyclonal line was further passaged and subcloned for more detailed knockdown analysis. Using the most active gRNA, we achieved 90%–99% knockdown of the gene of interest in a selected population of iPSCs after doxycycline treatment (Figure 3B). As expected, when we subcloned polyclonal populations via single-cell cloning, we observed a higher percentage of transcriptional knockdown. With immunofluorescence analysis we found that iPSC clones expressing gRNAs against *OCT4*, *NANOG*, *SOX2*, and *BAG3* showed complete loss of target protein expression 7 days after doxycycline induction. In cells expressing gRNAs against the core pluripotency transcription factors *OCT4*, *NANOG*, and *SOX2*, we observed clear morphological changes and a loss of pluripotency after doxycycline induction; however, loss of a non-pluripotency gene (*BAG3*) did not affect pluripotent morphology (Figure 3C).

Using the Gen1 CRISPRi knockin vector, we targeted non-iPSCs with a different genetic background to determine how broadly this technology can be applied to other cell types. A T-lymphocyte (CEM) CRISPRi line was generated, as described in Experimental Procedures. Similar to the iPSC experiments, gRNAs were introduced to the stable CEM CRISPRi cell line, and cells cultured in either the presence or absence of doxycycline for 10 days. Three gRNAs were tested to knock down *CD4* in CEM-CRISPRi cells, and all showed greater than 70% knockdown efficiency in polyclonal populations (Figure S4B). The most active gRNA-containing polyclone was subcloned, and three independent clonal lines were isolated and assayed for knockdown, where greater than 95% knockdown efficiency was observed (Figure S4C). These results clearly demonstrate the doxycycline-inducible CRISPRi vector system is highly versatile and transportable to other cell lines and shows high efficiency of knockdown across a range of cell types and genetic loci.

CRISPRi Knockdown Is Reversible and Tunable and Can Be Allele Specific

GCaMP is a calcium-sensitive modified GFP and, thus, can be used as a fluorescent reporter under steady-state levels of cytoplasmic Ca^{2+} (Apátı et al., 2013). Using GCaMP (driven off the strong constitutive promoter, CAG), we monitored the green-fluorescence signal in iPSCs to determine if we could knock down GCaMP and then reverse its expression by removing

doxycycline from the culture. We found that adding doxycycline for 7 days knocked down GCaMP expression by 98%, which was completely restored after removing doxycycline for 14 days (Figure 4A). Similarly, we targeted the *BAG3* endogenous locus and achieved efficient transcript knockdown post-doxycycline treatment. *BAG3* expression was fully restored after doxycycline withdrawal (Figure 4B). These findings indicate that CRISPRi knockdown is fully reversible in iPSCs.

To determine if we could achieve variable levels of knockdown with different gRNA sequences, we tested two additional gRNAs targeting GCaMP (g+24 and g+91) (Figure 4C). These gRNAs knocked down GCaMP expression by only ~30% and ~50%, as measured by flow cytometry (Figures 4D and 4E). Therefore, by changing the location of the gRNA-binding site, we can tune the level of knockdown when trying to mimic haploinsufficiency or reduced protein levels (rather than complete loss of function). In addition, we tested whether the knockdown level is tunable by titrating the doxycycline concentration. Careful titration of the doxycycline concentration enabled homogenous modulation of GCaMP expression (Figure S5).

We next sought to further test the tunability of knockdown with CRISPRi. We determined if we could use single-nucleotide polymorphisms (SNPs) to specifically target one allele for knockdown to achieve a heterozygous-like state. In our CRISPRi iPSCs, there is a SNP near the TSS of *OCT4*. Thus, we designed a gRNA in which the heterozygous SNP is located in the PAM sequence (AGG versus AGA). Because an “NGG” sequence is required for dCas9 to target DNA, we could selectively target only one *OCT4* allele (Figure 4F). After doxycycline induction, we found that the iPSC population carrying the SNP-specific *OCT4* gRNA (*OCT4* g-4) remained *OCT4* positive (~97%) by flow cytometry analysis. However, the median intensity of *OCT4* staining was reduced by ~40% after 7 days of doxycycline treatment, implying that *OCT4* expression was homogeneously reduced in all cells and not just a fraction of them (Figures 4G and 4H). We confirmed this finding with immunocytochemistry and TaqMan qPCR (data not shown).

CRISPRi Knockdown Is Highly Specific

To assess the specificity of CRISPRi targeting, we designed a gRNA that targets the GCaMP transgene, since its silencing should have few downstream transcriptional and cellular consequences. Indeed, expression of the GCaMP transcript was over 30-fold lower in the doxycycline-treated condition, while few other endogenous transcripts changed expression level with the exception of *VIM* as previously discussed (Figure 5A).

CRISPRi to Promote iPSC Differentiation

To show that our system can release iPSCs from their pluripotent state to promote differentiation, we tested the efficiency of CRISPRi in knocking down core pluripotency transcription factors (*OCT4*, *SOX2*, and *NANOG*) without adding small molecules or cytokines to the mTeSR media. We targeted gRNA against these genes and performed a time-course analysis of a selected number of transcripts by TaqMan qPCR (Figure 5B). We found that knocking down these target transcripts caused cell differentiation, as indicated by morphological changes and transient expression of the lineage-specific transcript *T* (mesoderm marker), and expression of *PAX6* (neuronal progenitor marker). After 3 days

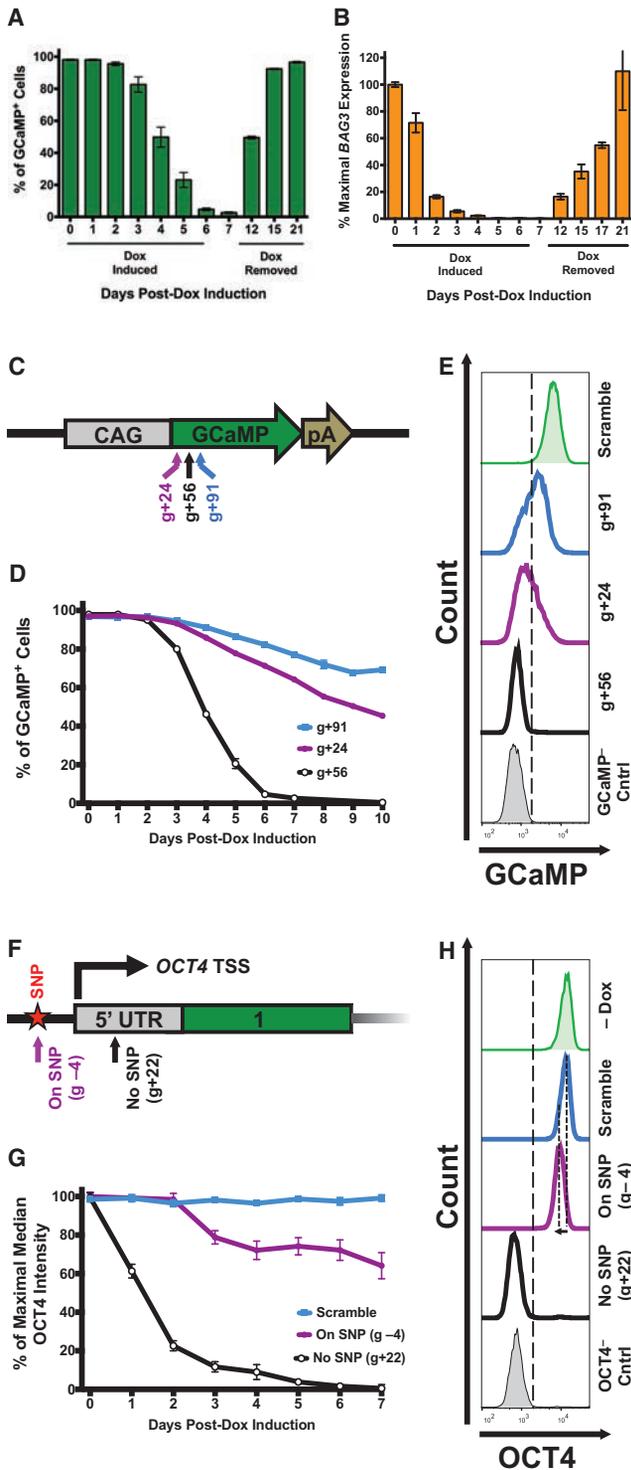


Figure 4. CRISPRi Knockdown Is Reversible and Tunable

A CRISPRi clone containing gRNA against the GCaMP transgene (GCaMP g+56) and endogenous *BAG3* locus were used to test the knockdown efficiency and reversibility of the CRISPRi system in iPSCs.

(A) Flow cytometry analysis of GCaMP expression showed that after 7 days of doxycycline induction, GCaMP was knocked down by ~99% and was completely restored after doxycycline withdrawal for 14 days.

(B) Using TaqMan qPCR, *BAG3* transcript levels were knocked down to nearly undetectable levels, and expression was restored after doxycycline withdrawal.

of doxycycline treatment, over 80% of the target transcript was depleted, indicating that CRISPRi can precisely and temporally control efficient knockdown of the transcript of interest.

CRISPRi Knockdown in Cardiac Mesoderm and iPSCs-CMs

To determine if loss-of-function approaches using CRISPRi can be applied in differentiated cell types, we targeted the cardiac mesoderm-specific transcription factor (*MESP1*) and two known cardiac-related disease-causing genes (*MYBPC3* and *HERG*). We established stable polyclonal lines of iPSCs containing gRNA against these three genes and differentiated them into cardiac mesoderm or iPSCs-CMs as described in Experimental Procedures (Figures S6A and S6B). Using a gRNA against these genes, *MESP1* was knocked down by ~90% in cardiac progenitor cells, and *MYBPC3* and *HERG* by ~90% and 60%, respectively, in lactate-purified iPSCs-CMs (Figure 6A). With western blots and immunocytochemistry, we observed ~90% MYBPC3 protein knockdown on day-35 lactate-purified iPSCs-CMs (Figures 6B and 6C).

Using flow cytometry, we analyzed the doxycycline response of CRISPRi cells based on mCherry expression (as a surrogate

(C) Schematic diagram of the GCaMP-expression vector in which the GCaMP open reading frame (ORF) is driven off the CAG promoter. The locations of three gRNAs (g+24, g+56, and g+91) are schematically highlighted on the GCaMP ORF. The coordinates of GCaMP gRNA are based on the translation start site. pA, poly A signal.

(D) Three stable CRISPRi colonies, each containing a different gRNA against GCaMP, were selected using blasticidin and cultured with doxycycline for 10 days. The percentage of GCaMP-positive cells for each gRNA-containing clone was plotted as a function of time based on flow cytometry analysis. Variable levels of GCaMP knockdown (~30%, ~50%, and ~99%) were achieved with different gRNA sequences. n = 1–3 technical replicates for each time point.

(E) Flow cytometry plots of GCaMP fluorescence of stable CRISPRi clones on day 10 of doxycycline treatment. Using different gRNAs that target near the same region, variable levels of knockdown can be achieved. A scramble gRNA-containing CRISPRi and a GCaMP-negative iPSC population are displayed as controls.

(F) Partial schematic diagram of the *OCT4* locus marked with the location of the TSS and two gRNA-binding locations. Asterisk, an SNP; green box, 5' UTR.

(G) Three stable CRISPRi colonies, two with different gRNAs against *OCT4* and one with a scrambled control, were selected with blasticidin. Stable clones that contain either a scramble gRNA, a gRNA that targets a PAM sequence containing a SNP (*OCT4* g-4), or a gRNA that does not target a SNP (*OCT4* g+22) were treated with doxycycline. The percentage of the maximal median intensity of *OCT4* staining for each gRNA-containing clone is plotted as a function of time by flow cytometry analysis. Complete loss of *OCT4* expression (>98% knockdown) was observed after 7 days of doxycycline induction only when both alleles were targeted using *OCT4* g+22. While using *OCT4* g-4, which targets only one *OCT4* allele (due to SNP in the PAM sequence), a gradual loss of *OCT4* staining intensity is observed over time (down by ~40% by day 7). Error bars represent SD; n = 1–3 technical replicates for each time point.

(H) Flow cytometry plots of *OCT4* staining on day 7 of doxycycline treatment. Dashed lines highlight the loss of *OCT4*-staining intensity (~40%) when using *OCT4* g-4 compared to the scramble control. By targeting only one allele of *OCT4*, the *OCT4*-staining intensity homogeneously shifts (while remaining *OCT4*-positive), indicating that each cell experiences approximately the same level of knockdown. Note that the x axis is a log-scale of *OCT4* intensity. Differentiated iPSC-derived fibroblasts (*OCT4*⁻ Cntrl) and a non-doxycycline-treated (-Dox) sample are displayed as controls. Error bars represent SD.

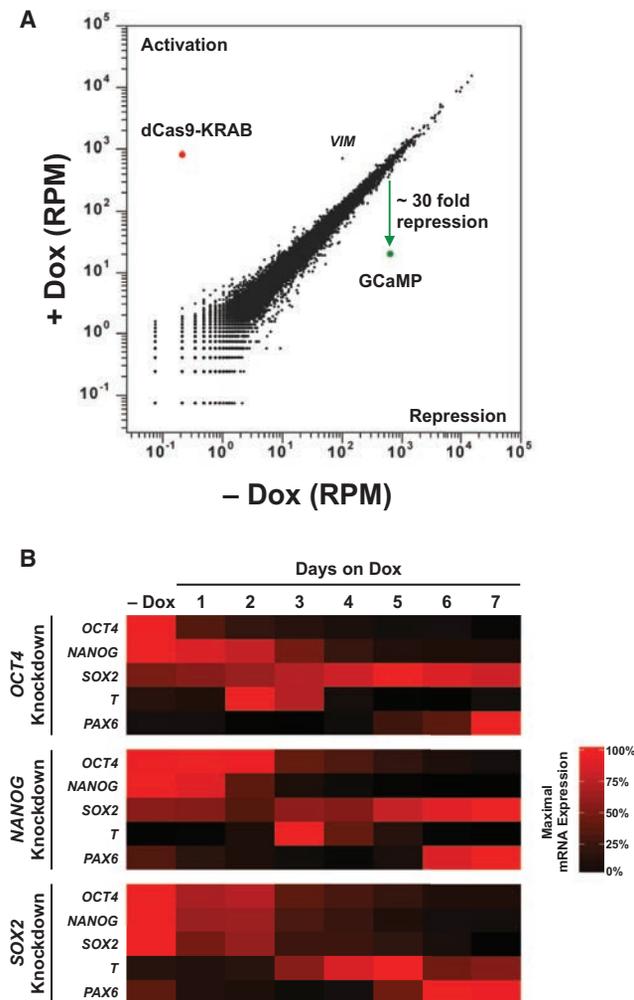


Figure 5. RNA-Seq and TaqMan qPCR Analysis

(A) RNA-sequencing RPMs (reads per million) are plotted for CRISPRi cells stably expressing a gRNA targeting the GCaMP transgene (GCaMP g+56) cultured in the absence or presence of doxycycline. CRISPRi knockdown is specific to the GCaMP transcript, and few off-target transcriptional changes were observed. Data represent two independent biological replicates.

(B) Heatmap of TaqMan qPCR of stable clones containing a single gRNA against the gene of interest (*OCT4*, *NANOG*, and *SOX2*) as a function of days after doxycycline treatment. Analysis shows that by day 3, over 80% of the target transcript is depleted. Three housekeeping genes (*18S*, *GAPDH*, and *UBC*) were used to measure relative transcript levels. Each data point is an average of two to four technical replicates. TaqMan probes are listed in Supplemental Experimental Procedures.

for dCas9-KRAB expression; Figure S5A). There was no silencing of the TetO promoter in low-passage and high-passage iPSCs, suggesting that long-term culturing (>3 months) does not cause silencing. However, cardiac progenitors (day 5) and iPSC-CMs (day 15) lose ~20% and 50%–80% of the doxycycline response, respectively. Prolonging the duration of doxycycline treatment (from 2 to 7 days) and splitting the cells improved doxycycline response (as measured by mCherry expression) in iPSC-CMs (Figure S6C). For this reason, we initiated all of our knockdowns on day 5 post-differentiation to obtain the

maximum amount of target gene silencing. It is worth noting that with CRISPRi, only minute amounts of the dCas9-KRAB protein are necessary to induce a knockdown. Hence, knockdown might occur even in cells that do not show detectable mCherry expression (Figure S5).

The knockdown of the *HERG* potassium channel in iPSCs was highly efficient (>95%), while in iPSC-CMs it was only 60% effective. We hypothesize that the reduction in the efficiency of *HERG* knockdown is partially due to activation of other *HERG* isoforms in iPSC-CMs. We further investigated whether knocking down the *HERG* potassium channel in iPSC-CMs would recapitulate a physiologically relevant cellular phenotype. We found that knocking down *HERG* in iPSC-CMs lead to a prolonged beat duration and the appearance of a shoulder during the downstroke, as measured using the GCaMP signal (which can be used as a surrogate for the action potential) (Huebsch et al., 2015) (Figures 6D and 6E). We confirmed the prolongation of action potential duration by patch-clamp electrophysiology in the *HERG* knockdown samples (Figures 6F). We expected this result, because the *HERG* potassium channel pumps potassium ions out of cells to lower the inner membrane potential during diastole. This cellular phenotype recapitulates aspects of the phenotype observed in LQT patients and their iPSC-CMs (Schwartz et al., 2012; Spencer et al., 2014).

DISCUSSION

In this study, we combined the power of human iPSC technology, which generates functional human cells, with inducible CRISPR-based genome editing and modulation technologies. Using the TetO inducible system, we deploy the newly developed CRISPRi system in the AAVS1 safe-harbor locus of human iPSCs to enable precise control of transcript silencing upon addition of doxycycline. With this approach, we rapidly and efficiently generated loss-of-function phenotypes in iPSCs and their cell-type derivatives to study mechanisms in development and disease. We introduced a single doxycycline-inducible vector system into the AAVS1 safe-harbor locus to gain tight transcriptional control of dCas9-KRAB (for CRISPRi) and Cas9 (for CRISPRn) for gene knockdown and knockout studies, respectively. This inducible vector system helped us precisely control the timing of knocking down the expression of target genes in a clonal iPSC line carrying the gRNA of interest. We were also able to efficiently target the CRISPRi vector into non-iPSC human cells (T-lymphocytes) and show efficient levels of transgene knockdown, which demonstrates the versatility of using the CRISPRi system in a wide range of cell types. This system can be readily targeted to other human cellular models in vitro and also to mouse models (Soriano, 1999) by exchanging the AAVS1-homology arms with the ROSA26-specific knockin arms.

We found that in iPSC populations, CRISPRi produced a homogeneous and rapid loss-of-function phenotype compared to CRISPRn. CRISPRi avoids potential complications associated with incomplete loss-of-function and gain-of-function phenotypes in cell populations produced by Cas9-induced hypomorphic alleles. Therefore, CRISPRi represents a powerful technology for repressing gene expression in bulk populations and especially when performing genome-scale phenotypic screens. Every CRISPRi iPSC that contained a target-specific gRNA

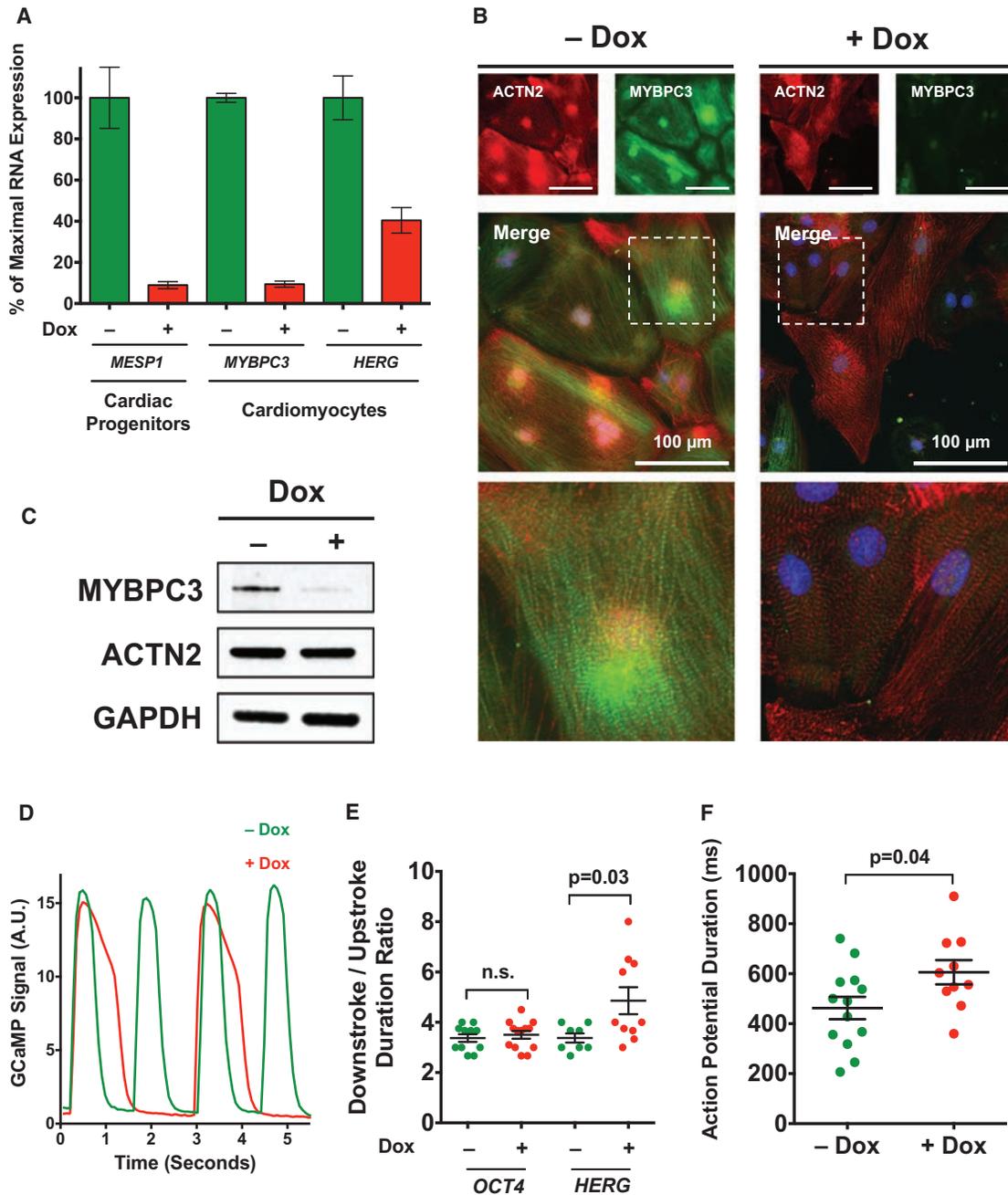


Figure 6. CRISPRi Knockdown in Differentiated Cell Types and Cardiac Disease Modeling

(A) Using CRISPRi, *MESP1* was knocked down by ~90% in polyclonal cardiac progenitors, and *MYBPC3* and *HERG* were knocked down by ~90% and 60% in polyclonal iPS-CMs, respectively.

(B) Immunostaining of day-35 lactate-purified iPS-CMs stained with antibodies against *MYBPC3* (green) and *ACTN2* (red). Using CRISPRi knockdown, loss of *MYBPC3* was observed in over 85% of analyzed cells in a polyclonal population. Nuclei were counterstained with DAPI. Scale bar, 100 μ m.

(C) Western blot of day-35 lactate-purified iPS-CMs with antibodies against *MYBPC3*, *ACTN2*, and *GAPDH*. Using CRISPRi, *MYBPC3* protein was knocked down by ~90%.

(D) GCaMP fluorescence in iPS-CMs containing gRNA against *HERG* and cultured in doxycycline (red). Recordings show a prolonged beat duration compared to untreated controls (green).

(E) Quantified ratio of the downstroke-to-upstroke duration of doxycycline-treated iPS-CMs shows a significant difference in untreated iPS-CMs containing a gRNA against *HERG*, but not in iPS-CMs containing gRNA against *OCT4* (negative control).

(F) Patch-clamp recordings from single iPS-CMs show prolonged action potential durations in doxycycline-treated samples containing *HERG* gRNA. Error bars represent SD.

displayed a rapid, uniform, and efficient transcriptional knockdown. This result was also validated across multiple endogenous loci in iPSCs, cardiac progenitors, and iPSC-CMs. By contrast, using CRISPRn, we found that while all cells harbored the gRNA-expression vector and had continuous expression of Cas9, they did not all display complete loss-of-function phenotypes. Indeed, up to one-third of the cells maintained expression of the target gene. When we sequenced the target alleles, we found that of the mutated alleles, over one-third had in-frame INDELS, potentially resulting in a hypomorphic protein encoded by a gene that is now resistant to further Cas9 cutting using the target gRNA. Statistically, we expect that one-third of the INDELS generated by double-strand breaks induced by Cas9 through the non-homologous end-joining pathway would produce in-frame mutations. This effect could cause partial loss-of-function or gain-of-function phenotypes. Additionally, the location and size of the in-frame INDEL might not change the function of the mutated protein compared with the wild-type protein (Boettcher and McManus, 2015; Shi et al., 2015; Sung et al., 2013).

CRISPRi gRNAs were only effective at promoter regions close to the TSS, which may reduce the likelihood of off-target effects by transcriptional interference elsewhere in the genome. Indeed, RNA-seq analysis showed that the knockdown of GCaMP was highly specific. Furthermore, expression of dCas9-KRAB did not cause significant off-target transcriptional changes as compared to Cas9 expression alone. Although CRISPRi is highly effective, there are cases when other genetic tools such as CRISPRn, TALENs, and RNAi may have advantages. For instance, we and others (Gilbert et al., 2014) have shown that CRISPRi gRNAs are only effective near the TSS, which restricts the efficiency of transcript for genes that have poorly defined or multiple TSSs. CRISPRn and TALENs can be effective at any exon as long as the genomic region is accessible (Doench et al., 2014; Kim et al., 2013b). Additionally, RNAi can target any constitutive portion of the mRNA and has already been approved for human therapy (Davidson and McCray, 2011; Haussecker, 2012); however, RNAi has been shown to have many off-target effects (Jackson et al., 2003; Kim et al., 2013b; Krueger et al., 2007).

We also demonstrated the feasibility of allele-specific interference and the tunable nature of CRISPRi-based knockdown, which can be used to study the dose-dependent effects of a gene involved in development and disease. The dosage of transcription factors plays a significant role during development and organogenesis (McFadden et al., 2005; Takeuchi et al., 2011). In addition, many human diseases result from haploinsufficiency in which a mutation in a single copy of a gene produces the disease phenotype (Armanios et al., 2005; Marston et al., 2012; Minami et al., 2014; Theodoris et al., 2015). Therefore, to study the dose-dependent effects of transcription factors in development and disease, CRISPRi can be used to homogeneously tune the level of repression in cells by either choosing the relevant gRNA sequences or empirically titrating the levels of doxycycline to achieve the desired knockdown level. Alternatively, introducing a single point mutation at different positions in the gRNA sequence (which leads to mismatches between the RNA-DNA homology sequence) can be used to tune CRISPRi knockdown activity (Gilbert et al., 2014). Finally, CRISPRi knock-

down was reversible in iPSCs upon doxycycline withdrawal, which would support studies involving transient knockdown of transcripts within a specific window during cell differentiation.

Our studies with CRISPRi in iPSCs showed that knocking down transcripts involved in maintaining pluripotency is highly efficient and rapidly causes a complete loss of pluripotent morphology, followed by cell differentiation in all cells expressing the appropriate gRNA. We also used this approach to knock down the *HERG* potassium channel to mimic an LQT2-type phenotype in iPSC-CMs. We found that the inducible TetO promoter is partially silenced during the cardiac differentiation process, which has been reported to be due to methylation at CpG dinucleotides (Oyer et al., 2009). This silencing is independent of integration at the AAVS1 locus, as CAG-driven transgenes integrated at the AAVS1 locus remain active after differentiation. To avoid the effects of promoter silencing, we initiated transcript knockdown in the iPSC state or progenitor cells (day 5 of differentiation), where the vast majority of the cells respond to doxycycline. This strategy has proved highly effective at transgene knockdown in cardiac progenitors and iPSC-CMs. To circumvent issues with silencing in future studies, we generated a non-inducible CRISPRi iPSC line (Gen3; in which dCas9-KRAB is driven off the CAG promoter), and the knockdown can be initiated upon introduction of gRNA. With this cell line, we expect to achieve highly efficient knockdown in differentiated cell types, such as iPSC-CMs.

Several groups have used the CRISPR/Cas9 system for loss-of-function genetic screens in human cells (Shalem et al., 2014; Wang et al., 2014). Furthermore, some groups have used genome-scale screens with CRISPRi and CRISPR activation (CRISPRa) to identify known and novel genes that control cell growth and sensitivity to cholera-diphtheria toxin (Gilbert et al., 2014). In this study, we present our CRISPRi iPSC lines as suitable model systems for performing screens to identify novel transcripts of pluripotency, drug resistance, and cell survival at the pluripotent stem cell stage. With genome-scale screens, we can identify factors that improve cell-specific differentiation into functional cell types that have been traditionally hard to obtain, and we can more rapidly generate mature functional cell types that better mimic *in vivo* cell counterparts. In addition, with CRISPRi, we can repress putative disease-associated genes in a medium- to high-throughput manner to unravel the molecular mechanisms underlying human disease *in vitro*. Finally, we can build on the current power of CRISPRi for developmental screens by using an orthogonal dCas9-effector system for gene activation via CRISPRa, which can synergistically modulate gene knockdown and activation and direct cell fate toward a particular lineage.

EXPERIMENTAL PROCEDURES

iPSC Culture

WTB and WTC iPSCs and derivative lines were maintained under feeder-free conditions on growth factor-reduced Matrigel (BD Biosciences) and fed daily with mTeSR medium (STEMCELL Technologies) (Ludwig et al., 2006). Accutase (STEMCELL Technologies) was used to enzymatically dissociate iPSCs into single cells. To promote cell survival during enzymatic passaging, cells were passaged with the p160-Rho-associated coiled-coil kinase (ROCK) inhibitor Y-27632 (10 μ M; Selleckchem) (Watanabe et al., 2007). iPSCs were frozen in 90% fetal bovine serum (HyClone) and 10%

DMSO (Sigma). The committee on Human Research at the University of California, San Francisco approved the iPSC research protocol (#10-02521).

Generation of Stable CRISPRi and CRISPRn iPSC Lines

iPSCs were singularized with accutase, resuspended in PBS, and counted with a Countess automated cell counter (Life Technologies). For plasmid transfections, the human stem cell nucleofector kit 1 solution was used on the Amaxa nucleofector 2b device (program A-23; Lonza). To generate the CRISPRi and CRISPRn iPSC lines, two million WTC or WTB iPSCs were nucleofected with the appropriate knockin vector (5 μ g) and each AAVS1 TALEN pair (2 μ g). Cells were then seeded in six-well plates in serial dilutions in mTeSR supplemented with Y-27632 (10 μ M). Selection was applied 3 days post-nucleofection with the appropriate antibiotic in mTeSR plus Y-27632 (10 μ M). To knock in the CRISPRi construct (carrying the Neomycin resistance gene cassette), Geneticin (Life Technologies) was applied at 100 μ g/ml. To knock in the CRISPRn and GCaMP constructs (carrying the Puromycin resistance gene cassette), 0.5 μ g/ml Puromycin (Life Technologies) was added. Selection was maintained for \sim 10 days until stable colonies appeared. Colonies with a diameter of greater than \sim 500 μ m were manually picked using a P200 pipette tip under an EVOS FL picking microscope (Life Technologies) and transferred to individual wells of a 24-well plate containing mTeSR medium supplemented with Y-27632 (10 μ M). Clones were then expanded into larger vessel formats.

Generation of CEM CRISPRi Cell Line

CEM CRISPRi cells were generated by electroporation of 0.5 μ g of each AAVS1 TALEN pair and 1 μ g of the Gen1 CRISPRi vector with an Amaxa nucleofector 2b device and Amaxa cell line nucleofector kit C (Lonza). Cells were selected in 1 μ g/ μ l G418, and clonal lines were generated by dilution in 96-well plates. Clonal populations were selected based on doxycycline induction of mCherry expression. Oligos encoding the CD4 protospacer were annealed and cloned into the pSLQ1371 vector using restriction sites BstXI and BlnI, and lentivirus was produced in HEK293T cells (Gilbert et al., 2014). To compare performance of CD4 gRNAs, each was transduced into CEM-CRISPRi cells. Transduced populations were incubated for 96 hr with doxycycline (2 μ M). Knockdown efficiency was calculated by gating all mCherry-expressing cells, and comparing cell-surface CD4 expression in the presence or absence of gRNA-expressing cells (BFP⁺). Three independent stable CEM CRISPRi clones were selected with 0.6 μ g/ml Puromycin and incubated in the presence or absence of doxycycline (2 μ M) for 14 days to assess maximal CD4 knockdown. Cells were stained using anti-CD4 APC-conjugated antibody and cell surface CD4 staining was quantified using a BD LSRII flow cytometer. CD4 knockdown was quantified as percent reduction relative to no doxycycline treatment condition.

gRNA Design and Cloning into the gRNA-Expression Vector

For CRISPRi, three to five gRNAs were designed to target near the TSS of the gene of interest (250 bp upstream and downstream, respectively). The location of the TSS was determined using NCBI (<http://www.ncbi.nlm.nih.gov/>). gRNA oligos were designed, phosphorylated, annealed, and cloned into the pgRNA-CKB vector using BsmBI ligation strategy. Additional details and a list of gRNA sequences are listed in supplemental experimental procedures.

gRNA Nucleofection and Selection of Stable CRISPRi and CRISPRn Clones

The gRNA-expression vector (pgRNA-CKB) was transfected into either the CRISPRi or CRISPRn cells with the human stem cell nucleofector kit 1 solution on the Amaxa nucleofector 2b device (program A-23; Lonza). Two million CRISPRi or CRISPRn iPSCs and 5 μ g of the circular gRNA-expression plasmid were used per nucleofection. Nucleofected cells were then seeded in a single well of a six-well plate in mTeSR supplemented with Y-27632 (10 μ M). Blastocidin selection (10 μ g/ml) was applied 24 hr post-nucleofection in mTeSR supplemented with Y-27632 (10 μ M) for 7–10 days, until stable colonies appeared. Stable colonies were then pooled and passaged at least three times in mTeSR plus Blastocidin and Y-27632 to enrich for cells with integration at transcriptionally active sites (Figure S3).

RNA Sequencing

For each sample, 1 μ g of total RNA was prepared using TRIzol as previously described. Strand-specific mRNA-seq libraries were prepared using TruSeq Stranded mRNA Library Prep Kit (Illumina). Upon completion, libraries were quantified and pooled using Qubit dsDNA HS assay and Agilent's Bioanalyzer high-sensitivity DNA assay. The indexed libraries were pooled and sequenced on Illumina HiSeq 4000 as 50-bp single-end reads. Reads were aligned to the hg19 genome assembly using the Ensembl 75 reference transcriptome customized to include the GCaMP6f constructs using TopHat2 (Kim et al., 2013a). Unaligned reads were subsequently aligned to the CRISPRi or CRISPRn knockin constructs where appropriate. Transcript alignments were then counted using SubRead v1.4.6 and analyzed with custom scripts written in Python (Liao et al., 2013). All data are displayed as reads per million (RPM) with a pseudocount of 0.075.

iPS-CM Differentiation and Lactate Purification

iPSCs were differentiated into iPS-CMs using the WNT modulation-differentiation method (Lian et al., 2012) (Figure S5A). iPS-CMs were purified via a modified version of the lactate metabolic-selection method (Tohyama et al., 2013). Additional details are outlined in Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the RNA-seq data reported in this paper is GEO: PRJNA307261.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one movie and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.01.022>.

AUTHOR CONTRIBUTIONS

M.A.M. and B.R.C. were primarily responsible for conception, design, and interpretation of the experiments. M.A.M. conducted most experiments with help from N.H., E.F., E.S., A.T., M.P.O., T.V.E., K.H., and L.M.J. Y.M. and A.H.C. generated the CRISPRn Gen1C iPSC line. C.I.S. performed electrophysiology experiments. D.E.G. generated the CEM CRISPRi cell line and provided knockdown analysis. L.A.G., J.S.W., and L.S.Q. provided technical expertise, the CRISPRi fusion cassette, and gRNA expression constructs. J.E.V. and M.A.H. conducted and analyzed the RNA-seq experiments. M.A.M., P.L.S., and B.R.C. wrote the manuscript with support from all authors.

ACKNOWLEDGMENTS

We thank members of the Conklin laboratory, Gladstone Institute of Cardiovascular Disease, Roddenberry Stem Cell Core, BioFulcrum, a Gladstone Institutes Enterprise, and Innovative Genomics Institute for technical assistance and helpful comments on the manuscript. We thank Tim Rand and Knut Woltjen for valuable discussions and helpful comments on the manuscript. We thank S. John Liu for RNA-seq analysis advice. We thank Jen Berman and Samantha Cooper at Bio-Rad for assistance with designing ddPCR probe-primer sets. Summer students Matthew Keller and Monique Morrison assisted with preliminary experiments. CEM CD4⁺ cells were obtained from Dr. J.P. Jacobs through the AIDS Reagent Program, Division of AIDS, NIAID, NIH. M.A.M. is supported by the Canadian Institutes of Health Research postdoctoral fellowship 129844. N.H. was supported by the CIRM training program TG2-01160 and T32 HL007544. E.F. was supported by a Bridges to Stem Cell Training grant TB1-01188 from CIRM. L.M.J. is supported by the CIRM Training Grant TG2-01160 and NICHD Career Development Award 1K12HD072222. Y.M. received fellowships from the Uehara Memorial Foundation Research and Gladstone-CIRM. D.G. is supported by UCSF-Gladstone Center for AIDS Research (CFAR), an NIH-funded program (P30 AI027763). T.V.E. was supported by Carlsberg Travel Grant (2013-01-0423), The Lundbeck Foundation (R140-2013-13348), and OUH Internationalisation Foundation. J.E.V., M.A.H., L.A.G., and J.S.W. were supported by the Howard Hughes

Medical Institute and the National Institutes of Health (R01 DA036858). B.R.C. received support from the US National Heart, Lung, and Blood Institute, the National Institutes of Health (U01-HL100406, U01-GM09614, R01-HL108677, U01-HL098179, U01-HL099997, P01-HL089707, R01HL130533, and R01-HL060664), and an Agilent University Relations Grant.

Received: August 20, 2015

Revised: December 21, 2015

Accepted: January 24, 2016

Published: March 10, 2016

REFERENCES

- Apáti, Á., Pászty, K., Hegedűs, L., Kolacsek, O., Orbán, T.I., Erdei, Z., Szébenyi, K., Péntek, A., Enyedi, Á., and Sarkadi, B. (2013). Characterization of calcium signals in human embryonic stem cells and in their differentiated offspring by a stably integrated calcium indicator protein. *Cell. Signal.* **25**, 752–759.
- Armanios, M., Chen, J.-L., Chang, Y.-P.C., Brodsky, R.A., Hawkins, A., Griffin, C.A., Eshleman, J.R., Cohen, A.R., Chakravarti, A., Hamosh, A., and Greider, C.W. (2005). Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc. Natl. Acad. Sci. USA* **102**, 15960–15964.
- Boettcher, M., and McManus, M.T. (2015). Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. *Mol. Cell* **58**, 575–585.
- Bolouri, H., and Davidson, E.H. (2003). Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. *Proc. Natl. Acad. Sci. USA* **100**, 9371–9376.
- Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823.
- Davidson, B.L., and McCray, P.B., Jr. (2011). Current prospects for RNA interference-based therapies. *Nat. Rev. Genet.* **12**, 329–340.
- Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32**, 1262–1267.
- Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096.
- Gaj, T., Gersbach, C.A., and Barbas, C.F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31**, 397–405.
- Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442–451.
- Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661.
- González, F., Zhu, Z., Shi, Z.-D., Lelli, K., Verma, N., Li, Q.V., and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* **15**, 215–226.
- Haussecker, D. (2012). The business of RNAi therapeutics in 2012. *Mol. Ther. Nucleic Acids* **1**, e8.
- Hayashi, Y., Caboni, L., Das, D., Yumoto, F., Clayton, T., Deller, M.C., Nguyen, P., Farr, C.L., Chiu, H.-J., Miller, M.D., et al. (2015). Structure-based discovery of NANOG variant with enhanced properties to promote self-renewal and reprogramming of pluripotent stem cells. *Proc. Natl. Acad. Sci. USA* **112**, 4666–4671.
- Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* **29**, 731–734.
- Huebsch, N., Loskill, P., Mandegar, M.A., Marks, N.C., Sheehan, A.S., Ma, Z., Mathur, A., Nguyen, T.N., Yoo, J.C., Judge, L.M., et al. (2015). Automated video-based analysis of contractility and calcium flux in human-induced pluripotent stem cell-derived cardiomyocytes cultured over different spatial scales. *Tissue Eng. Part C Methods* **21**, 467–479.
- Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P.S. (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21**, 635–637.
- Kearns, N.A., Genga, R.M.J., Enuameh, M.S., Garber, M., Wolfe, S.A., and Maehr, R. (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development* **141**, 219–223.
- Kim, H., and Kim, J.-S. (2014). A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* **15**, 321–334.
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013a). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36.
- Kim, Y., Kweon, J., Kim, A., Chon, J.K., Yoo, J.Y., Kim, H.J., Kim, S., Lee, C., Jeong, E., Chung, E., et al. (2013b). A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.* **31**, 251–258.
- Krueger, U., Bergauer, T., Kaufmann, B., Wolter, I., Pilk, S., Heider-Fabian, M., Kirch, S., Artz-Oppitz, C., Isselhorst, M., and Konrad, J. (2007). Insights into effective RNAi gained from large-scale siRNA validation screening. *Oligonucleotides* **17**, 237–250.
- Lian, X., Hsiao, C., Wilson, G., Zhu, K., Hazeltine, L.B., Azarin, S.M., Raval, K.K., Zhang, J., Kamp, T.J., and Palecek, S.P. (2012). Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. *Proc. Natl. Acad. Sci. USA* **109**, E1848–E1857.
- Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108.
- Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D.F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* **8**, 861–869.
- Ludwig, T.E., Bergendahl, V., Levenstein, M.E., Yu, J., Probasco, M.D., and Thomson, J.A. (2006). Feeder-independent culture of human embryonic stem cells. *Nat. Methods* **3**, 637–646.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826.
- Marston, S., Copeland, O., Gehmlich, K., Schlossarek, S., and Carrier, L. (2012). How do MYBPC3 mutations cause hypertrophic cardiomyopathy? *J. Muscle Res. Cell Motil.* **33**, 75–80.
- McFadden, D.G., Barbosa, A.C., Richardson, J.A., Schneider, M.D., Srivastava, D., and Olson, E.N. (2005). The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. *Development* **132**, 189–201.
- Minami, S.S., Min, S.-W., Krabbe, G., Wang, C., Zhou, Y., Asgarov, R., Li, Y., Martens, L.H., Elia, L.P., Ward, M.E., et al. (2014). Progranulin protects against amyloid β deposition and toxicity in Alzheimer's disease mouse models. *Nat. Med.* **20**, 1157–1164.
- Miyaoaka, Y., Chan, A.H., Judge, L.M., Yoo, J., Huang, M., Nguyen, T.D., Lizarraga, P.P., So, P.-L., and Conklin, B.R. (2014). Isolation of single-base genome-edited human iPS cells without antibiotic selection. *Nat. Methods* **11**, 291–293.
- Oyer, J.A., Chu, A., Brar, S., and Turker, M.S. (2009). Aberrant epigenetic silencing is triggered by a transient reduction in gene expression. *PLoS ONE* **4**, e4832.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183.

- Schwartz, P.J., Crotti, L., and Insolia, R. (2012). Long-QT syndrome: from genetics to management. *Circ Arrhythm Electrophysiol* 5, 868–877.
- Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343, 84–87.
- Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., and Vakoc, C.R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* 33, 661–667.
- Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* 21, 70–71.
- Spencer, C.I., Baba, S., Nakamura, K., Hua, E.A., Sears, M.A.F., Fu, C.C., Zhang, J., Balijepalli, S., Tomoda, K., Hayashi, Y., et al. (2014). Calcium transients closely reflect prolonged action potentials in iPSC models of inherited cardiac arrhythmia. *Stem Cell Reports* 3, 269–281.
- Sternecker, J.L., Reinhardt, P., and Schöler, H.R. (2014). Investigating human disease using stem cell models. *Nat. Rev. Genet.* 15, 625–639.
- Sung, Y.H., Baek, I.-J., Kim, D.H., Jeon, J., Lee, J., Lee, K., Jeong, D., Kim, J.-S., and Lee, H.-W. (2013). Knockout mice created by TALEN-mediated gene targeting. *Nat. Biotechnol.* 31, 23–24.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872.
- Takeuchi, J.K., Lou, X., Alexander, J.M., Sugizaki, H., Delgado-Olguín, P., Holloway, A.K., Mori, A.D., Wylie, J.N., Munson, C., Zhu, Y., et al. (2011). Chromatin remodelling complex dosage modulates transcription factor function in heart development. *Nat. Commun.* 2, 187.
- Theodoris, C.V., Li, M., White, M.P., Liu, L., He, D., Pollard, K.S., Bruneau, B.G., and Srivastava, D. (2015). Human disease modeling reveals integrated transcriptional and epigenetic mechanisms of NOTCH1 haploinsufficiency. *Cell* 160, 1072–1086.
- Tohyama, S., Hattori, F., Sano, M., Hishiki, T., Nagahata, Y., Matsuura, T., Hashimoto, H., Suzuki, T., Yamashita, H., Satoh, Y., et al. (2013). Distinct metabolic flow enables large-scale purification of mouse and human pluripotent stem cell-derived cardiomyocytes. *Cell Stem Cell* 12, 127–137.
- Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR/Cas9 system. *Science* 343, 80–84.
- Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. *Science* 350, 1096–1101.
- Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S., Muguruma, K., and Sasai, Y. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat. Biotechnol.* 25, 681–686.
- Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* 482, 331–338.

De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data

Dominic Grün,^{1,2,3,*} Mauro J. Muraro,^{1,2} Jean-Charles Boisset,^{1,2} Kay Wiebrands,^{1,2} Anna Lyubimova,^{1,2} Gitanjali Dharmadhikari,^{1,2,4} Maaïke van den Born,^{1,2} Johan van Es,^{1,2} Erik Jansen,^{1,2} Hans Clevers,^{1,2,5} Eelco J.P. de Koning,^{1,2,4} and Alexander van Oudenaarden^{1,2,*}

¹Hubrecht Institute, Royal Netherlands Academy of Arts and Sciences, 3584 CT Utrecht, the Netherlands

²Cancer Genomics Netherlands, University Medical Center Utrecht, 3584 CX Utrecht, the Netherlands

³Max Planck Institute of Immunobiology and Epigenetics, 79108 Freiburg, Germany

⁴Department of Medicine, Section of Nephrology and Section of Endocrinology, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands

⁵Princess Maxima Center for Pediatric Oncology, 3508 AB Utrecht, the Netherlands

*Correspondence: gruen@ie-freiburg.mpg.de (D.G.), a.vanoudenaarden@hubrecht.eu (A.v.O.)

<http://dx.doi.org/10.1016/j.stem.2016.05.010>

SUMMARY

Adult mitotic tissues like the intestine, skin, and blood undergo constant turnover throughout the life of an organism. Knowing the identity of the stem cell is crucial to understanding tissue homeostasis and its aberrations upon disease. Here we present a computational method for the derivation of a lineage tree from single-cell transcriptome data. By exploiting the tree topology and the transcriptome composition, we establish StemID, an algorithm for identifying stem cells among all detectable cell types within a population. We demonstrate that StemID recovers two known adult stem cell populations, *Lgr5*+ cells in the small intestine and hematopoietic stem cells in the bone marrow. We apply StemID to predict candidate multipotent cell populations in the human pancreas, a tissue with largely uncharacterized turnover dynamics. We hope that StemID will accelerate the search for novel stem cells by providing concrete markers for biological follow-up and validation.

INTRODUCTION

The identification of a stem cell in a tissue is a major challenge of pivotal importance. Being able to detect the stem cell population allows for powerful approaches to study cell differentiation dynamics by, for example, lineage tracing (Barker et al., 2007; Busch et al., 2015). Additionally, it provides a first step toward ex vivo propagation of primary stem cells in organoid cultures (Lancaster et al., 2013; Sato et al., 2009), important for applications in regenerative medicine. Moreover, stem cell populations relevant for disease progression, such as cancer stem cells, are promising targets for therapeutic intervention. Stem cells are typically rare, which makes their discovery by traditional population-based assays very difficult. For example, it took decades of dedicated research to define the population of hematopoietic stem cells (HSCs) (Eaves, 2015), but it remains an open question

how much heterogeneity exists within this subpopulation of bone marrow cells (Wilson et al., 2015). Similarly, the discovery of intestinal stem cells (van der Flier and Clevers, 2009) took years of work, and heterogeneity within this compartment remains under debate (Buczacki et al., 2013).

The recent availability of single-cell mRNA sequencing methods allows profiling of healthy and diseased tissues with single-cell resolution (Grün et al., 2015; Jaitin et al., 2014; Macosko et al., 2015; Patel et al., 2014; Paul et al., 2015; Treutlein et al., 2014; Zeisel et al., 2015). The transcriptome of a cell can be interpreted as a fingerprint, revealing its identity. However, biological gene expression noise (Eldar and Elowitz, 2010; Raj and van Oudenaarden, 2008) and technical noise because of amplification of minute amounts of mRNA from a single cell (Brennecke et al., 2013; Grün et al., 2014) affects the readout and makes it a challenge to discriminate cell types based on their transcriptome. By sequencing large numbers of randomly sampled single cells from a tissue, it is now possible to compile a nearly complete inventory of cell types.

These inventories can be screened for cell types of particular interest, such as stem cells. An obvious strategy for the identification of the stem cell is the derivation of a lineage tree from single-cell sequencing data. However, transcriptomes of randomly sampled cells only represent a snapshot of the system, and temporal differentiation dynamics cannot be directly derived. However, if the system of interest comprises all differentiation stages, such as the intestinal epithelium or the bone marrow, then attempts can be made to infer a lineage tree by assembling single-cell transcriptomes in a pseudo-temporal order. Existing approaches assume a continuous temporal change of transcript levels to assemble differentiation trajectories (Bendall et al., 2014; Haghverdi et al., 2015; Trapnell et al., 2014), but resolving the correct tree topology remains a challenge.

Here we present a method to identify rare and abundant cell types of a system and use these cell type classifications to guide the inference of a lineage tree. We investigate the general properties characterizing the position of a cell type within the lineage tree and identify the number of branches and the transcriptome uniformity of a cell type as features correlating with the degree of pluripotency. We show that our approach successfully recovers the identity of the stem cell in the intestine and in the bone

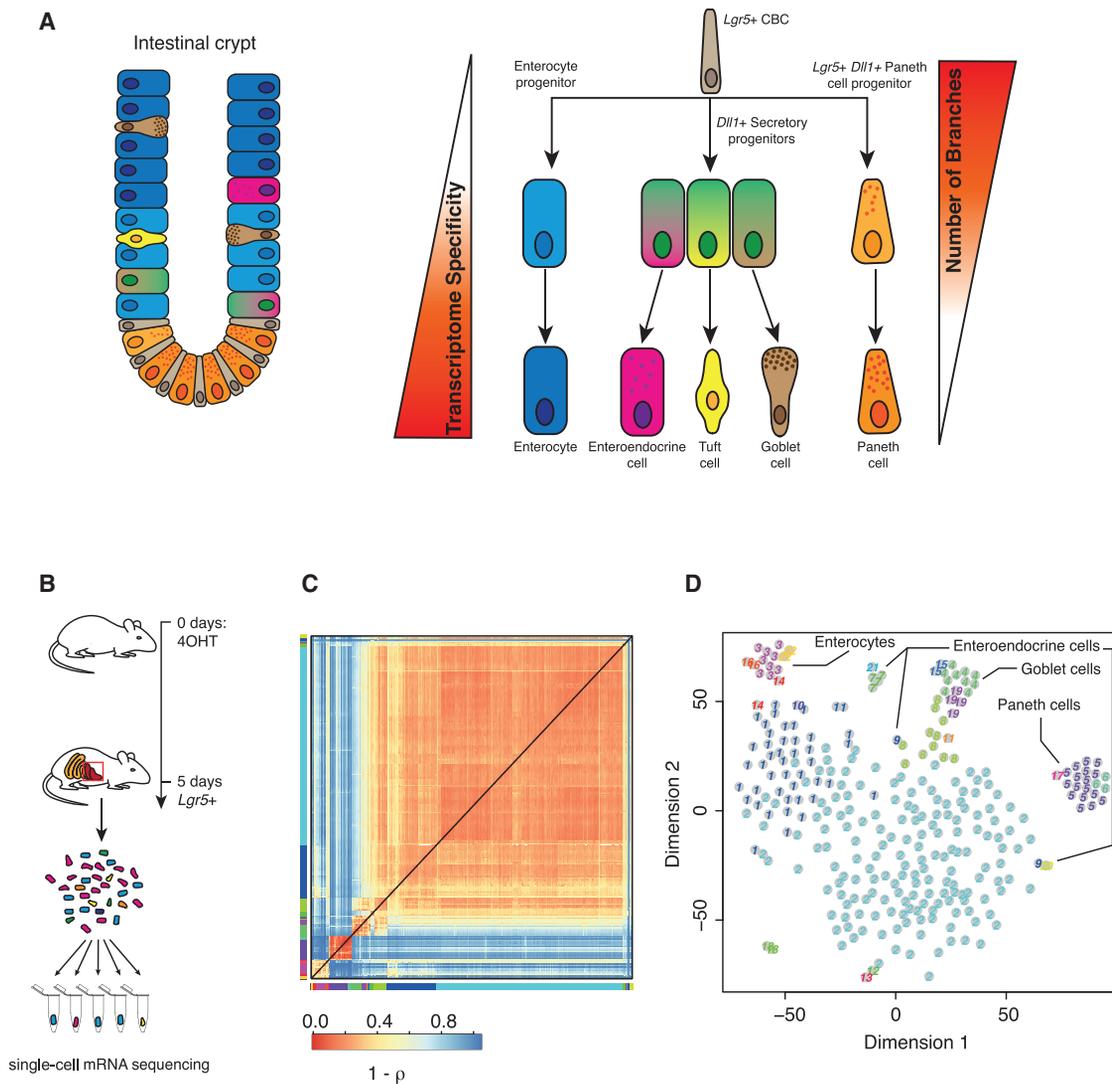


Figure 1. RacelD2 Recovers Intestinal Cell Types

(A) The intestinal epithelium is a well characterized differentiation system. *Lgr5*-positive stem cells give rise to secretory and absorptive precursors by WNT and NOTCH signaling that further differentiate into mature intestinal cell types.

(B) Summary of the lineage-tracing experiment performed to sequence single 5-day-old progeny of *Lgr5*-positive cells.

(C) Heatmap of cell-to-cell transcriptome distances measured by $1 - \rho$ (Pearson's correlation coefficient). RacelD2 clusters are color-coded along the boundaries.

(D) t-distributed stochastic neighbor embedding (t-SNE) map representation of transcriptome similarities between individual cells. The clusters identified in (C) are highlighted with different numbers and colors, and the corresponding intestinal cell types identified based on known marker genes are indicated.

See also Figure S1.

marrow, two systems with a well described stem cell population. We then use our method to predict multipotent cell populations in the adult human pancreas.

RESULTS

Robust Identification of Mouse Intestinal Cell Types by RacelD2

To develop a robust approach for the inference of differentiation trajectories, we used a previously published dataset from a lineage tracing experiment comprising the progeny of *Lgr5*-positive

mouse intestinal stem cells (Grün et al., 2015). This system is ideal for testing the inference of differentiation dynamics because the lineage tree is already well characterized (Figure 1A). The continuously self-renewing intestinal epithelium is arranged in crypts and villi, with a small number of *Lgr5*+ stem cells, also known as crypt base columnar cells (CBCs), residing near the crypt bottom. These CBCs give rise to rapidly proliferating transit-amplifying (TA) cells that migrate upward along the crypt-villus axis and develop into the terminally differentiated cell types (Barker, 2014; van der Flier and Clevers, 2009). Although absorptive enterocytes constitute the most abundant

cell type, the secretory lineage comprises rare cells, such as mucus-producing goblet cells, hormone-secreting enteroendocrine cells, and antimicrobial Paneth cells. Labeled cells were collected 5 days after label induction using an *Lgr5*-CreERT2 construct and a Rosa26-YFP reporter with a loxP-flanked transcriptional roadblock (Figure 1B).

We first improved the robustness of the initial clustering step of our previously developed RacelD algorithm (Grün et al., 2015) by replacing the k-means clustering with k-medoids clustering (Figure S1). Second, we noticed that the previously used gap statistic (Tibshirani et al., 2001) was not ideal for determining the cluster number. Although increasing the number of clusters in many cases leads to a growing gap statistic, the decrease of the within-cluster dispersion (Tibshirani et al., 2001) saturates quickly. A further increase of the cluster number, therefore, reduces cluster reproducibility. In RacelD2, we thus determine the cluster number by identifying the saturation point of the within-cluster dispersion. Together, these two changes lead to a more robust initial clustering of RacelD2 (Experimental Procedures; Figure S1).

For the intestinal lineage tracing data (Experimental Procedures), RacelD2 recovered a larger group of *Lgr5*+ stem cells (cluster 2) and early progeny (clusters 1 and 8) as well as the major mature cell types; i.e., enterocytes (cluster 3), goblet (clusters 4 and 19), Paneth (clusters 5 and 6), and enteroendocrine cells (cluster 7) (Figures 1C and 1D). These cell types could be unambiguously assigned based on the cluster-specific upregulation of marker genes inferred by RacelD2 (Table S1).

Inference of the Lineage Tree with Guided Topology

One of the major challenges for the inference of differentiation pathways in a system with multiple cell lineages is the determination of branching points. To overcome this problem, we predefined the topology of the lineage tree by allowing differentiation trajectories linking each pair of clusters. A putative differentiation trajectory links the medoids of two clusters, and the ensemble of all inter-cluster links defines the possible topology of the lineage tree. To minimize the effect of technical noise and, at the same time, the computational burden, we first reduce the dimensionality of the input space requiring maximal conservation of all point-to-point distances. In a second step, we assign each cell to its most likely position on a single inter-cluster link. To find this position, the vector connecting the medoid of a cluster to one of its cells is projected onto the links between the medoid of this and all remaining clusters, and the cell is assigned to the link with the longest projection after normalizing the length of each link to one. The projection also defines the most likely position of the cell on the link (Figure 2A), reflecting its differentiation state (Experimental Procedures). If this strategy is applied to the intestinal data, then only a subset of links is populated (Figure 2B). To determine links that are more highly populated than expected by chance and are therefore candidates for actual differentiation trajectories, we computed an enrichment *p* value based on comparison with a background distribution with randomized cell positions (Figure 2B; Figure S2A). Furthermore, we reasoned that the coverage of a link by cells indicates how likely it is that this link represents an actual differentiation trajectory and not only biased perturbations driving the transcriptome of a given cluster preferentially toward the transcriptome of another cluster without leading to actual differentiation events. We defined a link score as

one minus the maximum difference between the positions of each pair of neighboring cells on the link after normalizing the length of each link to one (Figure S2B). If this score is close to one, then the link is densely covered with cells with only small gaps in between. If the link score is close to zero, the cell density is only concentrated near the cluster centers connected by this link. A detailed description of the algorithm is given in the Experimental Procedures. The computationally inferred intestinal lineage tree is consistent with the known lineage tree (Figure 1A). Secretory cell types (clusters 4, 5, 6, and 7) populate individual branches emanating from the central *Lgr5*+ cluster, and absorptive enterocytes (cluster 3) differentiate from the same group via a more abundant group of TA cells (cluster 1).

We compared the inferred lineage tree to the tree predicted by Monocle (Trapnell et al., 2014), a recent method for the derivation of branched lineage trees that does not rely on a predefined tree topology, and found that Monocle could not resolve the different branches of secretory cells (Figure S2).

High Connectivity and High Transcriptome Entropy Reveals the Identity of the Stem Cell

Next we attempted to predict the stem cell identity from the lineage tree. Our working definition of a stem cell for this purpose purely relies on multipotency. More precisely, we try to identify, from the lineage tree, the cell population with the highest degree of multipotency. We noticed that different cell types showed a variable number of populated links to other clusters. The link score is reflected by the thickness of the line in our graphical representation (Figure 2B). We also show links with a low link score because they are informative about the associated cell state. For example, a cell type with many low-scoring links can fluctuate toward a diversity of fate biases, whereas cell types with only a few links are much more canalized. These two scenarios reflect a more promiscuous transcriptome, such as expected for stem cells, versus a more confined transcriptome, as expected for a mature cell type. In our data, cluster 2, which contains cells positive for *Lgr5* and other established stem cell markers (*Ascl2* and *Ctca4*) (Figure 2C), was the most highly connected cluster. Another putative property of stem cells is the tendency to exhibit a more uniform composition of the transcriptome in comparison with differentiated cells. Mature cell types frequently express a small number of genes at very high levels, crucial for cell type-specific functions. The transcriptome of Paneth cells, for instance, is dominated by high numbers of lysozymes and other host defense genes. The uniformity of the transcriptome is reflected by Shannon's entropy (Shannon, 1948), and this concept has previously been applied to study cellular differentiation (Anavy et al., 2014; Banerji et al., 2013; Piras et al., 2014) (Experimental Procedures). We anticipate that the transcriptome of a multipotent cell type is more uniform in each individual cell. In addition, multiple state biases could coexist within this population that can give rise to diverse mature cell types upon external stimuli, or stochastically, leading to high entropy (Banerji et al., 2013; Ridden et al., 2015). For the intestinal lineage tracing data, both Paneth and goblet cells had clearly reduced entropy compared with *Lgr5*-positive cells, whereas the entropy of enterocytes and enteroendocrine cells was comparable with stem cells (Figure 2D). We found that, for all analyzed datasets (see below), the number of links discriminates better between

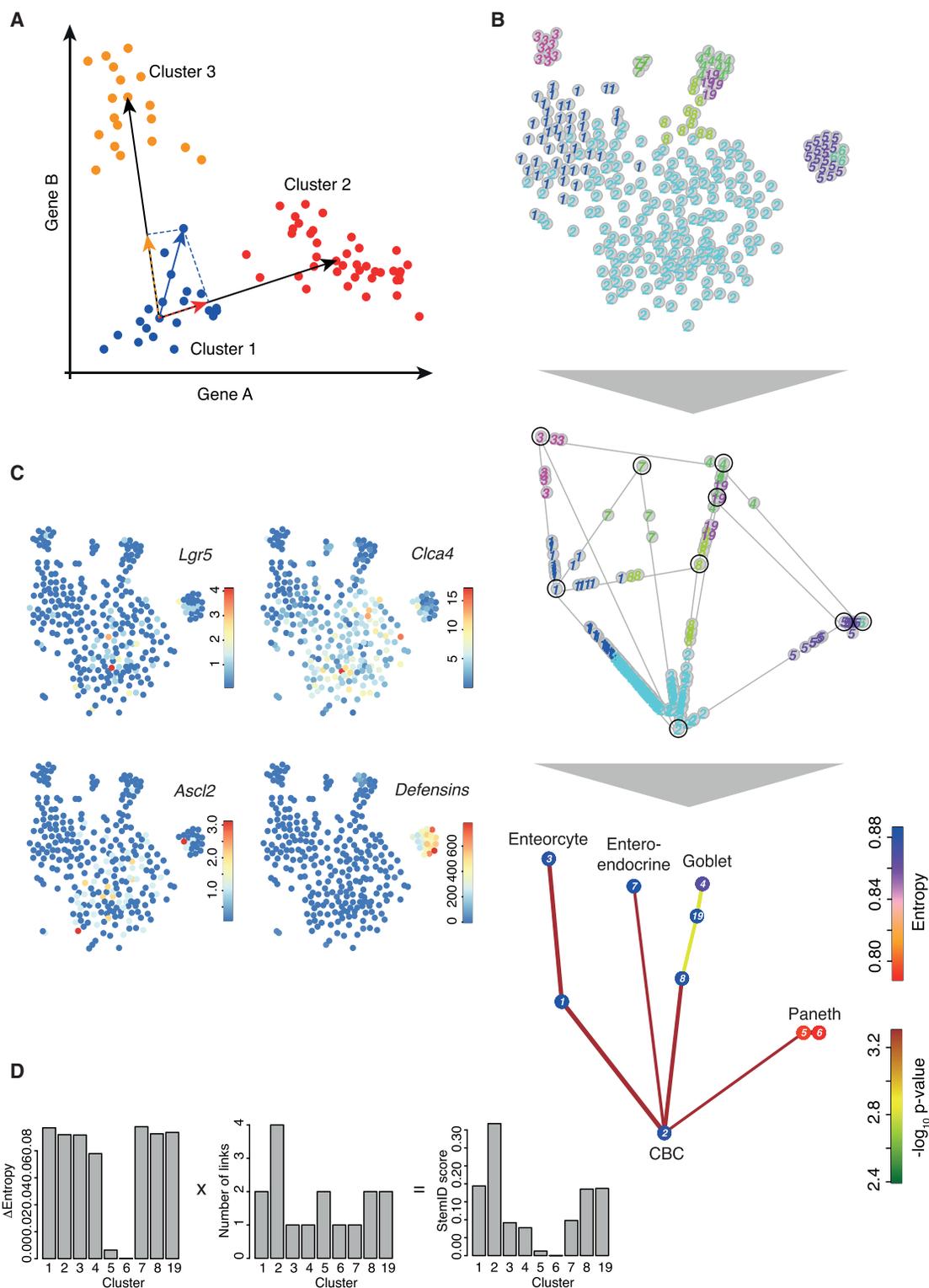


Figure 2. Lineage Tree Inference for Intestinal Stem Cell Progeny

(A) Schematic of the method used to infer differentiation trajectories (see main text and Experimental Procedures).

(B) Outline of the method visualized in the t-SNE-embedded space. All RacelD2 clusters with more than two cells (top) are connected by links, and, for each cell, the link with the maximum projection is determined as shown in (A). Only populated links are shown (center). Cluster centers are circled in black. Significant links are inferred by comparison with the background distribution with randomized cell positions (Experimental Procedures). Only significant links are

(legend continued on next page)

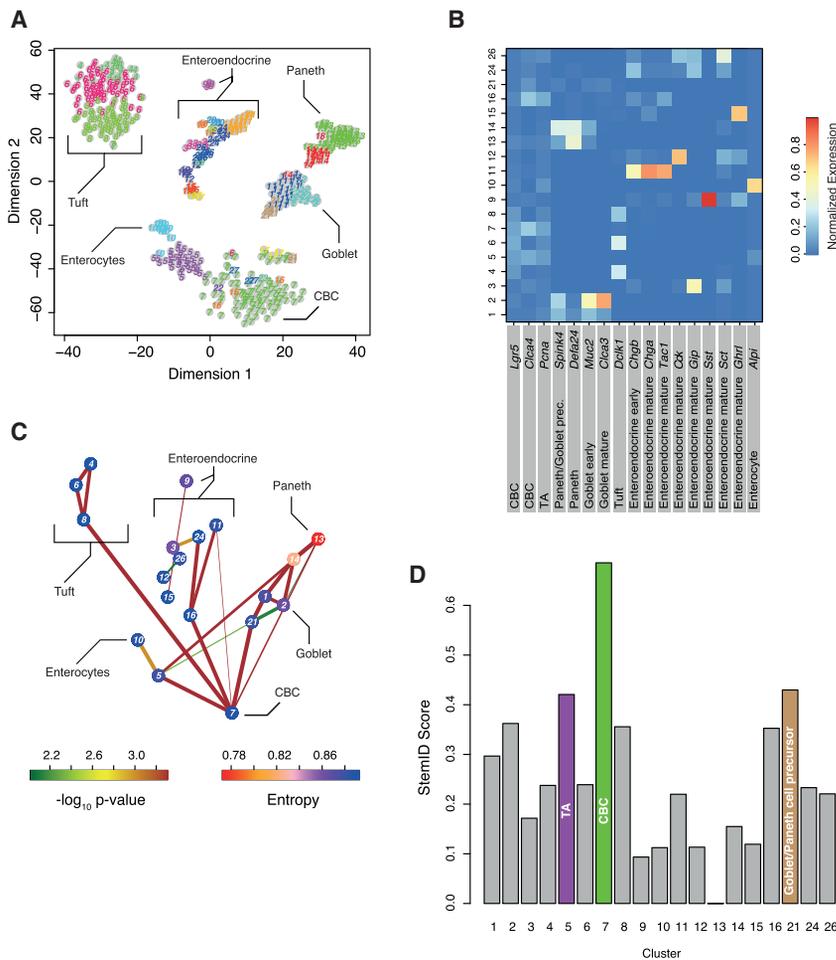


Figure 3. StemID Identifies Stem Cells in Complex with Non-random Mixtures of Intestinal Cells

(A) t-SNE map of transcriptome similarities of intestinal cells from a variety of single-cell mRNA sequencing experiments (main text and Figure S3). RacelD2 clusters are highlighted with different numbers and colors. Cell types identified based on marker gene expression are shown.

(B) Heatmap showing the average expression of known cell type markers across all clusters with more than five cells. For each gene, the sum of expression values over all clusters is normalized to one.

(C) Inferred intestinal lineage tree. Only significant links are shown ($p < 0.01$). The color of the link indicates the $-\log_{10}$ p value. The color of the vertices indicates the entropy. The thickness indicates the link score, reflecting how densely a link is covered with cells (Experimental Procedures).

(D) Barplot of StemID scores for intestinal clusters. In (B)–(D), only clusters with more than five cells were analyzed. See also Figures S3, S6, and S7.

multipotent and differentiated cells when rescaled by the entropy. Therefore, the simplest score that performs well in discriminating multipotent cells from the remaining cell types was a product of the median entropy (after subtracting the minimal entropy observed in the system) and the number of links of a cluster (Experimental Procedures). This score exhibits a clear maximum for cluster 2 comprising the *Lgr5*⁺ stem cells (Figure 2D). We named our algorithm StemID for the lineage tree inference and the derivation of this score.

StemID Recovers Intestinal Stem Cells in a Complex Dataset with Non-random Cell Type Frequencies

Next we wanted to test whether StemID could identify *Lgr5*⁺ cells in a larger and more complex dataset comprising intestinal cells of various independent experiments conducted in our lab. In this dataset, we combined 3 weeks and 8 weeks of *Lgr5* line-

age tracing data. A subset of those was enriched in secretory cells by fluorescence-activated cell sorting (FACS) on CD24 (van Es et al., 2012; Figure S3). For both time points, we also sorted non-traced CD24⁺ control cells (Experimental Procedures; Figure S3). RacelD2 revealed the known intestinal cell types within this dataset based on cluster-specific expression of known cell type marker genes and subdivided these into stages of differentiation or maturation (Figures 3A and 3B; Figure S3A). A full list of differentially expressed genes for each cluster is given in Table S2. For example, intestinal stem cells in cluster 7, marked by high expression of *Lgr5* and *Cla4* (Figure 3B), were connected directly to all secretory branches, whereas TA cells (cluster 5) primarily give rise to enterocytes (cluster 10) (Figure 3C; Figures S3C and S3D). Interestingly, we observed two distinct differentiation trajectories for Paneth cells (clusters 13 and 14), one via a *Dll1*-positive common precursor of Paneth and goblet cells (cluster 1) and another one directly connecting stem cells (cluster 7) or TA cells in cluster 5, marked by upregulation of the cell-cycle gene *Pcna*, directly to the mature Paneth cell clusters. Both the *Dll1*-dependent (van Es et al., 2012) and the direct route (Farin et al., 2014; Sawada et al., 1991), which was observed after ablation of Paneth cells, have been described. The recovery of alternative differentiation pathways demonstrates the power of our guided lineage inference. We were not able to recapitulate

shown ($p < 0.01$). The color of the link indicates the $-\log_{10}$ p value. The color of the vertices indicates the entropy. The thickness indicates the link score, reflecting how densely a link is covered with cells (Experimental Procedures).

(C) Transcript counts (color legend) of the intestinal stem cell markers *Lgr5*, *Cla4*, and *Ascl2* are highlighted in the t-SNE map. Expression of these genes is restricted to cluster 2 and clusters 5 and 6. Clusters 5 and 6 comprise Paneth cells, which were shown to co-express *Lgr5* (Grün et al., 2015). Accumulated transcript counts across all Defensin genes, which are markers of Paneth cells, are shown at the bottom right.

(D) Barplot of StemID scores for all clusters. The median transcriptome entropy of each cell type was computed across all cells in a cluster (left). The lowest entropy across all cell types was subtracted for each cell type because absolute differences were only small. This Δ entropy was multiplied by the number of significant links for each cluster (center), yielding the StemID score (right).

See also Figure S2.

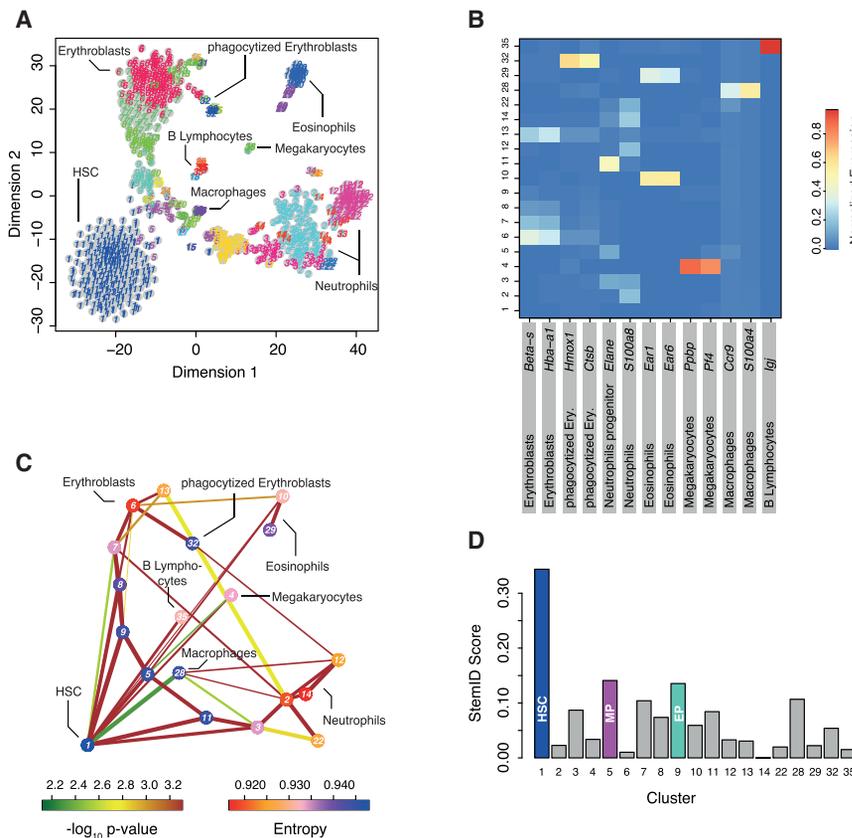


Figure 4. StemID Identifies Hematopoietic Stem Cells in Non-random Mixtures of Bone Marrow Cells

(A) t-SNE map of transcriptome similarities of hematopoietic cells sampled from physically interacting doublets or multiplets (main text and Figure S4). RaceID2 clusters are highlighted with different numbers and colors. Cell types identified based on marker gene expression are shown.

(B) Heatmap showing the average expression of known cell type markers across all clusters with more than five cells. For each gene, the sum of expression values over all clusters is normalized to one.

(C) Inferred hematopoietic lineage tree. Only significant links are shown ($p < 0.01$). The color of the link indicates the $-\log_{10}$ p value. The color of the vertices indicates the entropy. The thickness indicates the link score, reflecting how densely a link is covered with cells (Experimental Procedures).

(D) Barplot of StemID scores for hematopoietic clusters. MP, myeloid progenitor; EP, erythroblast progenitor.

See also Figures S4, S6, and S7.

this finding with a minimum spanning tree-based alternative approach (Figure S3E).

We then computed the StemID score and found that the *Lgr5+/Clca4+* cells (cluster 7) exhibit the highest score (Figure 3D). The second highest score was observed for cluster 21, which represents a common progenitor to Paneth and goblet cells. The TA cells in cluster 5, which our lineage inference identifies as progenitors with an enterocyte fate bias, acquire the third-highest StemID score.

Noticeably, Paneth cells in cluster 13 and mature goblet cells in cluster 2 show the same connectivity as the stem and progenitor cells in clusters 7, 5, and 21, but rescaling by entropy helps correctly assign a mature state to these cells (Figure S3F). In conclusion, StemID could identify intestinal stem cells and distinguish progenitor populations from more mature intestinal cell types.

StemID Recovers Hematopoietic Stem Cells within a Non-random Sample of Bone Marrow Cells

To test the performance of StemID in a different biological system, we applied the algorithm to single-cell sequencing data of mouse bone marrow cells. These cells were selected based on physical interactions between doublets or larger groups of cells and are thus not sampled randomly from all cell types in the bone marrow. This dataset was complemented with *Kit+Sca-1+Lin-CD48-CD150+* HSCs (Kiel et al., 2005) sorted from the bone marrow (Experimental Procedures; Figure S5B). Cell types identified by RaceID2 were dominated by the myeloid lineage and comprised HSCs, erythroblasts, megakaryocytes,

two groups of granulocytes (neutrophils and eosinophils), macrophages, a small group of B lymphocytes, and several clusters representing progenitor stages of the myeloid lineage (Figures 4A and 4B; Figure S6A). A full list of differentially expressed genes for each cluster is shown in Table S3. Cluster 1 comprises almost exclusively sorted HSCs (Figure S4B). The inferred lineage tree (Figure 4C; Figures S6C and S6D) indicates that HSCs differentiate into multipotent progenitor cells (cluster 5) but are also directly linked to mature lineages. HSCs and multipotent progenitors are both linked to megakaryocytes (cluster 4), eosinophils (clusters 10 and 29), macrophages (cluster 28), and two branches covering a spectrum of progenitor and mature states of the neutrophil (clusters 11, 3, 2, 14, 12, and 22) and erythroid lineage (clusters 9, 8, 7, 6, and 13), respectively. The B lymphocytes are only directly linked to the HSCs, suggesting that cluster 5 represents a myeloid progenitor population, and no lymphoid progenitors were present in our sample. The inferred lineage tree is therefore consistent with the existence of a common myeloid progenitor population giving rise to erythrocytes, megakaryocytes, granulocytes, and macrophages (Orkin and Zon, 2008). StemID determines the highest score for cluster 1 and, therefore, correctly recovers HSCs among all cell types in the mixture (Figure 4D; Figure S6). The second-highest score discriminates the multipotent myeloid progenitors (cluster 5) from the remaining cell types, and the third-highest score was assigned to the earliest progenitor of the erythroblast lineage. Therefore, the level of multipotency also correlates with the StemID score of bone marrow-derived cells.

The high connectivity of cluster 1 provides evidence for early fate biases already in HSCs. Moreover, the high entropy of HSCs reflects a more uniform transcriptome in individual cells of this population. The entropy distribution across all cells in this cluster is shifted in comparison with all other groups

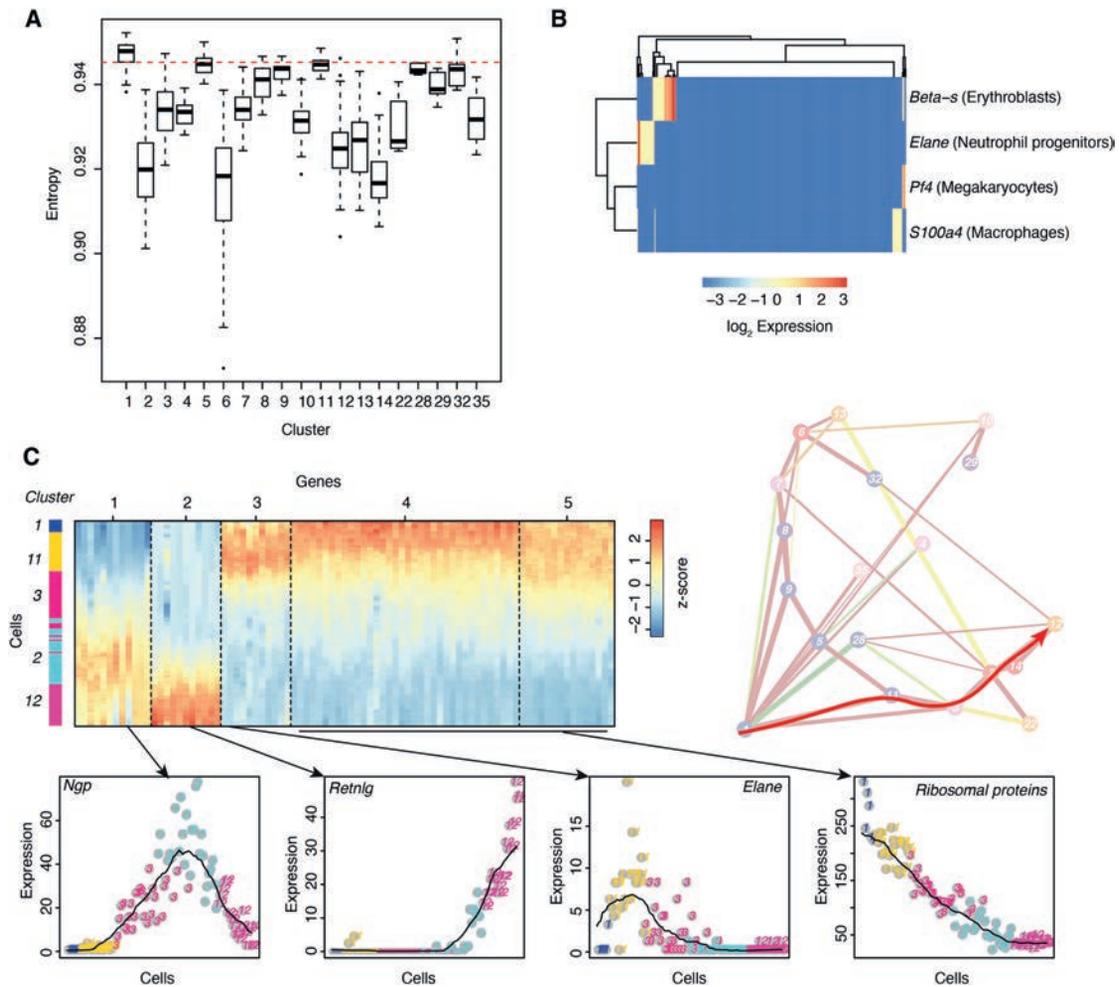


Figure 5. The Multipotency of HSCs Is Reflected by High Transcriptome Entropy

(A) Boxplot of the transcriptome entropy for all RacelD2-derived bone marrow cell types with more than five cells. The boundaries of the box represent the 25% and 75% quantiles, the thick line corresponds to the median, and whiskers extend to the 5% and 95% quantiles. The broken red line indicated the 25% quantile for HSCs (cluster 1).

(B) Two-dimensional clustering of lineage markers in all HSCs (cluster 1). The heatmap shows logarithmic expression.

(C) Self-organizing map (SOM) of Z-score-transformed, pseudo-temporal expression profiles along the neutrophil differentiation trajectory (clusters 1, 11, 3, 2, and 12), indicated by the red arrow superimposed on the lineage tree (Experimental Procedures). The pseudo-temporal order was inferred from the projection coordinates of all cells. The color-coding on the left indicates the cluster of origin. The SOM identified five different modules of co-regulated genes. Examples are shown at the bottom. The clusters of origin are indicated as colors and numbers. The black line represents a moving average (window size 25).

In (A)–(C), only clusters with more than five cells were analyzed.

(Figure 5A). In general, the inter-cluster variability substantially exceeds the intra-cluster variability. The narrow entropy distribution of cluster 1 also rules out a strong dependence on the cell cycle. However, we also observed that 54 of the 276 HSCs (20%) show distinct fate biases, revealed by low expression of lineage-specific marker genes (Figure 5B), a finding that is consistent with a recent report based on lineage tracing (Perié et al., 2015). Because the sensitivity of single-cell sequencing is limited, this number is almost certainly an underestimation. We note that most HSCs (112 of 276) are assigned to the link with the multipotent progenitor (cluster 5). We cannot address whether the observed fate bias persists during differentiation or whether stochastic switching between distinct cell fates occurs during differentiation. Our observation is also consistent

with a recent single-cell transcriptome analysis showing an unexpected heterogeneity of myeloid progenitor cell populations and suggests the existence of an early cell fate bias (Paul et al., 2015). We observe very similar sets of marker genes, as found in this study, but our lineage inference permits an analysis of the temporal dynamics of gene expression. As an example, we extracted all cells from the neutrophil branch (clusters 1, 11, 3, 2, and 12) in pseudo-temporal order derived from the projection coordinates and clustered temporal expression profiles by using self-organizing maps (Experimental Procedures). A Z-score of gene expression values along this trajectory reveals that the RacelD2 clusters represent sets of cells with common modules of co-expressed genes and that gene expression within these modules changes smoothly over time (Figure 5C). Although

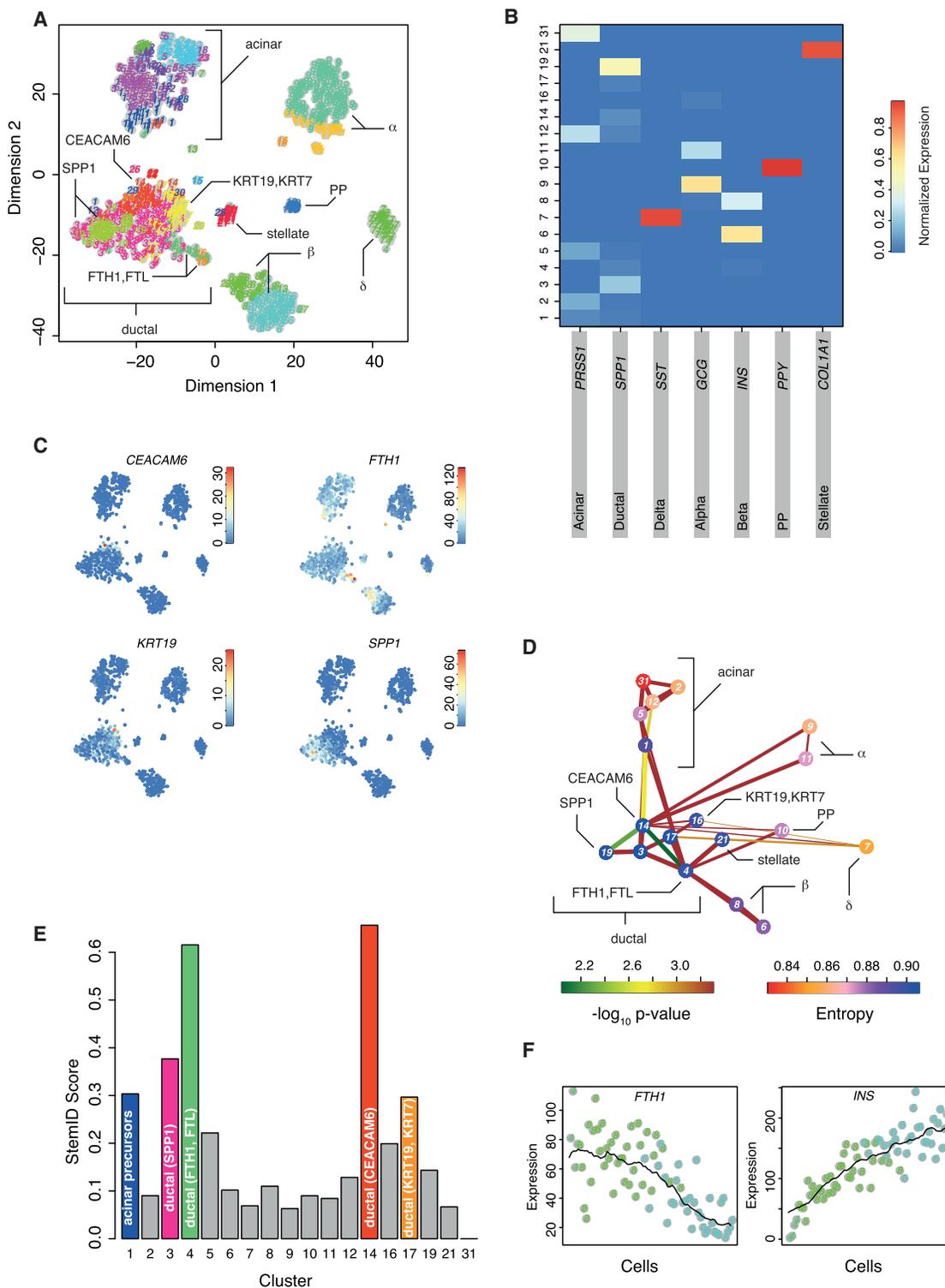


Figure 6. StemID Predicts Human Pancreatic Pluripotent Cells

(A) t-SNE map of transcriptome similarities of human pancreatic cells. RaceID2 clusters are highlighted with different numbers and colors. Cell types identified based on marker gene expression are shown. For ductal cells, marker genes of sub-populations are shown.

(B) Heatmap showing the average expression of known cell type markers across all clusters with more than five cells. For each gene, the sum of expression values over all clusters is normalized to one.

(C) Transcript counts (color legend) of the ductal sub-type markers *CEACAM6*, *FTH1*, *KRT19*, and *SPP1* are highlighted in the t-SNE map.

(D) Inferred pancreatic lineage tree. Only significant links are shown ($p < 0.01$). The color of the link indicates the $-\log_{10} p$ value. The color of the vertices indicates the entropy. The thickness indicates the link score reflecting how densely a link is covered with cells (Experimental Procedures).

(legend continued on next page)

ribosomal protein-encoding genes and other components of the translational machinery slowly decline during differentiation, other genes are transiently switched on in progenitor populations (e.g., *Elane*) or immature neutrophils (e.g., *Ngp*) or only upregulated in mature cells (e.g., *Retnlg*).

Finally, we note that the identification of the HSC population by StemID is robust to changing the contribution of this population to the mixed sample. For example, when only ten HSCs are randomly selected and all others are discarded from the dataset, StemID still assigns the highest score to the small HSC cluster (data not shown).

In summary, StemID could successfully identify the stem cell type in a complex mixture of cells isolated from bone marrow. The inferred lineage tree recovered known trajectories but suggested an early cell fate bias present already in HSCs.

StemID Predicts Multipotent Ductal Cell Populations among Human Adult Pancreatic Cells

After having demonstrated that StemID can robustly identify the stem cell population in two distinct biological systems, we applied the algorithm to predict multipotent cell populations in a less characterized system: the human pancreas. The pancreas consists of acinar cells that produce the digestive enzymes, ductal cells secreting bicarbonate to neutralize stomach acidity, and hormone-producing endocrine cells that regulate hormone metabolism (Jennings et al., 2015). It is unclear which multipotent cells maintain pancreatic homeostasis and can give rise to different mature cell types during regeneration upon injury. Although early studies have suggested that, in humans, these cell populations could reside within the exocrine compartment or that dedifferentiation of exocrine cells could give rise to endocrine cells (Bonner-Weir et al., 2000; Puri et al., 2015), the identity of multipotent cell populations is still unclear (Jiang and Morahan, 2014). We sequenced pancreatic cells from human donors (Experimental Procedures), and application of RaceID2 revealed all major cell types, including different subpopulations of acinar and ductal cells; hormone-producing α , β , δ , and pancreatic polypeptide producing (PP) cells; and stellate cells (Figures 6A and 6B; Figures S5A and S5B). A full list of differentially expressed genes for each cluster is shown in Table S4. In particular, we discovered novel subpopulations of ductal cells. In one of these groups (cluster 14), the cell surface glycoprotein *CEACAM6* was significantly upregulated ($p < 0.01$; Experimental Procedures), whereas components of the ferritin protein (*FTH1*, *FTL*), which is the major intracellular iron storage protein, were significantly upregulated ($p < 0.01$; Experimental Procedures) in the other group (cluster 4) (Figure 6C).

The inferred lineage tree assigns a central position to the ductal cells (Figure 6D; Figures S7C–S7E). Distinct subtypes of ductal cells appear to give rise to different endocrine sub-types and acinar cells. Although differentiation trajectories link cluster 4 to acinar, PP, and β cells, cluster 14 is linked to α and δ cells. Consistently, clusters 4 and 14 acquire the highest StemID score, indicating the highest level of multipotency among the

cell types detected in this system (Figure 6E; Figure S7F). The following ranks of the StemID score were occupied by other ductal sub-types and precursor cells that give rise to two sub-states of acinar cells. Interestingly, cluster 4 also directly connects to stellate cells. Upon injury, these cells can switch to an activated state and migrate to the injured location to participate in tissue repair (Omary et al., 2007).

To collect further evidence that cluster 4 is an endocrine progenitor cell, we plotted the expression of the cluster 4 marker *FTH1* and the β cell marker insulin (*INS*) in single cells residing on the differentiation trajectory connecting these two cell types. Cells were ordered by their projection coordinate. The genes exhibited smooth, anti-correlated gradients suggestive of a continuous transition between these two cell types (Figure 6F). To independently validate this observation, we performed antibody staining against insulin and FTL in human pancreatic tissue sections. We were able to detect individual cells co-expressing insulin and FTL within ductal structures, confirming the existence of cluster 4 cells (Figure 7A). Co-staining of glucagon revealed that these cells specifically produce insulin and not glucagon (Figure 7B), as suggested by our analysis (Figure 6C). Our results indicate that the ferritin-positive sub-population of ductal cells might differentiate into mature β cells.

DISCUSSION

In this study, we present an approach to identify stem cells using single-cell transcriptomics data. Because the physiological state of a cell is an approximate reflection of its transcriptome, it is a reasonable assumption that cell types can be discriminated based on their transcriptome. However, determining the stem cell identity among all rare cell types discovered also requires the derivation of a lineage tree.

To address this task, we combined cell type identification by RaceID2 with a tree reconstruction by guided topology. We first introduce an improved version of our previous RaceID algorithm (Grün et al., 2015) with a more robust initial clustering step. The replacement of k-means by k-medoids leads to increased robustness of clustering for all datasets analyzed in the paper. For the complex intestinal dataset (Figure 3), the fraction of clusters with Jaccard's similarity of > 0.7 is 40% for k-means versus 73% for k-medoids. The corresponding fractions are 58% versus 83% for the bone marrow data and 40% versus 90% for the pancreas data.

To infer differentiation trajectories, we assign every cell onto a specific link between its cluster of origin and another cluster based on the longest projection of the vector connecting the cluster center with the cell position onto these links. This adequately reflects how much a cell has moved from the most representative cell state in the same cluster (the medoid) toward another cell identity (or vice versa). If significantly more cells reside on a link than expected by chance, this provides strong evidence that cells of the cluster of origin exhibit a pronounced transcriptome bias toward another cell fate. In addition, if a

(E) Barplot of StemID scores for pancreatic clusters.

(F) Pseudo-temporal expression profiles for *INS* and *FTH1*. The transcript count is plotted for cells on the link, connecting clusters 4, 8, and 6. Cells are ordered by the projection coordinate.

In (B), (D), and (E), only clusters with more than five cells were analyzed. See also Figure S5.

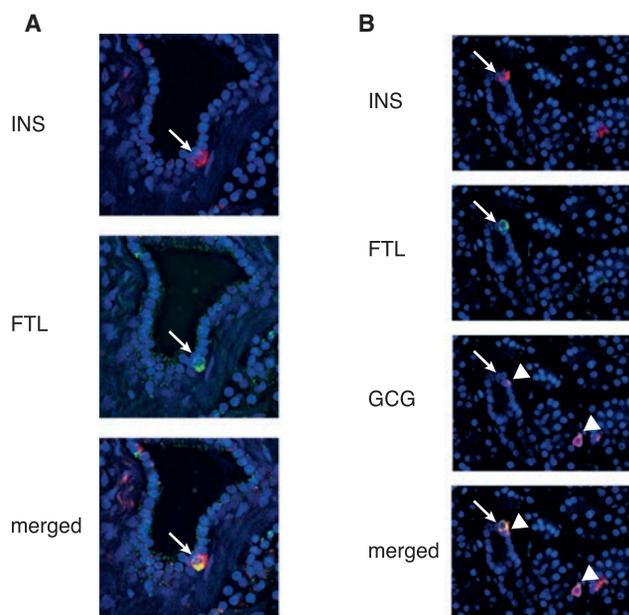


Figure 7. Validation of Putative Endocrine Precursor Cells in Ductal Subpopulations by Antibody Staining

(A and B) Antibody staining for INS and FTH1 in human pancreatic showing a single cell positive for INS and FTH1 residing in the lining of the duct (arrow). (B) Antibody staining for INS, FTH1, and GCG in human pancreatic tissue. Shown is a single cell positive for INS and FTH1 residing in the lining of the duct (arrow) next to a GCG-expressing cell (arrowhead). Another GCG-expressing cell is found nearby (arrowhead). Both GCG-expressing cells are FTH1-negative.

continuum of cell states covers a given link, as evidenced by a high link score, then this link represents a strong candidate for an actual differentiation trajectory. Significant links with reduced link scores, on the other hand, indicate plasticity of the connected cell types in a sense that the transcriptome of a cell type can, to some extent, fluctuate toward another fate.

The quality of our lineage inference is supported by the recovery of known differentiation trajectories in the intestinal epithelium and the bone marrow. Remarkably, we recovered a rare alternative differentiation pathway where *Lgr5*⁺ cells differentiate directly into Paneth cells without intermediate *Dll1*⁺ progenitors (Farin et al., 2014; Sawada et al., 1991). We could also show, for the intestinal and the bone marrow data, that StemID infers a lineage tree with substantially higher resolution in comparison with methods published previously (Haghverdi et al., 2015; Trapnell et al., 2014; Figure S6).

The derived lineage tree for the bone marrow suggested that, in contrast to the classical view of dichotomous differentiation via a hierarchy of increasingly restricted progenitor populations (Giebel and Punzel, 2008), a cell fate bias already exists at stages as early as the HSC stage (Figure 5B). This observation is consistent with a recent single-cell transcriptome analysis revealing heterogeneity of the common myeloid progenitor cell population, indicating early fate bias (Paul et al., 2015). Moreover, direct generation of progenitors restricted to the myeloid fate from mouse HSCs has been described in the past (Yamamoto et al., 2013), and the existence of unipotent

cells within human HSCs (Notta et al., 2016) and classically defined mouse multipotent progenitor populations was shown recently (Perié et al., 2015).

For both model systems, the StemID score, which quantifies very general properties of a cell type (i.e., the number of links and the entropy of the transcriptome), ranks RaceID2-predicted cell types by their level of multipotency. *Lgr5*⁺ CBCs and sorted HSCs acquire the highest score among all cell types of the intestine and bone marrow, respectively, demonstrating the performance of our algorithm. We could further demonstrate the performance of StemID on two previously published datasets (Figure S7) for cells from developing lung epithelium (Treutlein et al., 2014) and differentiating human radial glial cells (Pollen et al., 2015).

Potential problems for the StemID algorithm arise in the absence of intermediate progenitors or the occurrence of unrelated cell types. In the absence of intermediate progenitors, StemID infers a link to a more multipotent population. For example, B lymphocytes in the bone marrow dataset are directly linked to HSCs. It is known that a spectrum of progenitors will reside on this trajectory, and, as we have observed for the other lineages, an early fate bias toward lymphocytes could exist in HSCs. In the absence of intermediate progenitors, a link to a more multipotent population reflects all information on the lineage relationship that can be extracted from the data. If the stem cell itself is missing from the sample, StemID will identify the cell type with the highest level of multipotency. The presence of unrelated cell types in the mixture could lead to false positive links. However, because the feature space is high-dimensional, it is likely that none of the links between an unrelated cell type and the remaining lineage tree will be significantly populated. We also argue that links of mature cell types to related progenitor or stem cell populations were identified with high specificity (oftentimes only a single link in line with previous findings was detected). This makes the occurrence of significant links between unrelated cell types unlikely.

Finally, we used StemID to screen human adult pancreatic cells for multipotent cell populations. It is unclear which adult pancreatic cell types can give rise to the different mature pancreatic lineages during normal tissue turnover or regeneration. Although initial evidence suggested that multipotent cells within the ductal compartment could differentiate into endocrine cells both in humans and mice (Jiang and Morahan, 2014), subsequent lineage-tracing experiments produced contradictory results. Although mouse lineage tracing of carbonic anhydrase II (Ca²⁺)-positive ductal cells revealed that these cells give rise to β cells upon injury (Bonner-Weir et al., 2008), lineage tracing of *Sox9*⁻, *Muc1*⁻, or *Hnf1 β* ⁻ positive cells could not confirm this finding (Furuyama et al., 2011; Kopinke and Murtaugh, 2010; Kopp et al., 2011; Solar et al., 2009). Using StemID, we were able to predict distinct sub-populations of ductal cells with varying differentiation potential. Although ductal cells marked by high levels of CEACAM6 are predicted to differentiate into α , δ , and PP cells, another sub-population expressing high levels of the ferritin complex primarily appears to give rise to β cells and acinar cells. We note that the latter sub-population does not express any of the markers used in previous lineage-tracing experiments, but we caution that expression of these genes might be too low to be reliably detected by single-cell mRNA sequencing.

We further remark that β cell differentiation in the adult pancreas might not be conserved between human and mouse.

We provide the well documented R source code for RaceID2 and the StemID algorithm at <https://github.com/dgrun/StemID>. We hope that StemID will be useful for a better understanding of differentiation dynamics in a variety of systems.

EXPERIMENTAL PROCEDURES

Lineage-Tracing Experiments

For lineage-tracing experiments, we injected 0.4 mg tamoxifen into 3-month-old Lgr5-CreERT2 C57Bl6/J mice bred to Rosa26LSL-YFP reporter mice.

Isolation of Crypts from Mouse Small Intestine

Crypts were isolated from mice as described previously (Sato et al., 2009). See the Supplemental Experimental Procedures for more details.

Human Islet Isolation, Dispersion, and Sorting

Pancreatic cadaveric tissue was procured from a multiorgan donor program and only used when the pancreas could not be used for clinical pancreas or islet transplantation, according to national laws, and when research consent was present. Human islet isolations were performed in the islet isolation facility of the Leiden University Medical Center according to a modified protocol originally described by Ricordi et al. (1988). See the Supplemental Experimental Procedures for details regarding culturing and cell sorting.

Immunofluorescence

Pancreatic tissue samples were fixed overnight in 4% formaldehyde (Klinipath), stored in 70% ethanol, and subsequently embedded in paraffin. After deparaffinization and rehydration in xylene and ethanol, respectively, antigen retrieval was performed in citric buffer for 20 min. Sections were blocked with 2% normal donkey serum and 1% lamb serum in PBS. Primary antibodies were rabbit anti-Ftl (ab69090), mouse anti-glucagon (ab10988), and guinea pig anti-insulin (ab7842). Alexa Fluor-conjugated secondary antibodies against rabbit, mouse, and guinea pig immunoglobulin G (IgG) (Life Technologies; A11008, A10037, and A21450) were used at a dilution of 1:200. Nuclear counterstaining was done by embedding with DAPI Vectashield (Vector Laboratories, H-1500). Imaging was performed on a Leica SP8 confocal microscope using hybrid detectors.

Preparation of Mouse Hematopoietic Cells

We used C57Bl/6 female or male mice from 23 to 52 weeks bred in our facility. Experimental procedures were approved by the Dier Experimenten Commissie of the Royal Netherlands Academy of Arts and Sciences and performed according to the guidelines. Bone marrow was isolated from femur and tibia by flushing Hank's balanced salt solution (HBSS, Invitrogen) without calcium or magnesium, supplemented with 1% heat-inactivated fetal calf serum (FCS) (Sigma). See the Supplemental Experimental Procedures for details regarding single cell isolation.

Single-Cell Sequencing Library Preparation

The protocol was carried out as described previously (Grün et al., 2015). See the Supplemental Experimental Procedures for a detailed description.

Quantification of Transcript Abundance

Read mapping and quantification were done as described previously (Grün et al., 2015). See the Supplemental Experimental Procedures for a detailed description.

RaceID2 and StemID

A brief overview is given in the Results. The algorithm and follow-up analyses are described in full detail in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession numbers for the RNA sequencing datasets reported in this paper are GEO: GSE76408, GSE76983, and GSE81076.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and four tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2016.05.010>.

AUTHOR CONTRIBUTIONS

D.G. and A.v.O. conceived the study. D.G. developed the algorithm and performed all computational analyses. Single-cell sequencing of pancreatic cells and antibody staining were performed by M.J.M. with the help of G.D. Single-cell sequencing of intestinal cells was performed by K.W. with the help of A.L., J.v.E., and M.v.d.B. Single-cell sequencing of bone marrow cells was performed by J.C.B. E.J. helped with antibody staining. D.G. wrote the manuscript, and all authors read and edited the manuscript. A.v.O. supervised D.G., M.J.M., K.W., and J.C.B. and the project itself. E.J.P.d.K. supervised G.D. and E.J. H.C. supervised M.v.d.B. and J.v.E.

ACKNOWLEDGMENTS

This work was supported by European Research Council Advanced Grant ERC-AdG 294325-GeneNoiseControl, a Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) Vici award, the DON Foundation, and the Dutch Diabetes Research Foundation.

Received: February 24, 2016

Revised: April 4, 2016

Accepted: May 12, 2016

Published: June 23, 2016

REFERENCES

- Anavy, L., Levin, M., Khair, S., Nakanishi, N., Fernandez-Valverde, S.L., Degnan, B.M., and Yanai, I. (2014). BLIND ordering of large-scale transcriptional developmental timecourses. *Development* *141*, 1161–1166.
- Banerji, C.R.S., Miranda-Saavedra, D., Severini, S., Widschwendter, M., Enver, T., Zhou, J.X., and Teschendorff, A.E. (2013). Cellular network entropy as the energy potential in Waddington's differentiation landscape. *Sci. Rep.* *3*, 3039.
- Barker, N. (2014). Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat. Rev. Mol. Cell Biol.* *15*, 19–33.
- 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.
- Bendall, S.C., Davis, K.L., Amir, A.D., Tadmor, M.D., Simonds, E.F., Chen, T.J., Shenfeld, D.K., Nolan, G.P., and Pe'er, D. (2014). Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell* *157*, 714–725.
- Bonner-Weir, S., Taneja, M., Weir, G.C., Tatkiewicz, K., Song, K.H., Sharma, A., and O'Neil, J.J. (2000). In vitro cultivation of human islets from expanded ductal tissue. *Proc. Natl. Acad. Sci. USA* *97*, 7999–8004.
- Bonner-Weir, S., Inada, A., Yatoh, S., Li, W.-C., Aye, T., Toschi, E., and Sharma, A. (2008). Transdifferentiation of pancreatic ductal cells to endocrine beta-cells. *Biochem. Soc. Trans.* *36*, 353–356.
- Brennecke, P., Anders, S., Kim, J.K., Kołodziejczyk, A.A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S.A., Marioni, J.C., and Heisler, M.G. (2013). Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* *10*, 1093–1095.
- Buczacki, S.J.A., Zecchini, H.I., Nicholson, A.M., Russell, R., Vermeulen, L., Kemp, R., and Winton, D.J. (2013). Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* *495*, 65–69.
- Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S.M., Reth, M., Höfer, T., and Rodewald, H.-R. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* *518*, 542–546.

- Eaves, C.J. (2015). Hematopoietic stem cells: concepts, definitions, and the new reality. *Blood* 125, 2605–2613.
- Eldar, A., and Elowitz, M.B. (2010). Functional roles for noise in genetic circuits. *Nature* 467, 167–173.
- Farin, H.F., Karthaus, W.R., Kujala, P., Rakhshandehroo, M., Schwank, G., Vries, R.G.J., Kalkhoven, E., Nieuwenhuis, E.E.S., and Clevers, H. (2014). Paneth cell extrusion and release of antimicrobial products is directly controlled by immune cell-derived IFN- γ . *J. Exp. Med.* 211, 1393–1405.
- Furuyama, K., Kawaguchi, Y., Akiyama, H., Horiguchi, M., Kodama, S., Kuhara, T., Hosokawa, S., Elbahrawy, A., Soeda, T., Koizumi, M., et al. (2011). Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. *Nat. Genet.* 43, 34–41.
- Giebel, B., and Punzel, M. (2008). Lineage development of hematopoietic stem and progenitor cells. *Biol. Chem.* 389, 813–824.
- Grün, D., Kester, L., and van Oudenaarden, A. (2014). Validation of noise models for single-cell transcriptomics. *Nat. Methods* 11, 637–640.
- Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 525, 251–255.
- Haghverdi, L., Buettner, F., and Theis, F.J. (2015). Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* 31, 2989–2998.
- Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., and Amit, I. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 343, 776–779.
- Jennings, R.E., Berry, A.A., Strutt, J.P., Gerrard, D.T., and Hanley, N.A. (2015). Human pancreas development. *Development* 142, 3126–3137.
- Jiang, F.-X., and Morahan, G. (2014). Pancreatic stem cells remain unresolved. *Stem Cells Dev.* 23, 2803–2812.
- Kiel, M.J., Yilmaz, O.H., Iwashita, T., Yilmaz, O.H., Terhorst, C., and Morrison, S.J. (2005). SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* 121, 1109–1121.
- Kopinke, D., and Murtaugh, L.C. (2010). Exocrine-to-endocrine differentiation is detectable only prior to birth in the uninjured mouse pancreas. *BMC Dev. Biol.* 10, 38.
- Kopp, J.L., Dubois, C.L., Schaffer, A.E., Hao, E., Shih, H.P., Seymour, P.A., Ma, J., and Sander, M. (2011). Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. *Development* 138, 653–665.
- Lancaster, M.A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L.S., Hurler, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature* 501, 373–379.
- Macosko, E.Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161, 1202–1214.
- Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., et al. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* 351, 126–127.
- Omary, M.B., Lugea, A., Lowe, A.W., and Pandol, S.J. (2007). The pancreatic stellate cell: a star on the rise in pancreatic diseases. *J. Clin. Invest.* 117, 50–59.
- Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* 132, 631–644.
- Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344, 1396–1401.
- Paul, F., Arkin, Y., Giladi, A., Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Winter, D., Lara-Astiaso, D., Gury, M., Weiner, A., et al. (2015). Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* 163, 1663–1677.
- Perié, L., Duffy, K.R., Kok, L., de Boer, R.J., and Schumacher, T.N. (2015). The Branching Point in Erythro-Myeloid Differentiation. *Cell* 163, 1655–1662.
- Piras, V., Tomita, M., and Selvarajoo, K. (2014). Transcriptome-wide variability in single embryonic development cells. *Sci. Rep.* 4, 7137.
- Pollen, A.A., Nowakowski, T.J., Chen, J., Retallack, H., Sandoval-Espinosa, C., Nicholas, C.R., Shuga, J., Liu, S.J., Oldham, M.C., Diaz, A., et al. (2015). Molecular identity of human outer radial glia during cortical development. *Cell* 163, 55–67.
- Puri, S., Folias, A.E., and Hebrok, M. (2015). Plasticity and dedifferentiation within the pancreas: development, homeostasis, and disease. *Cell Stem Cell* 16, 18–31.
- Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135, 216–226.
- Ricordi, C., Lacy, P.E., Finke, E.H., Olack, B.J., and Scharp, D.W. (1988). Automated method for isolation of human pancreatic islets. *Diabetes* 37, 413–420.
- Ridden, S.J., Chang, H.H., Zygalkakis, K.C., and MacArthur, B.D. (2015). Entropy, Ergodicity, and Stem Cell Multipotency. *Phys. Rev. Lett.* 115, 208103.
- 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.
- Sawada, M., Takahashi, K., Sawada, S., and Midorikawa, O. (1991). Selective killing of Paneth cells by intravenous administration of dithizone in rats. *Int. J. Exp. Pathol.* 72, 407–421.
- Shannon, C.E. (1948). *A Mathematical Theory of Communication*. Bell Syst. Tech. J. 27, 379–423, 623–656.
- Solar, M., Cardalda, C., Houbracken, I., Martín, M., Maestro, M.A., De Medts, N., Xu, X., Grau, V., Heimberg, H., Bouwens, L., and Ferrer, J. (2009). Pancreatic exocrine duct cells give rise to insulin-producing beta cells during embryogenesis but not after birth. *Dev. Cell* 17, 849–860.
- Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* 63, 411–423.
- Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32, 381–386.
- Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 509, 371–375.
- van der Flier, L.G., and Clevers, H. (2009). Stem cells, self-renewal, and differentiation in the intestinal epithelium. *Annu. Rev. Physiol.* 71, 241–260.
- van Es, J.H., Sato, T., van de Wetering, M., Lyubimova, A., Nee, A.N.Y., Gregorieff, A., Sasaki, N., Zeinstra, L., van den Born, M., Korving, J., et al. (2012). Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nat. Cell Biol.* 14, 1099–1104.
- Wilson, N.K., Kent, D.G., Buettner, F., Shehata, M., Macaulay, I.C., Calero-Nieto, F.J., Sánchez Castillo, M., Oedekoven, C.A., Diamanti, E., Schulte, R., et al. (2015). Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* 16, 712–724.
- Yamamoto, R., Morita, Y., Oebara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* 154, 1112–1126.
- Zeisel, A., Muñoz-Manchado, A.B., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347, 1138–1142.

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?

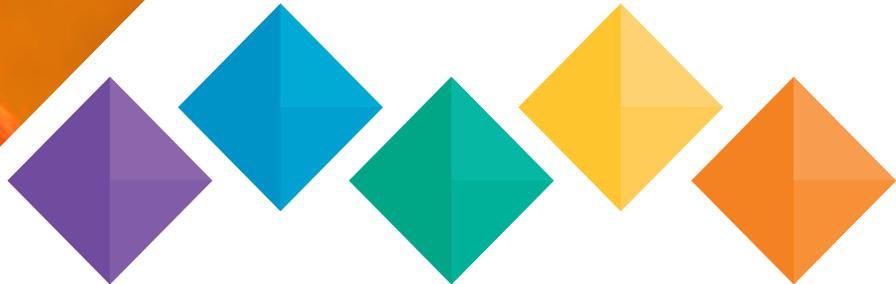
Check it out at www.cell.com/figure360

Figure360

CellPress

GUANGZHOU CHINA 2017

10-12 NOVEMBER



**STEM CELLS:
THE NEXT GENERATION**

KEYNOTE SPEAKER



Rudolf Jaenisch

*Whitehead Institute for
Biomedical Research, US*

THE ISSCR INTERNATIONAL SYMPOSIA RETURN TO CHINA! Join us in Guangzhou 10-12 November 2017, for forward-thinking sessions that explore the development of stem cell therapy. From the foundational to the translational, the ISSCR continues to bring you the latest information from the field's leading minds: the "next generation" of stem cell research and therapeutic application.

SESSION TOPICS INCLUDE:

Stem Cell State and Differentiation
Organoids and Disease Modeling
Cell-Based Therapies
Ethics of Genome Engineering

Registration Open: 10 May 2017
Early Registration Deadline: 12 July 2017
Advance Registration Closes: 11 October 2017

Abstract Submissions Open: 10 May 2017
Abstract Submission Deadline: 12 July 2017

LEARN MORE OR REGISTER AT ISSCR.ORG/GUANGZHOU

/// NEW STEMFLEX MEDIUM ///
DESIGNED FOR BETTER EVERYTHING



Enhanced flexibility and superior performance in today's applications

Introducing a media that works for the applications stem cell researchers are doing today, including gene editing and single-cell passaging. Gibco™ StemFlex™ Medium supports optimal, reproducible results while also letting you choose your matrix, passaging reagent, and feeding schedule.

Do what you do better at
thermofisher.com/stemflex

ThermoFisher
SCIENTIFIC