

# Cell Stem Cell

Best of 2014



# PHENOTYPING: A MAJOR LEAP FORWARD

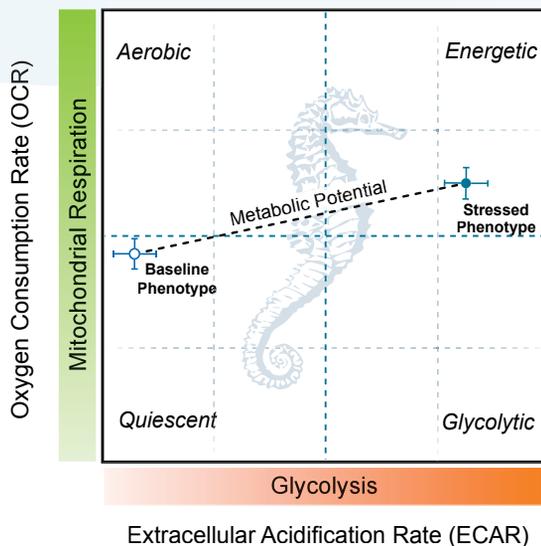
Determine the functional impact of genes on metabolism in an hour.



## XFP Cell Energy Phenotype Test Kit

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

Get a **FREE** energy phenotype analysis of your cells at [www.seahorsebio.com/pheno](http://www.seahorsebio.com/pheno)





don whitley  
scientific

# Simulate the low oxygen environment of the stem cell niche

Closed cell culture with O<sub>2</sub>, CO<sub>2</sub>, temperature & humidity control



## NEW: WHITLEY H135 HYPOXYSTATION

The tallest, widest, deepest hypoxic chamber in the Whitley range can accommodate a variety of items of equipment such as live cell imaging devices, microscopes, plate readers, etc. Cell manipulations can be performed in your required low oxygen conditions as the entire chamber is a working and incubation space.

- Control O<sub>2</sub> in 0.1% increments to 20% and CO<sub>2</sub> in 0.1% increments to 15%.
- A removable front fitted (with two patented oval ports) to facilitate the transfer of large items of equipment for use inside the chamber.
- Innovative full colour touch screen that is Ethernet-enabled for remote access.
- A rapid 12 litre airlock that completes a cycle in just 60 seconds.

### OPTIONS INCLUDE:

- The Whitley Automatic Humidification System – sterile, on-demand moisture control.
- The Whitley Internal HEPA Filtration System with Enhanced Biological Containment.

North American distributors:

**HYPOXYGEN**

Sales: +44 (0)1274 595728 / [www.hypoxystation.co.uk](http://www.hypoxystation.co.uk)

# AUTOMATED CELL SORTING

NOW AVAILABLE AT A BENCH NEAR YOU



## Presenting automation that lets you sort cells right in your lab.

The S3e™ Cell Sorter is designed to make cell sorting simple and accessible to all. Advanced automation features eliminate the need for extensive flow cytometry expertise by allowing instrument setup and operation with minimal user input while a small footprint and affordable price ensure that the S3e fits into almost any lab. It's time to put the focus on research and take the how, when, and where out of cell sorting.

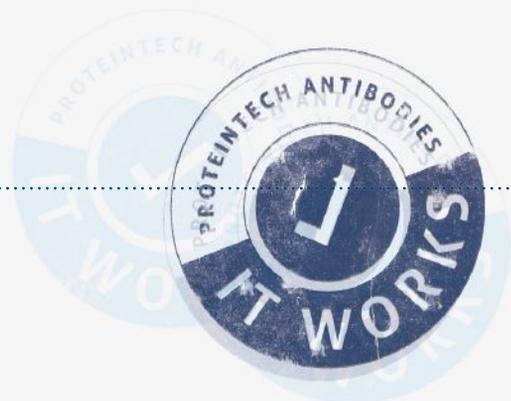
Learn more about the S3e Cell Sorter at [bio-rad.com/info/PersonalSorter](https://www.bio-rad.com/info/PersonalSorter)

**BIO-RAD**



# SETTING THE BENCHMARK IN ANTIBODIES

- Proteintech makes every single antibody in its 12,000-strong catalog
- Exclusively available through Proteintech and its approved distributors.
- Antibody specificity verified with siRNA treated samples
- It works in all species and applications or get a full money back refund - No questions asked

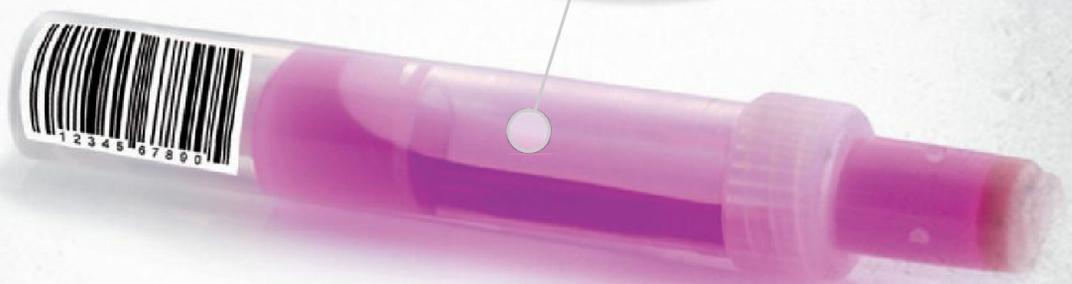
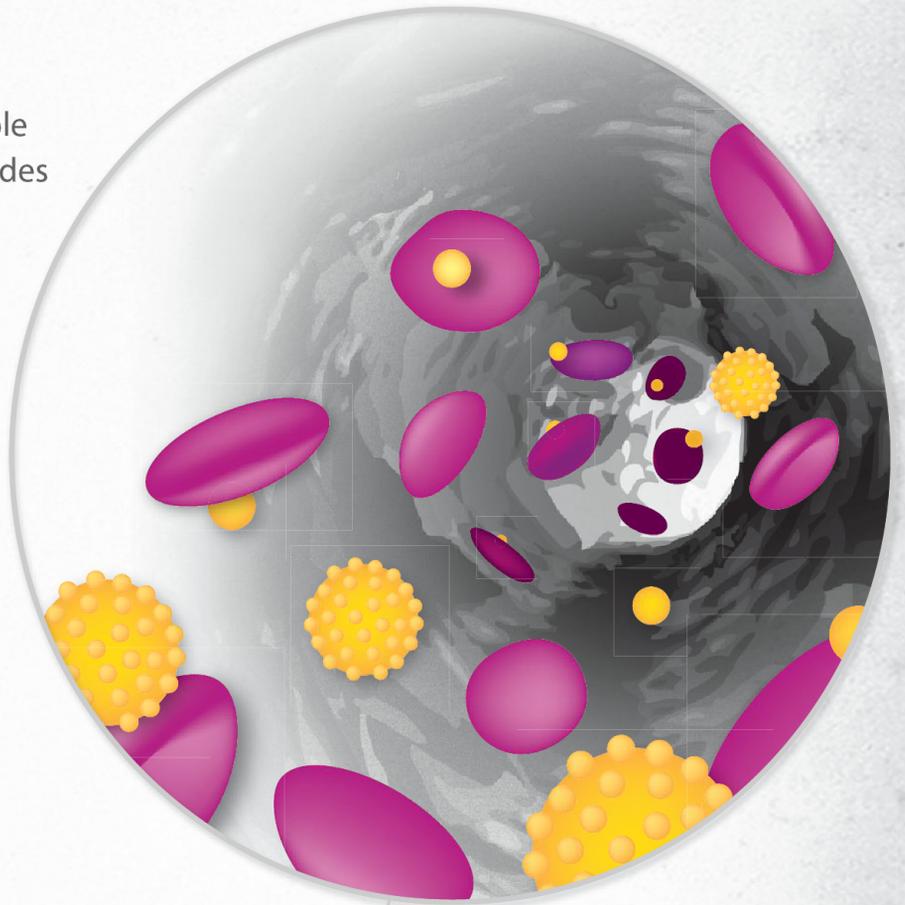


# truCulture®

Investigate Immune Responses Using Myriad RBM's  
TruCulture®\* System.

TruCulture is an innovative whole blood culture system that provides a standardized collection, culturing, and processing method for reliable and functional *ex vivo* analysis of immune responses.

It is ideal for clinical research, drug development and characterization of inflammatory, autoimmune and infectious diseases.



Contact Myriad RBM and find out why this novel tool is considered indispensable by your peers.

**MyriadRBMIS@myriadrbm.com 512-835-8026 or 1-800-RBM-MAPS**

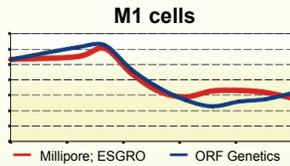
MYRIAD  RBM®

[myriadrbm.com/truculture](http://myriadrbm.com/truculture)

\* For Research Use Only. Not Intended for Use in Diagnostic Procedures.

# THE PATENT FOR mLIF HAS EXPIRED IN USA

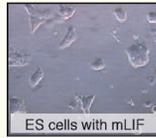
## EQUAL QUALITY; LOWER COST ALTERNATIVE



### Equal bioassay performance

„We have tested the recombinant murine LIF from ORF Genetics and found it to be very active and exactly as good as ESGRO@LIF (Millipore) in mLIF standard Bioassay“

Raphael Nir, PhD  
President SBH Sciences, USA



### Comparable results in mESC cultivation

„We found that mLIF from ORF Genetics and Millipore were comparable and both gave nice looking ES cell colonies. After testing it, we only use mLIF from ORF.“

Dr. Gudrun Valdimarsdottir  
Faculty of Medicine, University of Iceland



### Proven track record in knock-out/ knock-in mice development

„We have used mLIF from ORF Genetics for many years in our services of developing knock-out/knock-in mice with great success. We are very pleased to have switched our mLIF source from Millipore to ORF Genetics.“

Dr. Doron Shmerling  
Managing Director of PolyGene AG, Switzerland

- ▶ Order at: [orders@orfgenetics.com](mailto:orders@orfgenetics.com)
- ▶ Visit us at: [www.orfgenetics.com/ISOkine/mLIF/](http://www.orfgenetics.com/ISOkine/mLIF/)

- Examples of the many publications in 2014 applying mLIF from ORF
- Pfaffeneder et al. *Nature Chemical Biology*. 2014. Vol 10, Issue 3, p. 574-581
  - Pearton et al. *Developmental Biology*. 2014. Vol 392, p. 344-357

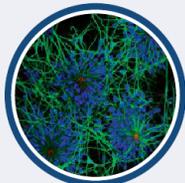
Promotion  
10µg  
for \$10\*

\*Promotional offer expires 31<sup>st</sup> of August 2015

LIST PRICES: 10µg: \$110 100µg: \$390 1000µg: \$2450



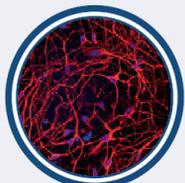
## Succeed with Axol - your human cell culture specialists!



We produce a highly validated, human iPSC-derived cells and are providing a growing range of human primary cells. The package wouldn't be complete without optimized media, coating solutions and other reagents.



Axol Innovation Labs work hard to improve on existing media and reagents as well develop new products for human cell culture. A recent development is our xeno-free neural cell culture system; free from animal components, reducing variability and improving consistency between experiments!



We have a passion for great science, delivering epic support and innovating future products to help our customers advance faster in their research.

Find out more at

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





**“Together we  
can advance  
cell therapies  
worldwide.”**

**Pall Life Sciences**

**Pall Life Sciences delivers the expertise and experience to help our customers realize efficient, effective development and commercialization of cell therapies.**

Working with Pall, you'll benefit from our industry-leading cell manufacturing technologies and services – from our unique single-use bioreactor systems to our innovative SoloHill microcarriers. You'll be fully equipped to overcome the challenges of manufacturing live cells for therapeutics, making your path to industrialization smoother than ever.

Pall's business philosophy is about collaboration and helping you achieve your goals. Together, we can advance the development of cutting-edge cell therapies... and dramatically improve the lives of patients worldwide.

**Pall Life Sciences**

*Your vision. Our expertise. Their future.*

**[www.pall.com/cellculture](http://www.pall.com/cellculture)**

## Foreword

---



Three years ago, Cell Press launched the “Best of” reprint collections across a number of our journals, including *Cell Stem Cell*. We are pleased to introduce the latest edition of *Best of Cell Stem Cell*, which focuses on articles published over the course of 2014. For this edition, we made our selection by looking at review and research articles with the highest number of full-text HTML and PDF downloads since publication and then choosing a representative group from each of the two published volumes. We have organized the highlighted articles in the order that you would see them in a monthly issue and hope you will enjoy browsing through them in the same way you would browse through the journal each month.

We recognize that no one measurement can be indicative of “the best” research 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 2014.

We hope that you will enjoy reading this special collection and that you will visit [www.cell.com/cell-stem-cell/home](http://www.cell.com/cell-stem-cell/home) to see the latest findings that we have had the privilege to publish, presented in the new website format that we launched last year. While you're there, check out our regular rotation of Featured Five Reviews and enjoy free access to all Cell Press articles published over a year ago. If you prefer to read articles on a mobile device, you may also want to use our Journal Reader app, now available across multiple platforms from [www.cell.com/mobile](http://www.cell.com/mobile). Also, be sure to visit [www.cell.com](http://www.cell.com) to find other high-quality articles published in the full collection of Cell Press journals.

Last but certainly not least, we are grateful for the generosity of our many 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](mailto:jchristison@cell.com)  
p: 617-397-2893  
t: @CellPressBiz



DEF-CS 500 Culture System

# Direct comparison of the Cellartis® DEF-CS™ 500 Culture System with different vendors' systems for culturing human pluripotent stem cells

The Cellartis DEF-CS 500 Culture System is a fully defined, feeder-free culture system for the expansion of human induced pluripotent stem (hiPS) cells.

With this system the cells are passaged enzymatically, and the cells maintain pluripotency and stable karyotype for more than 30 passages.

In this application note, the Cellartis DEF-CS system was directly compared to four vendors' feeder-free culture systems for expansion and maintenance of hiPS cells. A feeder cell-based culture system was used as a control. Pluripotency and growth rate analysis indicate that the Cellartis culture system provides robust growth while efficiently maintaining cells in an undifferentiated state.

## Methods

### Cell culture

The hiPS cell line 253G1 was thawed, seeded at a cell density of  $1-3 \times 10^4$  cells/cm<sup>2</sup>, and maintained for five weeks in DEF-CS, and in culture systems from four different vendors (vendor 1, 2, 3 and 4). The hiPS cells were cultured on feeder-free coatings specific for each culture system and were handled according to each manufacturer's recommendations. As a control, 253G1 hiPS cells were cultured on a mitomycin-C treated STO feeder cell layer.

### Growth rate and flow cytometry

After three weeks of adaptation to each culture system, the growth rate of 253G1 hiPS cells was calculated for the next 20 days by plotting the number of cells against days for each culture system. Briefly, 253G1 hiPS cells from each culture system were detached using the reagents recommended by each vendor. To count cells, single cell suspensions were generated by introducing a second digestion step for the aggregate cell systems using TrypLE™ Select enzyme (Life Technologies). Cell number in the single-cell suspensions was calculated manually.

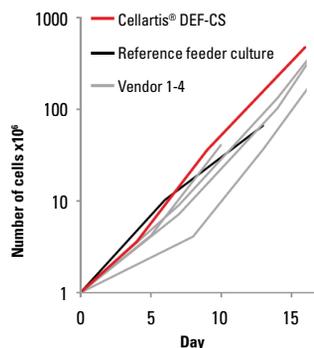


Figure 1. Comparison of growth rate between culture systems

After five weeks of culture, the hiPS cells were collected and incubated with TRA-1-60 and SSEA-4 antibodies conjugated to Alexa Fluor 488 and phycoerythrin, respectively. The labeled cells were analyzed by flow cytometry (FC); side scatter (SSC), forward scatter (FSC), and percentage of TRA-1-60 and SSEA-4 positive cells were quantified for cells grown in the different culture systems.

## Results

The cells were acclimated to each culture system for three weeks. Then, growth rate and pluripotency were characterized for cells growing in each system. The Cellartis DEF-CS culture system promoted a high and robust growth rate for 253G1 hiPS cells (Figure 1).

It has previously been shown that SSC is highly heterogeneous in undifferentiated pluripotent stem cells and predicts clonogenic self-renewal (Ramirez *et al.*, 2013). Cells grown in the various culture systems were collected and analyzed by FC. Cells grown in the DEF-CS system displayed more heterogeneity in SSC than the cell populations grown in the other vendors' culture systems (Figure 2).

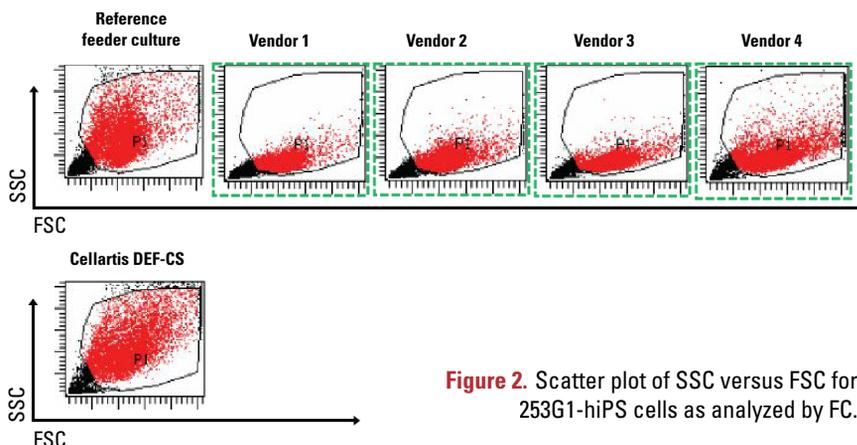


Figure 2. Scatter plot of SSC versus FSC for 253G1-hiPS cells as analyzed by FC.

Cells were also analyzed for expression of the stem cell markers TRA-1-60 and SSEA-4. FC analysis indicated that cells grown in the Cellartis DEF-CS culture system maintain the highest proportion of TRA-1-60+ and SSEA-4+ cells, markers indicative of pluripotency (Figure 3).

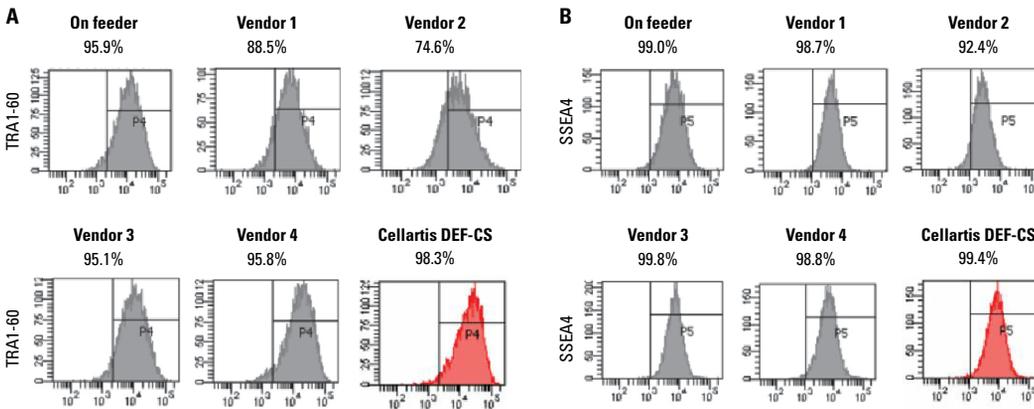
## Conclusions

The Cellartis DEF-CS culture system allowed for efficient and robust expansion of 253G1 hiPS cells similar to other vendors' culture systems. Culture in the DEF-CS system resulted in a higher proportion of pluripotent stem cells that express the TRA-1-60 and SSEA-4 stem cell markers than the majority of the other systems

tested. The highly heterogeneous SSC versus FSC pattern was similar for cells grown in the Cellartis DEF-CS system and the control feeder culture system, but was different for other vendors' culturing systems. Taken together, these data indicate that the DEF-CS system resulted in the most robust and stable growth and was best able to maintain hiPS cells in an undifferentiated state.

## Reference

Ramirez J.M. *et al.* (2013) Side Scatter Intensity Is Highly Heterogeneous in Undifferentiated Pluripotent Stem Cells and Predicts Clonogenic Self-Renewal. *Stem Cells Dev.* **22**:1851–60.



**Figure 3.** Expression of the pluripotency markers TRA1-60 and SSEA-4. After five weeks in culture, TRA1-60 (A) and SSEA-4 (B) expression in 253G1-hiPS cells was analyzed by FC.

## PRODUCTS

Cat. #	Product	Description
Y30010	DEF-CS 500 Culture System	Complete kit for culturing human pluripotent stem cells, including 500 mL basal medium, additives and coating compound <a href="mailto:info-cellartis@takara-clontech.eu">Free sample available, contact info-cellartis@takara-clontech.eu</a>
Related products		
Y00260	DEF-hiPSC™ ChiPSC4	From source cell line DEF-hiPSC™ ChiPSC4 and DEF-hiPSC™ ChiPSC18, cultured and frozen in DEF-CS – Approximately 3x10 <sup>6</sup> cells/vial
Y00300	DEF-hiPSC™ ChiPSC18	

Learn more at [www.clontech.com/stemcells](http://www.clontech.com/stemcells)



### Notice to Purchaser

Your use of these products and technologies is subject to compliance with any applicable licensing requirements described on the product's web page at <http://www.clontech.com>. It is your responsibility to review, understand and adhere to any restrictions imposed by such statements. Cellartis® products are to be used for research purposes only. They may not be used for any other purpose, including, but not limited to, use in drugs, in vitro diagnostic purposes, therapeutics, or in humans. Cellartis® products may not be transferred to third parties, resold, modified for resale, or used to manufacture commercial products or to provide a service to third parties without prior written approval of Takara Bio Europe AB.

### Takara Bio Europe

orders@takara-clontech.eu • tech@takara-clontech.eu  
 Europe : +33.(0)1 3904 6880 • Austria : 0800 296 141 • Germany : 0800 182 5178 • Switzerland : 0800 563 629 • United Kingdom : 0800 234 8063  
 For research use only. Not for use in diagnostic or therapeutic procedures. Not for resale. Takara and the Takara logo are trademarks of TAKARA HOLDINGS, Kyoto, Japan. Clontech, the Clontech logo, and that's GOOD science! are trademarks of Clontech Laboratories, Inc. Cellartis, DEF-CS, DEF-hiPSC and the Cellartis logo are trademarks of Takara Bio Europe AB. TrypLE is a trademark of Life Technologies Corporation. All other marks are the property of their respective owners. Certain trademarks may not be registered in all jurisdictions. © Takara Bio Europe SAS

Now with new Atmospheric Control Unit (ACU) for live cell-based assays

# Monochromator. Reinvented!

CLARIOstar

## CLARIOstar®: revolutionary LVF Monochromator™ technology

The CLARIOstar is the ideal multi-mode reader for assay development. BMG LABTECH's proprietary LVF Monochromators allow for highest flexibility without compromising on sensitivity.

- LVF Monochromators with Linear Variable Filters and Dichroic Mirror
- Filter-like performance with increased sensitivity over conventional monochromators
- Continuously adjustable bandwidths (8-100 nm)
- Use monochromators, filters, or a combination of both
- Integrated fluorophore library for easy wavelength selection
- Ultrafast absorbance scans from 220 to 1000 nm in < 1 sec/well



Made in Germany

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

**BMG LABTECH**

*The Microplate Reader Company*

## LIGHT RE-ENGINEERED

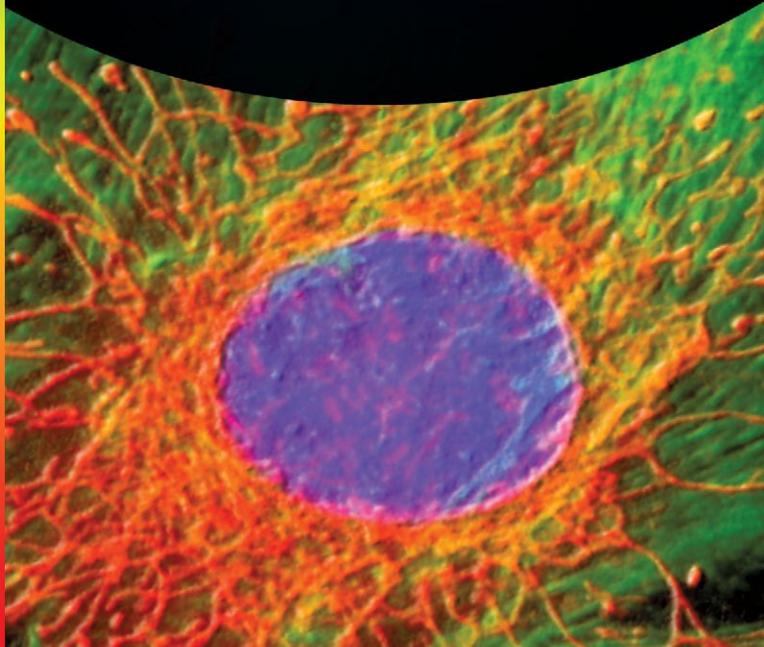


**Custom Engineering  
and Manufacturing**

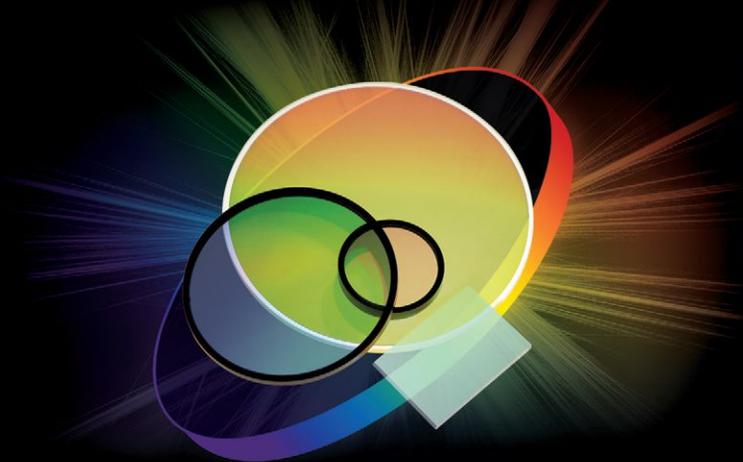
Optical System Layout and Optimization

Optical Components, Custom and OTS • Low and High Volume Builds

Sourcing and Supply Chain Management • Industrial Design



## PROTOTYPE TO PRODUCTION

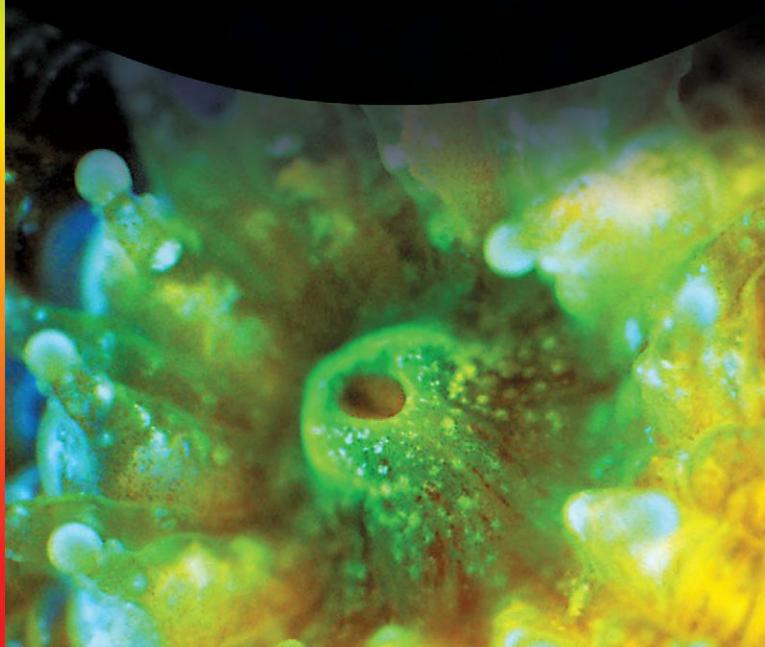


**Durable High Transmission  
Sputtered Filters from UV to NIR**

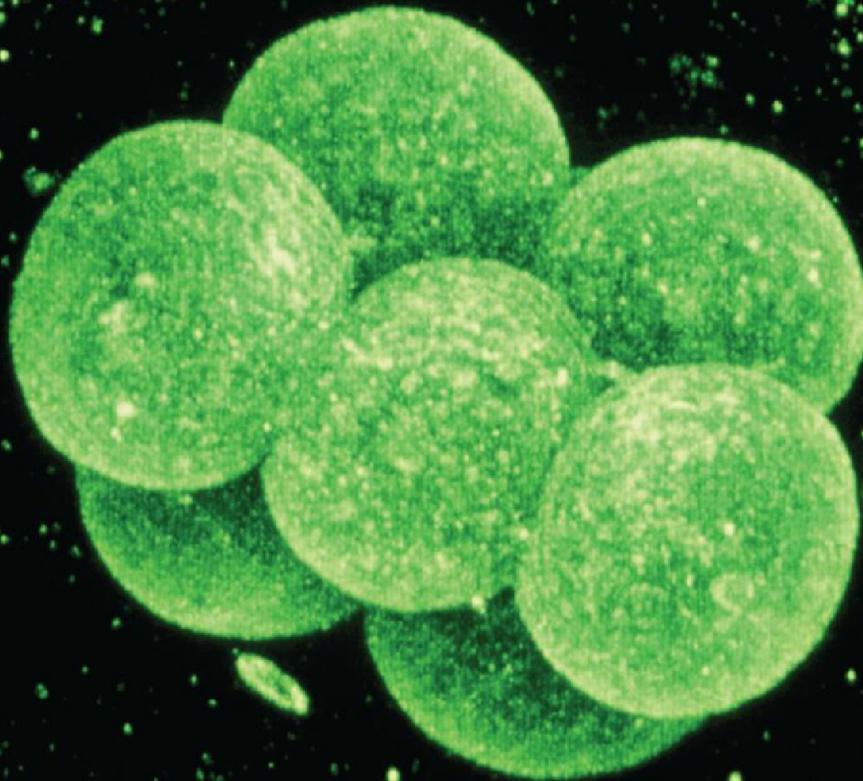
Flow Cytometry • Machine Vision • Point-of-Care

Laser Applications • Raman • Fluorescence

Astronomy • High Resolution Imaging



# Pluripotency when you need it. . .



## ... ATCC® Stem Cells

Our growing selection of pertinent stem cells, media, and reagents is only part of the story. We support every stage of your stem cell research and assay development: from assuring pluripotency, to reprogramming your cells into functional tissues or disease models, to providing expert stem cell culture guidance. Our resources include:

- Mesenchymal stem cells
- Gender and ethnic diverse iPSCs
- Lineage differentiation kits
- Hematopoietic stem cells
- Transfection protocols
- Starter kits, media, and reagents



Discover more with ATCC stem cell resources at [www.atcc.org/stemcells](http://www.atcc.org/stemcells)

© 2015 American Type Culture Collection. The ATCC trademark and trade name, and any other trademarks listed in this publication are trademarks owned by the American Type Culture Collection unless indicated otherwise. ATCC products are intended for laboratory research only. They are not intended for use in humans, animals, or diagnostics.

Mouse embryo confocal image courtesy of Dr. David Becker

# Cell Stem Cell

Best of 2014

## Review

Evolution of the Cancer Stem Cell Model

*Antonija Kreso and John E. Dick*

## Articles

Leptin-Receptor-Expressing Mesenchymal Stromal Cells Represent the Main Source of Bone Formed by Adult Bone Marrow

*Bo O. Zhou, Rui Yue, Malea M. Murphy, James G. Peyer, and Sean J. Morrison*

Conversion of Danger Signals into Cytokine Signals by Hematopoietic Stem and Progenitor Cells for Regulation of Stress-Induced Hematopoiesis

*Jimmy L. Zhao, Chao Ma, Ryan M. O'Connell, Arnav Mehta, Race DiLoreto, James R. Heath, and David Baltimore*

Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming

*Xiao Hu, Lei Zhang, Shi-Qing Mao, Zheng Li, Jiekai Chen, Run-Rui Zhang, Hai-Ping Wu, Juan Gao, Fan Guo, Wei Liu, Gui-Fang Xu, Hai-Qiang Dai, Yujiang Geno Shi, Xianlong Li, Boqiang Hu, Fuchou Tang, Duanqing Pei, and Guo-Liang Xu*

Targeting Self-Renewal in High-Grade Brain Tumors Leads to Loss of Brain Tumor Stem Cells and Prolonged Survival

*Zhe Zhu, Muhammad Amir Khan, Markus Weiler, Jonas Blaes, Leonie Jestaedt, Madeleine Geibert, Peng Zou, Jan Gronych, Olga Bernhardt, Andrey Korshunov, Verena Bugner, Peter Lichter, Bernhard Radlwimmer, Sabine Heiland, Martin Bendszus, Wolfgang Wick, and Hai-Kun Liu*

In Vivo Direct Reprogramming of Reactive Glial Cells into Functional Neurons after Brain Injury and in an Alzheimer's Disease Model

*Ziyuan Guo, Lei Zhang, Zheng Wu, Yuchen Chen, Fan Wang, and Gong Chen*

The Developmental Potential of iPSCs Is Greatly Influenced by Reprogramming Factor Selection

*Yosef Buganim, Styliani Markoulaki, Niek van Wietmarschen, Heather Hoke, Tao Wu, Kibibi Ganz, Batool Akhtar-Zaidi, Yupeng He, Brian J. Abraham, David Porubsky, Elisabeth Kulenkampff, Dina A. Faddah, Linyu Shi, Qing Gao, Sovan Sarkar, Malkiel Cohen, Johanna Goldmann, Joseph R. Nery, Matthew D. Schultz, Joseph R. Ecker, Andrew Xiao, Richard A. Young, Peter M. Lansdorp, and Rudolf Jaenisch*

Prolonged Fasting Reduces IGF-1/PKA to Promote Hematopoietic-Stem-Cell-Based Regeneration and Reverse Immunosuppression

*Chia-Wei Cheng, Gregor B. Adams, Laura Perin, Min Wei, Xiaoying Zhou, Ben S. Lam, Stefano Da Sacco, Mario Mirisola, David I. Quinn, Tanya B. Dorff, John J. Kopchick, and Valter D. Longo*

## Resource

An iCRISPR Platform for Rapid, Multiplexable, and Inducible Genome Editing in Human Pluripotent Stem Cells

*Federico González, Zengrong Zhu, Zhong-Dong Shi, Katherine Lelli, Nipun Verma, Qing V. Li, and Danwei Huangfu*

**BAKER**

## Environments For Science™

From the moment when  
the pieces fell into place,  
and you saw the big picture,

through late nights reviewing literature,  
designing procedures,  
testing hypotheses —

nothing fell through the cracks,  
because too much was at stake.

The work you did then  
advanced your field,  
and your career.

We are proud to have been there,  
protecting your work, and you,  
with the very best in air containment,  
contamination control, and  
controlled environment technology.

Now, as your work gives rise  
to new discoveries —

# We can't wait to see what you do next.



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

Biological Safety Cabinets • Clean Benches • Fume Hoods • CO<sub>2</sub> Incubators • Hypoxia & Anaerobic Workstations

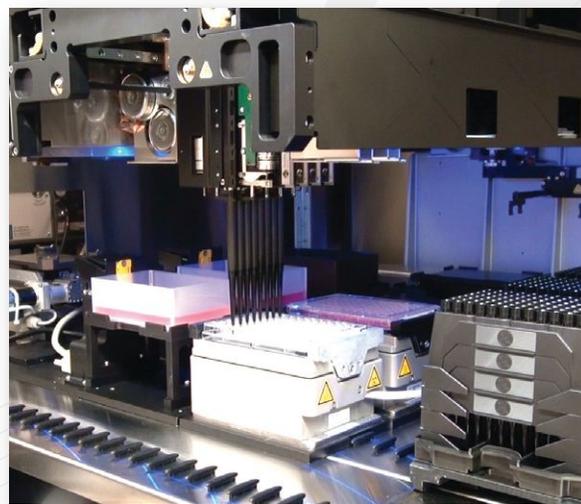
# NYSCF The New York Stem Cell Foundation **Research Institute**

*Catalyst for cures*

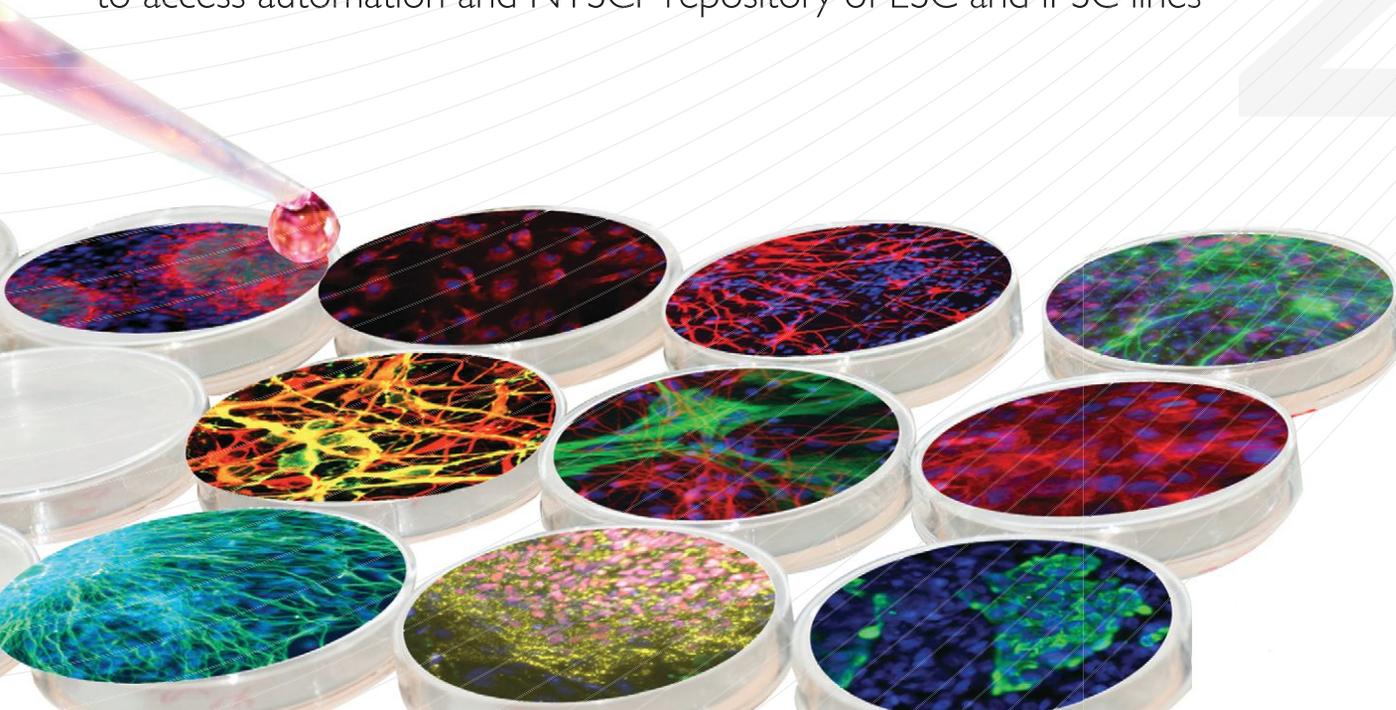
NYSCF has developed a high-throughput robotic platform for iPSC generation, differentiation, manipulation, and characterization

## Partnership opportunities include:

- Production of iPSC lines
- Large-scale parallel differentiation from and genomic engineering of ESC/iPSCs
- Adaptation and optimization of existing differentiation protocols to automation
- Assay development, validation, and scale-up
- Access to NYSCF repository of ESC/iPSC lines (genetically diverse and diseased populations)



Potential users should contact [array@nyscf.org](mailto:array@nyscf.org) to learn about potential opportunities to access automation and NYSCF repository of ESC and iPSC lines



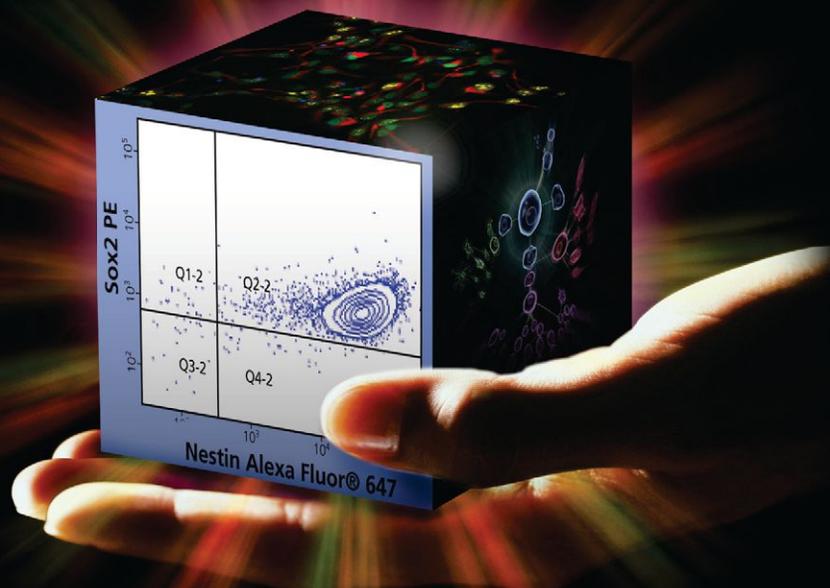


- HUMAN BIOSPECIMENS
- MOLECULAR SERVICES
- STEM AND PRIMARY CELL REAGENTS
- 3D CELL CULTURE SCAFFOLDS
- DIFFERENTIATED iPS CELLS
- CONTRACT SERVICES

**ReproCELL** is an expanding global leader in providing researchers with integrated tools for translational research. ReproCELL, with its **BioServe, Reinnervate and Stemgent** brands offers a comprehensive suite of products and services from human biospecimen procurement to 3D cell culture, through cellular reprogramming and differentiation to readily-available differentiated cells.

LEARN MORE AT ISSCR,  
**BOOTH B08:17!**

# Reveal Stem Cell Diversity



- Evaluate the stemness of cell cultures
- Assess the expression of differentiation markers
- Run functional assays on cell populations
- Isolate your preferred cell types using cell sorting

## Our expert, Dr Nil Emre:

Dr. Emre focused her post-doctoral research on human embryonic stem cell differentiation. She joined BD Biosciences in 2008 and she is currently a manager for the Stem Cell and Cell Function group at BD Biosciences developing antibody-based research reagents and applications for stem cell biology research.

## Meet Dr Emre at:

- **The poster - F-1506 on June 26th at 6pm till 8pm**  
Combination of intracellular flow cytometry with cell surface marker screening identifies a novel cell surface marker signature for human embryonic stem cell derived neurons.
- **The innovation showcase - Room A2 on June 27th at 12pm**  
Flow Cytometric Analysis for Determining the Quality of Stem Cell Cultures
- **The BD Biosciences booth - B05:20 during session and lunch breaks**



# Evolution of the Cancer Stem Cell Model

Antonija Kreso<sup>1</sup> and John E. Dick<sup>1,\*</sup>

<sup>1</sup>Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario M5G 1L7, Canada and Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

\*Correspondence: [jdick@uhnresearch.ca](mailto:jdick@uhnresearch.ca)

<http://dx.doi.org/10.1016/j.stem.2014.02.006>

Genetic analyses have shaped much of our understanding of cancer. However, it is becoming increasingly clear that cancer cells display features of normal tissue organization, where cancer stem cells (CSCs) can drive tumor growth. Although often considered as mutually exclusive models to describe tumor heterogeneity, we propose that the genetic and CSC models of cancer can be harmonized by considering the role of genetic diversity and nongenetic influences in contributing to tumor heterogeneity. We offer an approach to integrating CSCs and cancer genetic data that will guide the field in interpreting past observations and designing future studies.

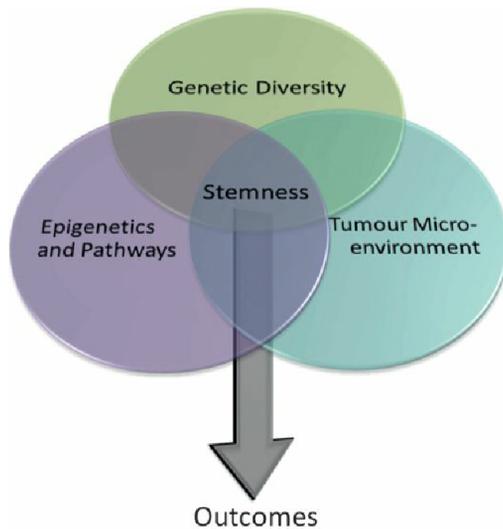
## Introduction

Despite advances in cancer treatment, many patients still fail therapy, resulting in disease progression, recurrence, and reduced overall survival. Historically, much focus has been on the genetic and biochemical mechanisms that cause drug resistance. However, cancer is widely understood to be a heterogeneous disease and there is increasing awareness that intratumoral heterogeneity contributes to therapy failure and disease progression (Hanahan and Weinberg, 2011). A tumor is not simply a “bag” of homogeneous malignant cells. Rather, a tumor is a complex ecosystem containing tumor cells, as well as various infiltrating endothelial, hematopoietic, stromal, and other cell types that can influence the function of the tumor as a whole. These extraneous cell types can influence tumor cells directly and can create metabolic changes such as a hypoxic environment and nutrient fluctuations, which contribute to heterogeneity in the function of malignant cells. By functioning as a complex ecosystem, overall tumor fitness may be enhanced, ultimately impacting therapy failure (Junttila and de Sauvage, 2013). Aside from these non-cell-autonomous effects, even individual malignant cells within a tumor can possess variation in growth, apoptosis, metabolism, and other “hallmarks of cancer.” However, the mechanisms driving intratumoral variation in cellular function have, until recently, been uncertain.

Three avenues of cancer research are coming together to provide increasing clarity to the underlying mechanisms of tumor heterogeneity and uncovering how these are linked to therapy resistance, tumor progression, and recurrence. Advanced genome sequencing has demonstrated that cancer within a single patient is a heterogeneous mixture of genetically distinct subclones that arise through branching evolution (Burrell et al., 2013; Greaves and Maley, 2012). The unique driver mutations within each subclone can impact the cancer hallmarks differently, thereby contributing to functional heterogeneity. In parallel, strong evidence is emerging that nongenetic determinants, largely related to developmental pathways and epigenetic modifications (DNA methylation, histone modification, chromatin openness, microRNA [miRNA], and other noncoding RNA) contribute to functional heterogeneity (Dick, 2008; Meacham and Morrison, 2013; Nguyen et al., 2012). These determinants are generally ascribed to the maintenance of normal tissue

stem cell hierarchies. Similarly, nongenetic determinants create hierarchically organized tumor tissues where a subpopulation of self-renewing cancer stem cells (CSCs) sustains the long-term clonal maintenance of the neoplasm. Although considerable controversy remains as to which tumor types are hierarchically organized and how best to define CSCs, this developmental and/or hierarchical model has generated considerable interest because CSCs appear to possess properties that make them clinically relevant. Evidence from both experimental models and clinical studies indicate that CSCs survive many commonly employed cancer therapeutics. Moreover, the properties and transcriptional signatures specific to CSC are highly predictive of overall patient survival pointing to their clinical relevance. Although this area will not be discussed in our review, an additional promising avenue is the recognition that there are many nontumor cell elements associated with tumors, referred to collectively as the tumor microenvironment (TME) (Hanahan and Coussens, 2012). The juxtaposition of a tumor cell with the TME influences the function of that cell, resulting in significant variation in cellular function. The complexity imposed by the TME is amplified due to crosstalk between tumor cells and the TME. The TME plays a role in adaptive drug resistance, as cells of the same genetic make-up can be sensitive or resistant to drugs depending on the context they are in. Recent studies also point to the potential for the TME to initiate stem cell-like programs in cancer cells (Charles et al., 2010; Vermeulen et al., 2010). Collectively, all three mechanisms are strongly linked to therapy failure and tumor recurrence and all are important determinants of tumor fitness (Figure 1).

We will focus our Review on the genetic and developmental mechanisms that generate tumor heterogeneity, and we will emphasize human studies. Although often considered as mutually exclusive models to describe tumor heterogeneity, we propose that the genetic and developmental and/or hierarchical models of cancer can be harmonized. Indeed, recent findings in leukemia and solid tumors indicate that gene-expression signatures specific to CSC and normal stem cells are highly prognostic for outcome across a wide spectrum of patients with diverse driver mutations (Bartholdy et al., 2014; Eppert et al., 2011; Gentles et al., 2010; Merlos-Suárez et al., 2011), suggesting that stemness is a central biological property or process



**Figure 1. Stemness as a Guiding Principle that Governs Therapeutic Response**

Three fields in biology—cancer genetics, epigenetics, and microenvironment—are coming together to provide increasing clarity to the processes that determine stemness and in turn influence clinical outcome. These three factors can influence stemness simultaneously, but they can also act independently over time. Through evolutionary time, different forces can impact a cell's stemness properties and thereby shape tumor progression and therapeutic response.

upon which many driver mutations coalesce. Thus, our central hypothesis is that three facets—genetic diversity, epigenetics, and the TME—contribute to tumor heterogeneity, and the clinical relevance of each is related to the extent to which it impinges on stemness and thereby influences patient survival (Figure 1).

### Defining Stemness

The modern era of stem cell research began in 1961 with the pioneering studies of Till and McCulloch who developed a clonal *in vivo* repopulation assay and used it to show that a single hematopoietic cell had multilineage differentiation potential while still retaining the property of self-renewal. Although multilineage differentiation potential is often considered a stem cell property, studies in the hematopoietic system have clearly identified distinct cells capable of both repopulation and multilineage potential but lacking self-renewal potential. Thus, the cardinal property of a stem cell is self-renewal, whether normal or malignant. Self-renewal is the key biological process where, upon cell division, a stem cell produces one (asymmetric division) or two (symmetric division) daughters that retain the capacity for self-renewal, ensuring that the stem cell population is maintained or expanded for long-term clonal growth. Operationally, the gold standard measure of a stem cell is maintenance of long-term clonal growth in functional repopulation assays, involving either transplantation into serial recipients or *in situ* tracking. Indeed the lack of adherence to this principle has generated much confusion in the CSC field. Many studies employ surrogate *in vitro* assays such as serial replating of tumorspheres or report on serial passage of bulk tumor cells. However, only clonal serial *in vivo* repopulation assays can formally test self-renewal of stem cells. The molecular programs that underlie the stem cell state

are only just emerging as studies are defining critical epigenetic states and the transcription factors and epigenetic modifiers (e.g., Polycomb complexes and miRNA) that are responsible for endowing self-renewal to a cell. The term “stemness” is increasingly being used in the literature to refer collectively to the integrated functioning of molecular programs that govern and maintain the stem cell state. We will adopt the term “stemness” throughout this review to denote this meaning.

### Cancer Stem Cells and Tumor-Initiating Cells

By definition, both CSCs and normal tissue stem cells possess self-renewal capacity; however, self-renewal is typically deregulated in CSCs. For many cancers, CSCs represent a distinct population that can be prospectively isolated from the remainder of the tumor cells and can be shown to have clonal long-term repopulation and self-renewal capacity—the defining features of a CSC (Clarke et al., 2006; Nguyen et al., 2012). However, in some cancer types it has not been possible to distinguish CSCs from non-CSCs because most cells have CSC function. Such tumors seem to be homogeneous or possess a very shallow hierarchy. As well, some evidence is emerging that certain cancer cells exhibit plasticity by reversibly transitioning between a stem and non-stem-cell state (although this is a controversial and intensely debated topic). Thus, even though some tumors may not be organized into a rigid hierarchy, the stemness state contributes a variety of functions that enable cells to survive therapy. A key proposition of our review is that the determinants of stemness are the core contributors that affect therapy failure, regardless of whether these determinants are present within a transitory state or in well-defined CSC populations. Like CSCs, transitory cells also possess clonal tumor-initiation capacity; however, prospective isolation is difficult. Thus, in terms of nomenclature they pose a problem and formally they should not be termed CSCs, a term restricted to cases where self-renewing CSC can be prospectively purified. We and others refer to such cells on the basis of the functional tumor-initiating cell (T-IC) or leukemia-initiating cell (L-IC) assays that identify them. T-IC or L-IC are defined by their ability to: (1) generate a xenograft that is representative of the parent tumor, (2) self-renew as demonstrated by serial passage in a xenograft assay at clonal cell doses, and (3) give rise to daughter cells that may possess proliferative capacity but are unable to establish or maintain the tumor clone upon serial passage (Clarke et al., 2006). The T-IC/L-IC terms can also be applied in situations where a bona-fide CSC exists, but the proper combination of cell surface markers required for their prospective isolation has not been found. For ease of reading, we have adopted the term T-IC/L-IC throughout our review to refer to all cells with clonal long-term tumor initiating function and not just to those where prospective isolation has been possible.

### Historical Perspectives on Tumor Heterogeneity

Heterogeneity in the cellular morphology of tumors was noted by the great experimental pathologists of the 1800s. Aside from cellular morphology and tumor histology (Heppner, 1984), improved technology has uncovered additional features of heterogeneity between tumors, including variation in cell surface markers (Dexter et al., 1978; Pertschuk et al., 1978; Poste et al., 1980; Raz et al., 1980), genetic abnormalities (Mitelman

et al., 1972; Shapiro et al., 1981), growth rates (Danielson et al., 1980; Dexter et al., 1978; Gray and Pierce, 1964), and response to therapy (Barranco et al., 1972; Heppner et al., 1978). Early evidence pointed to the existence of multiple tumor cell subpopulations within single cancers, including melanoma (Gray and Pierce, 1964), sarcoma (Mitelman, 1971; Prehn, 1970), mammary tumors (Dexter et al., 1978; Henderson and Rous, 1962; Heppner et al., 1978), colon cancer (Dexter et al., 1979), and other solid tumors (Klein and Klein, 1956). Along the same lines, when single cells were cloned from a metastatic mouse melanoma cell line and injected into syngeneic hosts, the degree of metastasis varied extensively, indicating that diversity existed within the parental tumor cells enabling only some clones to metastasize (Fidler and Kripke, 1977). Important evidence for diversification of tumor cell characteristics came from studies in malignant glioma (Shapiro et al., 1981), where primary human tumor cells were isolated and mitoses analyzed by karyotyping. The established karyotypic heterogeneity in the primary tumor was used as a marker for clonal subpopulations derived from primary cells through limiting dilution plating. Cloned subpopulations differed with respect to their sensitivity to chemotherapeutics (Yung et al., 1982) and genetic stability (Shapiro et al., 1981). Evidence that functional tumor cell heterogeneity exists *in vivo* came directly from human acute myeloid and lymphoblastic leukemia patients, where *in vivo*  $^3\text{H-TdR}$  radiolabeling showed marked differences in the proliferation kinetics of individual leukemic cells that could be distinguished on the basis of morphology (Clarkson et al., 1970; Gavosto et al., 1967; Killmann et al., 1963). Thus, this era yielded many observations describing variation in functional parameters and established that growth properties of individual cells within a tumor were far from homogeneous.

Of particular importance from this earlier era of cancer research was quantitative evidence from syngeneic mouse tumor grafting experiments showing that the capacity to initiate a new tumor and sustain disease was variable, with not every cell able to function as a T-IC (Bruce and Van Der Gaag, 1963; Hewitt, 1958). The same observations were made in studies that were carried out in human patients, where tumors were autotransplanted subcutaneously into the same patient (Southam et al., 1962). These studies not only illustrated that tumor reinitiation was variable, but that even in syngeneic recipients T-IC were rare. Collectively, these clonal studies established that tumors are not a collection of homogeneous cells with equal capacity for proliferation. Instead, analogous to an intricate ecosystem, tumors are complex networks where individual cells display a diverse set of characteristics and function together to support the growth and maintenance of the tumor as a whole.

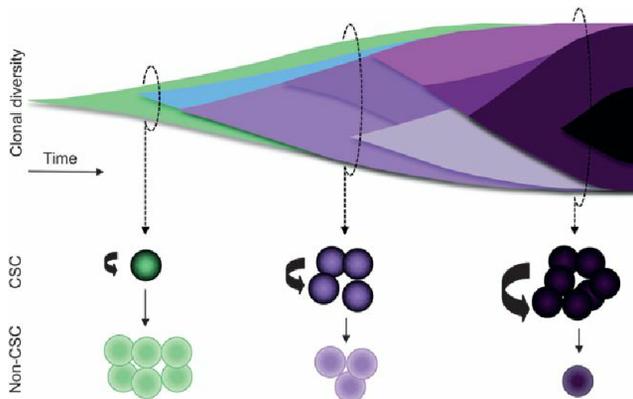
Since the original conception of evolutionary reasoning (Darwin, 1859), it has become evident that genetic diversity within a species' gene pool enhances its ability to survive and adapt to changing environments over time. Likewise, the stability and robustness of ecosystems depends on the degree of biodiversity (Loreau et al., 2001; Tilman et al., 2006). In developmental biology, different specialized cell types need to exist for the effective functioning of organs. For example, for the proper functioning of the blood system, hematopoietic stem cells (HSCs) need to produce a heterogeneous pool of specialized cell types that differ in structure and function. Heterogeneity even within

the HSC pool has been described (Cheung et al., 2013; Yamamoto et al., 2013). Without this diversity, the function of the blood system would be compromised. While the evolution of species and biodiversity in ecology are consequences of changes at the genetic level, diversity in cell function and tissue development within an organism are the result of nongenetic, developmental programs.

### Genetic Mechanisms as the Source of Tumor Heterogeneity

A guiding principle in cancer research is that tumor initiation and progression result from the sequential acquisition of genetic mutations that contribute to subsequent clonal expansions (Nowell, 1976). This view is strongly supported by early studies where genetic mutations were analyzed across different stages of colorectal cancer (Vogelstein et al., 1988). These studies established that genetic changes *cause* phenotypic manifestations, a finding that added significant weight to the idea that cancer development follows the rules of Darwinian evolution (Cairns, 1975; Nowell, 1976). The basic premise of this long-standing idea is that a cell that is endowed with an advantageous heritable mutation generates progeny that has a survival advantage over other cells that lack this mutation. Consequently, the progeny of the cell with increased fitness will flourish and produce a clonal population that dominates the site where it originated. Over time, additional advantageous mutations can arise, endowing a further growth advantage to another cell within the clone. As unique subclones arise, different outcomes are possible: less fit subclones can be completely lost with the most fit subclone dominating, or many minor subclones can persist alongside the dominant clone, forming reservoirs from which evolution can continue. Overall, models where subclones persist and/or contribute to independent phylogenetic lineage trees within single tumors are highly reminiscent of the branching evolution that Darwin described as leading to increased fitness and overall robustness of a species (Figure 2).

Technological advances have made high-throughput sequencing of tumor genomes possible. The last 6 years have seen a flood of whole exome sequencing (WES) and whole genome sequencing (WGS) of thousands of tumors, enabling complex analyses of the mutations that are present within a single tumor and across multiple tumors (Garraway and Lander, 2013). Several principles are emerging from this work, including that fact that the mutational burden is highly variable across tumor types (Lawrence et al., 2013). For example, leukemias tend to have the lowest number of mutations per tumor compared to adult solid tumors. Even within the same tumor type, there is considerable variation in driver mutations and the same driver mutations can occur in different tumor types, suggesting that the same pathways can be active in different tumors (Alexandrov et al., 2013; Kandoth et al., 2013). The high intertumor and intratumor heterogeneity makes it difficult to establish without functional testing whether a particular somatic polymorphism is a driver mutation or a passenger variant. WGS shows that tumors contain thousands of variants, making resolution of the passenger and driver issue a substantial challenge. The origin of passenger mutations was recently elegantly documented in acute myeloid leukemia (AML) genetic studies. Of the many hundreds of mutations that are found in AML blasts,



**Figure 2. Unified Model of Clonal Evolution and Cancer Stem Cells**

Top panel shows that acquisition of favorable mutations can result in clonal expansion of the founder cell. In parallel, another cell may gain a different mutation that allows it to form a new subclone. Over time, genetic mutations accumulate and subclones evolve in parallel. Bottom panel shows that it may be that CSCs are not static entities but can evolve over the lifetime of a cancer as genetic changes can influence CSC frequency. Some subclones may contain a steep developmental hierarchy (left), where only few self-renewing CSCs exist among a large number of non-CSCs. Other subclones (middle) may contain an intermediate hierarchy, where the number of CSCs is relatively high but a hierarchy still exists. Some subclones may have the genetic alterations that confer high self-renewal potential, where most cells are tumorigenic. In this scenario, applying the CSC concept to such homogeneous subclones is not warranted because most cells can self-renew and few non-CSC progeny are generated.

many were actually present in the founder HSC that was the cell of origin for the AML subclone; these mutations arose every time the normal HSC divided and remained functionally neutral and thus can be considered passengers (Jan et al., 2012; Welch et al., 2012). The acquisition of the oncogenic driver within one such HSC “trapped” the preexisting spectrum of mutations within the AML subclone that expanded and progressed from this initiating cell. After development of leukemia, very few additional mutations are needed to drive the last population expansion, although hits continue to be acquired as the disease progresses (Jan et al., 2012; Welch et al., 2012). Thus, the driver and cooperating mutations that the AML subclone acquires during leukemic progression need to be filtered out from the large spectrum of passengers that preexisted and will continue to arise as AML cells proliferate. Similar approaches have been taken in breast cancer and other solid tumors to distinguish drivers from passengers (Nik-Zainal et al., 2012; Stephens et al., 2012). While these approaches clearly document how the mutational landscape of individual patients can influence the heterogeneous properties between patients, more sophisticated approaches are needed to determine how genetic mechanisms contribute to heterogeneity within tumors.

### Intratumoral Genetic Diversity

A key proposition of the multistep tumorigenesis model put forth by Cairns and Nowell is that there are sequential sweeps of clonal dominance that are variably detected depending on when a tumor is sampled. However, they also conceived that the tumor might contain multiple branches or subclones that are evolving independently (Figure 2). Indeed, with deeper sequencing and improved bioinformatic methods, it is becoming

clear that tumors are often composed of a dominant genetic clone plus one or more genetically distinct subclones. For example, topological sampling of tumors has shown that different regions possess distinct mutations that are reflective of genetic subclones seeding different parts of a single tumor (Gerlinger et al., 2012). In the case of metastatic renal cancer, 70% of somatic variants were not found in all biopsies of the same tumor; only a *VHL* mutation and loss of a region on 3p were ubiquitous. Even gene-expression signatures of good and poor prognosis were detected in different regions of the same tumor. Intratumoral diversity with respect to metastatic progression was shown in pancreatic cancer. By sequencing the genomes of metastases and different regions of matched primary tumors obtained through rapid autopsies from seven individuals with end stage pancreatic cancer, the primary tumor was found to harbor geographically and genetically distinct subclones that gave rise to lung, liver, or peritoneal metastases within the same patient (Yachida et al., 2010). Importantly, despite the presence of founder mutations within the parental clones, the cells giving rise to metastatic lesions had a large number of additional mutations, indicating that further clonal evolution had taken place during metastasis. Others have also reported genetic heterogeneity between metastasis-initiating cells in pancreatic cancer (Campbell et al., 2010). These studies highlight the complexity in predicting which subclones will progress to metastasis, even after the genomic architecture of the primary tumor is established.

With the ability to detect genetic subclones within tumors, it is now possible to create lineage maps that provide insight into the subclonal evolution. Such advancements have made it possible to reconstruct the life histories of breast cancers (Nik-Zainal et al., 2012). By sequencing 20 breast cancers to an average 30–40× coverage and one cancer to 188× depth and applying a new bioinformatics algorithm (Greenman et al., 2012) to reconstruct the genomic history, the authors showed that breast cancer evolves through acquisition of driver mutations that produce clonal expansions. Interestingly, the driver mutations occur infrequently in long-lived lineages that passively accumulate mutations without expansions. The most recent common ancestor appeared surprisingly early, indicating that much of the time is spent driving subclonal diversification and evolution among the nascent cancer cells (Nik-Zainal et al., 2012). These studies do not just give a snapshot of the tumor, but narrate the steps it has taken before it was diagnosed, providing promising avenues for earlier screening. One caveat of these studies is that interpatient tumor genetic variability is likely extensive and it may thus be challenging to delineate a common set of steps that are characteristic of different breast cancer subtypes. The key question that the discovery of subclonal diversity raises is which clones will survive therapy and progress to cause recurrence and/or metastasis.

Studies in leukemia have been particularly instructive in revealing the presence of subclones and their role in tumor progression. Analysis of chromosomal translocation breakpoints and DNA copy-number alteration (CNA) profiling in twins with *ETV6-RUNX1* positive acute lymphoblastic leukemia (ALL) showed that a preleukemic clone is initiated in utero that expands, seeds both twins, but then evolves with different kinetics and CNA acquisition in each twin (Bateman et al., 2010; Hong

et al., 2008; Li et al., 2003; Zuna et al., 2004). Genome-wide CNA profiling of paired diagnostic and relapse samples of ALL has been informative (Inaba et al., 2013; Mullighan et al., 2008). In approximately 40% of cases, the leukemic subclone present at relapse was identical to the subclone present at diagnosis or it was a direct evolutionary product. However in 50% of cases, the relapse subclone shared only limited genetic identity with the diagnostic subclone and did not evolve from it. Similar findings came from WGS studies of paired diagnosis and relapse AML samples: the major population at relapse shared only limited genetic identity with the major population at diagnosis and did not evolve from it. These findings from B-ALL and AML predicted the presence of genetically distinct subclones at diagnosis and the existence of ancestral, prediagnostic subclones. These results indicate that tumor evolution may occur through a more complex branching model that gives rise to genetically distinct subclones at diagnosis that vary in aggressiveness and response to therapy (Greaves, 2009, 2010). Moreover, these data establish that there is subclonal variation in both the response to therapy and the probability that a subclone will survive and regenerate a new tumor.

However, these are still *in silico* depictions of intratumoral diversity and they are inferred from bulk tumor tissue. This limitation makes it difficult to determine how and when population expansions occurred to generate subclones and there are questions of whether the sequencing was deep enough to reveal the entire population substructure. Like population studies of human evolution or in ecology, lineage trees that describe evolutionary history are best undertaken with single cells. If large numbers of single cells are analyzed, their relationship to one another can be mapped and phylogenetic lineage trees can be created (Melchor et al., 2014; Navin et al., 2011; Potter et al., 2013; Shlush et al., 2012). Early studies of this type have now been reported for some leukemias and lymphomas, and they reveal a high degree of complexity within single tumors. Shlush et al. tracked polymorphic somatic mutations in large numbers of single leukemia cells taken at diagnosis and relapse and reconstructed cell lineage trees based on their divisional history. The reconstructed lineage trees from cells at relapse were shallow (indicating that they divide rarely) compared to cells at diagnosis, which showed many more subpopulations. Interestingly, relapse cells were closely related to the L-IC enriched subpopulation from the diagnostic sample, which is known to be relatively quiescent. This result implies that in these instances, relapse might have originated from rarely dividing L-IC. Given the importance of L-ICs to tumor growth, it will be important to broaden this type of single cell analysis to more samples and to determine the extent to which L-IC are involved in establishing a genetically diverse relapse.

Collectively, the identification of genetically diverse subclones within single tumors provides strong evidence that intratumoral heterogeneity can be driven by the unique mutation spectrum present within each subclone (Figure 1). However, key questions remain that sequencing studies alone cannot resolve. One major challenge is determining which mutations are able to drive tumor growth and how to link these drivers to the clonal propagation potential of subclones. Does genetic diversity exist in tumor cells that are responsible for long-term tumor propagation? Which subclones will evolve further? Will all cells within a subclone be

equally sensitive to therapy? Which clones will recur or metastasize? Studies to answer these questions require functional assays.

### **Nongenetic Mechanisms as the Source of Heterogeneity—the Cancer Stem Cell Model**

Although the idea that cancer retains features of embryological development has a long history (Cohnheim, 1875), the modern idea that developmental programs underlying normal tissue organization may still function to some extent in cancer began with seminal studies of teratocarcinoma (Pierce et al., 1960), small cell lung carcinoma (Baylin et al., 1978), and mammary carcinoma (Bennett et al., 1978; Hager et al., 1981). They suggested that many tumor cells were differentiated and that these “differentiated” cells were generated by tumor “stem” cells, similar to normal tissue stem cells producing normally differentiated tissues. Thus, tumors can be considered as caricatures of embryogenesis or normal tissue renewal (Pierce and Cox, 1978; Pierce and Speers, 1988). Early studies in the hematopoietic system were also instructive. There was clear evidence from cytokinetic labeling studies that the majority of leukemia blasts were postmitotic and needed to be replenished from a small population of highly proliferative cells (Clarkson et al., 1967; Clarkson et al., 1965; Clarkson, 1969). Presciently, these studies also predicted the existence of a rare leukemic population that cycled very slowly and showed resistance to antiproliferative therapies and therefore was thought to be the source of recurrence. Since similar cytokinetics were observed for normal hematopoietic stem cells, it was proposed that the slow-cycling leukemia cells were responsible for the continued generation of the proliferative fraction, representing a leukemic “stem cell” population (Clarkson, 1974). These early studies, together with efforts to identify clonogenic AML progenitors (Buick et al., 1977; McCulloch, 1983; Metcalf et al., 1969; Moore et al., 1973), sparked an interest in thinking about leukemia in terms of hierarchical organization, as was being established for normal hematopoiesis at that time.

In order to demonstrate that a tumor is organized in a hierarchical manner, it is crucial to establish that it consists of functionally distinct cell types that can be prospectively purified and assayed. With the development of fluorescence-activated cell sorting techniques (Bonner et al., 1972), coupled with refinements in xenografting techniques in immune-deficient mice, it was possible to engraft normal human hematopoietic cells (Kamel-Reid and Dick, 1988; Lapidot et al., 1992) and leukemic cells in mice (Dick et al., 1991; Kamel-Reid et al., 1991; Kamel-Reid et al., 1989). These tools, along with quantitative assays, set the stage for the first purification of T-IC, the operational term for human CSCs (Clarke et al., 2006).

Flow sorting using cell surface markers CD34 and CD38 was used to prospectively isolate human T-IC in AML, termed leukemia-initiating cells (L-IC) (Lapidot et al., 1994). The leukemia initiation potential was in the CD34<sup>+</sup>CD38<sup>-</sup> fraction and no engraftment was detected from the CD34<sup>+</sup>CD38<sup>+</sup> or CD34<sup>-</sup> fractions. By injecting different numbers of cells per mouse and establishing a linear correlation with engraftment, it was calculated that 1 in 2.5 × 10<sup>5</sup> cells could initiate a leukemic graft (Lapidot et al., 1994). Analysis of additional AML samples in a more sensitive immune-deficient mouse model (using non-obese diabetic/severe combined immunodeficiency [NOD/SCID] mice)

followed this initial study (Bonnet and Dick, 1997), further establishing that AML is organized as a hierarchy with CD34<sup>+</sup>CD38<sup>-</sup> L-IC at the apex. These studies provided proof for the hypotheses from the 1960s and established that not every AML cell was equal and only rare cells were L-IC.

The initial studies in AML laid the foundation for the generation of CSC studies in solid tumors that followed. The first identification of CSCs in a solid tumor was achieved over ten years ago in human breast cancer (Al-Hajj et al., 2003). A subset of breast cancer cells (CD44<sup>+</sup>CD24<sup>-</sup>) was prospectively isolated and shown to be solely responsible for sustaining the disease in immune-deficient mice. The CSC subset could be serially passaged and the xenografts generated were histologically heterogeneous, resembling the parent tumor from which they were derived. These results demonstrated that the same CSC principles that had previously been shown to apply in an AML model could also be translated to a solid tumor. Since the initial publication in breast cancer, a plethora of papers have been published identifying CSCs in numerous cancers including brain (Singh et al., 2004), head and neck (Prince et al., 2007), pancreas (Hermann et al., 2007; Li et al., 2007), lung (Eramo et al., 2008), prostate (Collins et al., 2005; Patrawala et al., 2006), colon (O'Brien et al., 2007; Ricci-Vitiani et al., 2007), and sarcoma (Wu et al., 2007). In all cases, non-T-ICs were generated in the xenografts providing evidence for hierarchical organization. However, in most cases where patient-derived cancer samples were used, no genetic analysis was undertaken to compare the xenografts to the primary tumor to determine whether there was selective outgrowth of one or more subclones. Nonetheless, CSCs appeared to be a common feature across different cancer subtypes and tumors from different tissues. Collectively, the studies illustrated that the ability to initiate and propagate tumor growth varies between different cells within a cancer and that this variation is due to a hierarchical relationship between tumorigenic and nontumorigenic cells. This relationship is comparable to developmental hierarchies seen in normal tissues where stem cells reside at the apex and are responsible for generating progeny that in turn exhibit increasing commitment and lineage restriction.

### Xenografting and CSC Detection

Because tumor initiation is one of the defining features of CSCs, xenografting is central to the CSC model. A limitation of xenograft studies is that even orthotopic transplantation may not faithfully reproduce the TME or the growth factor milieu found within a patient's tumor. Some murine growth factors are not cross-species reactive (e.g., TNF) (Bossen et al., 2006; Rongvaux et al., 2013). These environmental differences can impart selective forces on tumor cells. As a result, some cells that would possess T-IC activity in humans might not display growth as xenografts. Moreover, the experimental techniques necessary to obtain single cells to test for T-IC activity are harsh. Digestion of a solid tumor into single cells causes a loss of stromal components and cellular architecture. Cells are under atmospheric oxygen levels and are subjected to abrupt changes in nutrients and pH. Furthermore, cells are stained with antibodies to cell-surface molecules and passed through a sorting machine to separate putative CSCs from non-CSCs. After hours of preparation under conditions that are drastically different

from the native environment of the tumor, cells are then finally injected back into a xenogeneic environment and assayed for growth potential. Given these harsh experimental procedures, testing for the presence of CSCs effectively tests for the most robust cell that can grow. In addition, key aspects of the TME are altered in the transplantation process. All of these changes may affect a cancer cell's growth properties in the xenograft assay.

Over the past two decades, there have been steady improvements to the xenograft assay, including development of more immune-deficient recipient mice, better methods for transplantation, and humanizing recipients with human TME and/or growth factors (Rongvaux et al., 2013). Accordingly, some aspects of the initial CSC model have needed to be refined. For example, L-ICs were thought to reside solely in the CD34<sup>+</sup>CD38<sup>-</sup> fraction of AML (Bonnet and Dick, 1997; Lapidot et al., 1994). However, by using more immune-deficient NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG) mice, L-IC can now be detected in other fractions (Taussig et al., 2008). Careful examination of larger sample numbers using NSG mice together with intrafemoral transplantation to improve the sensitivity for L-IC detection has confirmed the essential conclusion of the original studies: in virtually all samples, L-IC reside in the CD34<sup>+</sup>CD38<sup>-</sup> fraction. Still, in at least half of the samples (>100 samples tested, data not shown), L-ICs are also found in at least one other fraction (usually CD34<sup>+</sup>CD38<sup>+</sup>). In addition to analyzing the phenotypic fractions that contain L-ICs, the size of the population and the L-IC frequency therein need to be taken into account. Nevertheless, in the vast majority of cases evidence of a hierarchy is still seen since fractions devoid of L-IC are found (Eppert et al., 2011). In those samples where cell surface marker analysis using CD34 and CD38 cannot identify a cell fraction that is devoid of L-IC activity, we have found that sorting cells on the basis of miRNA expression levels can be used to prospectively separate L-IC and non-L-IC fractions (Gentner et al., 2010; Lechman et al., 2012). Thus, even in those samples we cannot conclude a priori that the tumor does not follow the CSC model. Rather, we suggest that the cell surface markers are uninformative and not valid to make a determination of hierarchical organization in those cases. The use of sorting based on differential expression of miRNA represents a potentially powerful method that has already shown utility for sorting human HSC. This method needs to be explored further as a means to fractionate cells when cell surface markers are not available (as for many solid tumor types) or where they are uninformative (Amendola et al., 2013). In addition to miRNA, the use of reporter assays that measure cellular signaling pathway activity, such as the Wnt reporter (Vermeulen et al., 2010), can be used as alternative means of measuring distinct cellular fractions that may segregate T-IC and non-T-IC. The tools for analyses of other important pathways in human cells is only beginning to emerge and it will be interesting to see how intracellular signaling markers will impact the identification of T-ICs. At least in AML, the strongest independent piece of evidence supporting the utility of the xenograft assay to detect bona fide L-IC is that only gene signatures from functionally validated L-IC populations are strongly prognostic for patient survival (Eppert et al., 2011). This and other evidence is presented in the Linking Stemness, Prognosis, and Therapy section.

In the context of solid tumors, there have been many more discordant findings regarding the phenotype and properties of T-IC depending on experimental conditions and the type of xenograft assays employed. Through the use of recipients with increased immune-deficiency, the assayed frequency of T-IC in melanoma changed by many orders of magnitude such that the most permissive recipient read out a T-IC frequency of virtually 1 in 1 (Quintana et al., 2010). As most tumor cells were T-IC, such tumors appear to be homogeneous and not following a hierarchical model of tumor organization, although some have argued that there may be methodological explanations (Boiko et al., 2010). Nevertheless, use of more immune-deficient mouse models does not necessarily change the evidence documenting the existence of CSC and for many tumors a hierarchy containing T-IC and non-T-IC was seen (Ishizawa et al., 2010; O'Brien et al., 2012). In others cases such as breast cancer, the implantation of human stromal elements appears to mitigate many problems and enables reliable detection of T-IC (Kuperwasser et al., 2004). Overall, future use of recipients that express components of the human immune system, as well as cross-species reactive growth factors, should be valuable modifications to the xenograft assay and enable more reliable evaluation of hierarchical organization in tumors.

Given the central importance of the xenograft assay to measure functional cancer cell properties and its heavy use in CSC research and experimental drug studies, it is important to fully describe the characteristics, including the genetic make-up, of cancers that grow as xenografts. A major limitation of CSC studies to date is that there has been a lack of integration of genomic and functional properties of T-ICs as we describe later in this Review. Notably, it remains to be determined which genetic clones can generate grafts in mice and how this influences the corresponding CSC measurements.

#### Clonal Dynamics, Dormancy, and Therapy Failure

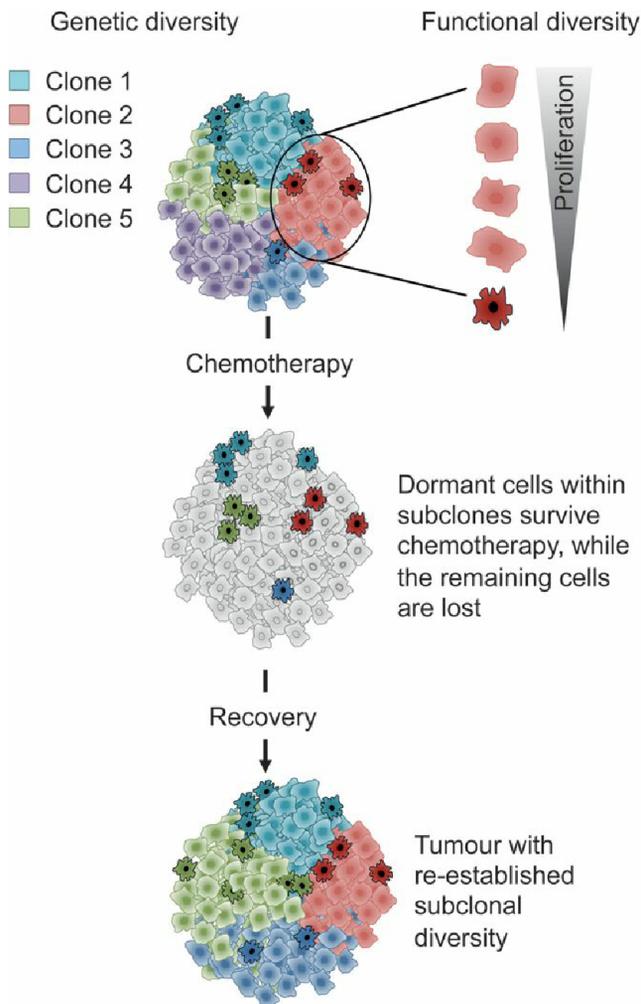
Although the studies described above indicate that not all tumor cells possess T-IC function, another major question is whether all T-IC are equal in their tumor propagation ability. If there is variation, it will be important to establish whether the variation exists within cells of a single genetic clone. The answer to this question holds major importance to the design of future cancer therapeutics. Addressing this question requires genetic analyses combined with functional assays that measure tumor propagation at the resolution of individual clones. We have recently characterized colon cancer xenografts and shown that single genetic subclones from the patient tumor can be separated and stably propagated over multiple passages in mice. Being able to propagate a genetic clone allowed us to track the behavior of cells within unigenetic lineages. By using lentivirus-mediated cell marking, we mapped the growth dynamics of 150 marked cells from ten primary human colon cancer samples in serial transplants that spanned 387 days of tumor growth on average (Kreso et al., 2013). In every genetic clone that we analyzed, we detected significant variation of cellular behaviors: some marked cells were proliferative and persisted at every transplant, whereas others were less robust and could not be detected at later points during transplantation. Thus, these results directly identify functional diversity among cells that are part of a single genetic clone in a solid tumor.

In addition, approximately 20% of marked cells were initially undetectable, but following serial transplantation such cells became activated and continued to function (Kreso et al., 2013). This provided formal evidence for the existence of dormant cell populations that drive tumor growth in primary human colon cancer. Moreover, by treating xenografts with conventional chemotherapy, we discovered that while some long-term persisting cells were eradicated, the dormant cells survived treatment and contributed to tumor regrowth. These changes were not accompanied by selection of distinct genetic subclones, as the control and treated tumors displayed close genetic identity. Tumor cell dormancy has been observed in other systems, including breast cancer, melanoma, and leukemia (Pece et al., 2010; Roesch et al., 2010; Saito et al., 2010). Because most conventional chemotherapies are largely cytotoxic to dividing cells, dormancy may provide cells with a means of escape or survival (Figure 3), although other mechanisms, such as acquisition of new mutations or selection of cells with preexisting genetic mutations, could also be at work to ensure survival following therapy. Collectively, these studies provide evidence that even within a single genetic clone, cancer cells are heterogeneous in their ability to survive chemotherapeutic insults. This added layer of functional diversity adds a new tier of complexity within tumors.

#### Plasticity and CSC Detection

In vitro studies have often been used as surrogate means of studying T-IC. A number of reports using cell lines that have been cultured in vitro have shown that the T-IC state is not static. Sorted T-IC enriched populations generated non-T-IC, but some studies found that sorted non-T-IC populations could generate T-IC (Gupta et al., 2011; Magee et al., 2012; Sharma et al., 2010). Studies of JARID1B, a histone demethylase, have been informative and highlight the complexity of the T-IC state in melanomas. JARID1B was shown to mark slowly cycling melanoma cells that are essential for continuous tumor growth of established melanomas and metastatic progression, but are not required for tumor initiation (Roesch et al., 2010). JARID1B expression was limited to a small subpopulation of melanoma cells, but the maintenance of this subpopulation was dynamic: while purified JARID1B-positive cells generated JARID1B-negative cells, as expected by the CSC model, single JARID1B-negative cells also gave rise to heterogeneous progeny, including JARID1B-positive cells. This study indicates that some cells that are essential for tumor maintenance may not be static entities, but rather can acquire tumor maintenance capabilities depending on the context.

The finding that normal stem cells can reenter the stem cell state (Mani et al., 2008) gave way to the idea that it may be possible to generate T-IC from non-T-IC under some conditions. Indeed, EMT factors have been used to generate T-IC from non-T-IC in breast cancer (Chaffer et al., 2013). The environment in which tumor cells reside can also induce stem-like states in cancer cells. For example, myofibroblast-secreted factors, including hepatocyte growth factor, can induce Wnt signaling in colon cancer cells and consequently induce a T-IC-like state in more differentiated tumor cells in vivo (Vermeulen et al., 2010). In mouse models of intestinal tumor initiation, epithelial nonstem cells can reexpress stem cell markers upon Wnt activation and



**Figure 3. Functional Diversity between Cells within Subclones Impacts Response to Therapy**

Each clone (depicted by the different colors) contains a mixture of cells that vary with respect to their stemness and/or proliferative ability, including relatively dormant cells. Together these factors represent the functional diversity present within single genetic subclones. Chemotherapy can reduce tumor burden by eliminating the highly proliferative cells within subclones, while sparing the relatively dormant cells; following therapy, these cells can seed a new cancer. Thereby, subclonal diversity can be altered with chemotherapy and can allow for the selection of cells with additional genetic mutations that confer a survival advantage. Not depicted in the diagram is the concept that chemotherapy-resistant cells can exist before treatment and can be selected following chemotherapy. Thus, chemotherapy can introduce new mutations to confer treatment resistance, but it can also select preexisting cells that accumulated mutations, which confer chemotherapy resistance during the long evolution of the tumor before it was diagnosed.

can “dedifferentiate” to T-ICs (Schwitalla et al., 2013). Likewise, perivascular nitric oxide that is released by endothelial cells can activate Notch signaling and induce a stem-like state in PDGF-induced gliomas (Charles et al., 2010). These studies highlight the dynamic nature of cancer cells and show the importance of the stem cell state in tumor generation.

Given the importance of these concepts, it will be important to show whether other cancers possess such properties. As well, it will be critical to determine to what extent plasticity exists in primary tumor tissue, as opposed to cell lines, and whether it is

induced in vivo. Although provocative, some studies reporting plasticity were not done clonally and this is essential to understand the homogeneity of cells in each state and the frequency of cells that are able to change states. Is every non-T-IC able to generate a new T-IC, or are only some non-T-ICs responsible for the generation of new T-IC? If only some, does this reflect heterogeneity of the non-T-IC population? Clearly tumors with a high probability of interconversion between T-IC and non-T-IC states render hierarchical cellular organization less meaningful than if such interconversions are rare. Normal tissue stem cells can also “dedifferentiate” into a more primitive state when normal tissue homeostasis is perturbed, for example during transplantation procedures or following stem cell ablation (Rinkevich et al., 2011; Tata et al., 2013; Van Keymeulen et al., 2011). Thus, it will be important to determine the probability of being in one state versus another and the factors that influence such interconversions (Gupta et al., 2009). However, even in tumors where the interconversion rate is high, the available data indicates that when a cancer cell possesses stemness properties it is more likely to progress, metastasize, resist therapy, and self-renew, compared to when it is in the opposite state. Thus, even for tumors that do not strictly follow the CSC model, the concept that stemness is an important aspect of the biology of that cell remains strong. As such, novel approaches will be needed to eradicate cells that display determinants of stemness.

### Epigenetics and Stemness

The primary, nonmutational mechanism that governs developmental hierarchies is epigenetic regulation of the genome. Epigenetic modifications of DNA, histones, and nucleosomes as well as noncoding RNAs, including miRNA, allow for modification of gene expression (Baylin and Jones, 2011; Iorio and Croce, 2012). Alterations in the epigenome dictate cell fate specification and have been used as means of reprogramming noncancerous cells. Although epigenetic modifications are not as stable as mutational changes and can be reversed, some types of modification are a stable, heritable means by which distinct cellular states and functions can be generated. The importance of epigenetic regulation in generating diversity apart from genetic mutation has been shown in several systems. For example, a small proportion of slowly cycling melanoma cells that are essential for tumor growth can be purified based on the expression of JARID1B, a member of the jumonji/ARID1 histone 3 K4 demethylases (Roesch et al., 2010). Other epigenetic factors including members of the Polycomb group of transcriptional repressors (BMI-1 and EZH2) that are linked to normal stem cell self-renewal have been shown to exhibit variation in expression levels within tumors and play a role in tumor progression (Sparmann and van Lohuizen, 2006). Further support for the role of stemness in cancer biology is emerging from cancer genome-sequencing efforts showing that genetic disruption of epigenetic regulators of normal stem cell function is critical for cancer pathogenesis. Mutation in *DNMT3A*, which is highly recurrent in AML, causes major dysregulation of gene expression leading to upregulation of stemness genes and increased repopulation and self-renewal of normal HSC (Ley et al., 2010; Shah and Licht, 2011). Other highly recurrent mutations in genes such as *IDH1/IDH2* and *TET2* affect epigenetic programs that underlie stemness for many cancers, including AML (Abdel-Wahab and Levine,

2010). Thus, epigenetic factors, classically ascribed to govern normal cell diversification, are becoming increasingly relevant for the maintenance of different cancer cell states.

Epigenetic mechanisms can also be important for the observed variability in response to therapy (Glasspool et al., 2006). A small population of cells that remain drug-tolerant following treatment has been reported across several cell lines, including cells derived from melanoma, lung, gastric, colon, and breast cancers (Sharma et al., 2010). Following treatment of these drug-sensitive cell lines with anticancer agents, the authors observed a small proportion of cells that persist, remaining viable while the majority of cells are killed by the therapy. This drug-tolerant phenotype was related to changes at the level of global chromatin, with high expression of the histone demethylase JARID1A and IGF-1R signaling in drug tolerant cells (Sharma et al., 2010). Importantly, heterogeneity in drug response can be generated even when cultures are initiated from single cancer cells, indicating a nongenetic mechanism. Others have found that escape of cells from anticancer drug treatment involves a survival advantage conferred by cell-to-cell variability in the dynamics of specific proteins (Cohen et al., 2008). Substantial variation between daughter cells in response to antimetabolic drugs has also been reported that is not the result of genetic differences, but rather due to competing intracellular networks involving caspase activation and cyclin B1 levels (Gascoigne and Taylor, 2008). Overall, these studies highlight the importance of nongenetic mechanisms governing both cellular fates and drug response. It will be important to discern how these *in vitro* studies translate to *in vivo* growth properties of cancer cells following drug administration.

Gene-expression analysis is another important means by which different cellular states can be identified. By using single cell multiplex PCR analysis in combination with fluorescence-activated cell sorting, it has been shown that colon tumors contain subpopulations of cells whose transcriptional states mirror those of the lineages found in the normal colon epithelium (Dalerba et al., 2011). Importantly, these authors show that a tumor derived from a single cell can exhibit the morphological diversity and transcriptional variability reflective of multilineage differentiation seen in normal colon tissue. These gene-expression programs are also important indicators of patient survival (Dalerba et al., 2011), formally proving that epigenetic heterogeneity due to multilineage differentiation processes can establish phenotypic and functional diversity in tumor clones. As such, it is evident that tumor heterogeneity can arise due to transcriptional programs that are reminiscent of normal tissue differentiation, which are independent of genetic diversity.

### Noisy Gene Expression and Heterogeneity

Studies in lower organisms have found that stochastic nongenetic processes involving protein production or degradation can account for numerous phenotypic effects (Losick and Desplan, 2008; Süel et al., 2007; Wernet et al., 2006). In mammalian cells, survival of apparently homogeneous cells can be dictated by natural differences in protein levels, which regulate receptor-mediated apoptosis between cells and illustrate the dramatic effects that noise in gene expression can have (Spencer et al., 2009). The variability in levels of proteins within cells, albeit

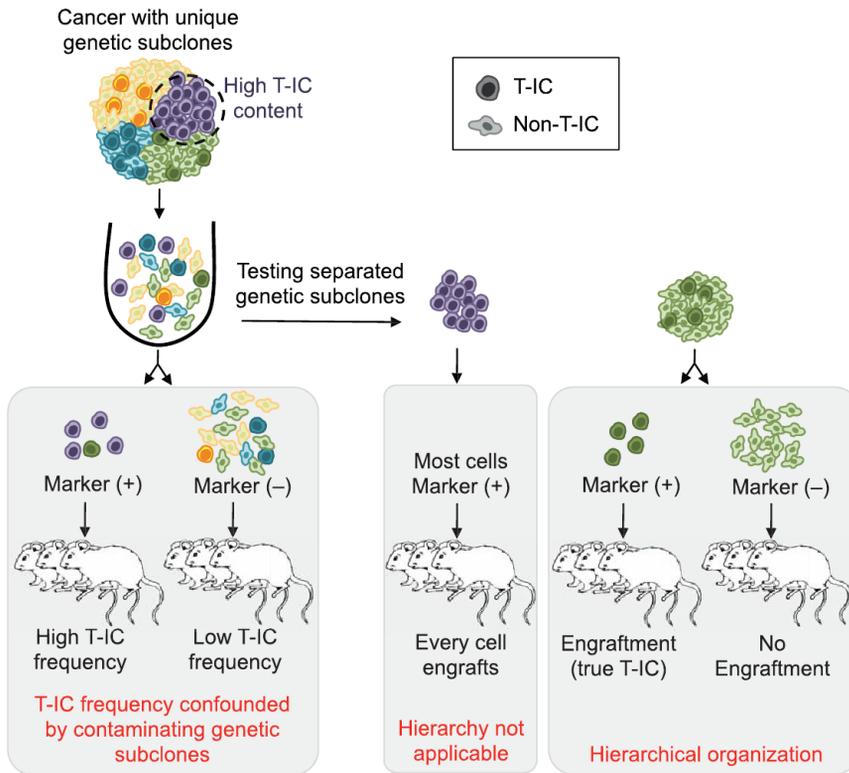
transiently heritable, quickly changes in daughter cells due to different growth rates and noise in gene expression. As such, it is inherently different from epigenetic regulation discussed above. Nonetheless, the variability between cells with respect to noise in gene expression and variability in signal transduction components has implications for tumor biology and therapeutics. Traditionally, the failure of a therapy to eradicate all cells has been ascribed to genetic differences, proliferative status, or the microenvironment, but it is possible that the variability of cells to respond can also be governed by natural differences in protein levels.

At the level of cell populations, evidence is emerging to support stochastic processes governing cell state equilibria. For example, breast cancer cell lines, separated into different phenotypic fractions or “states” based on cell surface marker expression, return to equilibrium proportions over time *in vitro*. This progression toward equilibrium proportions was the result of interconversion between different phenotypic states, which can be modeled as stochastic processes that occur with each cell division using the Markov process, where interconversion rates depend only on the cell’s current state (Gupta et al., 2011). This study provided a theoretical framework for explaining phenotypic equilibria in breast cancer cell lines. It will be important to see these principles established in primary cancers where T-IC can be highly resolved by sorting and tested in robust clonal serial T-IC assays.

Collectively, these studies indicate that in apparently homogeneous environments, cells of the same genotype can exist in different states that influence their behavior. The implications of such variability to biology and medicine are important. The implicit assumption most experimental studies take is that a uniform cell population reacts in a uniform manner. However, given the biological noise between individual cells, the effects of treatments on populations are likely underestimated, as averaging data across many cells can have the net effect of masking heterogeneity at the single cell level. With new technological advances, an increasing number of single-cell studies are being reported that demonstrate considerable cell-to-cell variability in apparently homogeneous populations. For example, quantitative PCR gene-expression analysis of 280 genes was undertaken for 1,500 single cells that span a variety of highly purified mouse HSC and progenitor populations. This study uncovered a large degree of heterogeneity within cell types that were classically thought to represent a uniform collection of cells (Guo et al., 2013). It will be important to establish whether such variation is the result of technical variability in the assays used, or whether it represents true differences in biological function. In the context of the CSC model, extensive cellular variability within what is thought to be a uniform CSC population has implications for therapeutic targeting. Therapy directed against a molecular target might not be effective if intrinsic variability in the cellular context renders subsets of cells within the population unequally responsive to drug targeting.

### Limitations of Genetic and Nongenetic Models

As we have argued throughout this Review, both genetic and nongenetic determinants influence tumor heterogeneity and often these two views have been presented as mutually exclusive models, stimulating intense debate (Clevers, 2011; Marusyk



**Figure 4. Failure to Separate Genetic Subclones May Confound Conclusions Regarding Source of Cancer Heterogeneity**  
Left panel shows that if cancer cells are not separated into distinct genetic subclones before they are tested for the presence of a hierarchical organization, then T-IC measurement may not reflect the complexity of the parental cancer. Right panel shows that cells from the purple clone have a high number of engrafting cells irrespective of a marker that is used to prospectively purify T-ICs. On the contrary, T-ICs can be prospectively purified from the green clone. Of note, the clones depicted in yellow and blue contain T-IC but in this example they are not positive for the marker of choice. This highlights that multiple markers may be required to identify T-ICs from distinct genetic subclones.

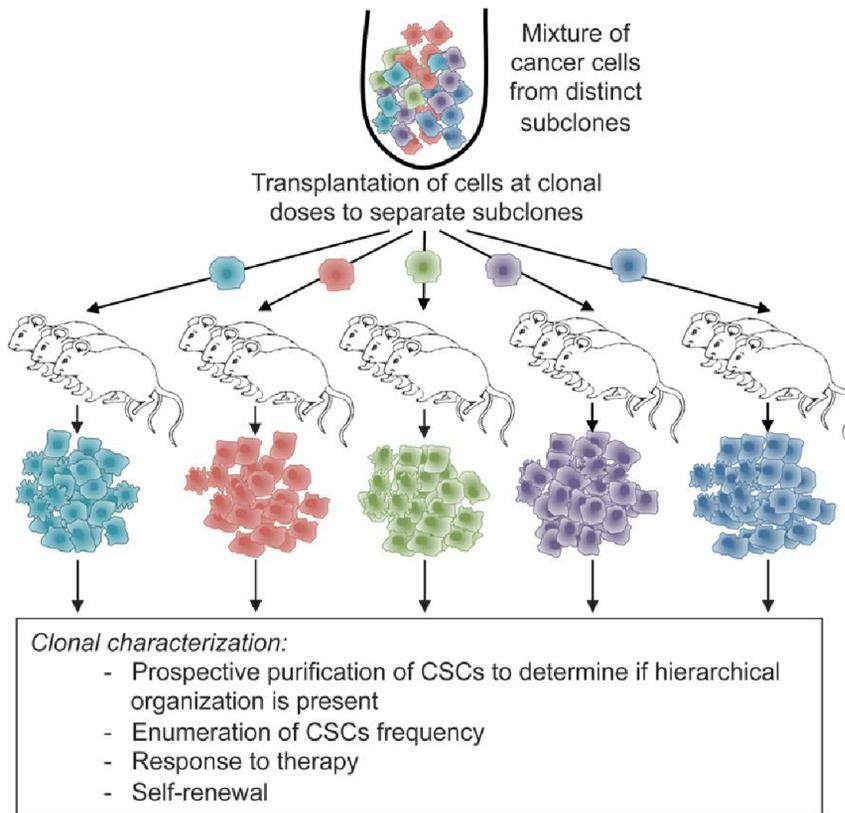
**Unification of the CSC and Clonal Evolution Models**

From a conceptual standpoint it is clear that therapy failure and recurrence are not simply due to the acquisition of new mutations. Rather, the surviving tumor cells must also have regenerative potential in order to regrow the tumor; cells contributing to recurrence must behave like T-IC. Further, tumors are dynamic entities: cells are dying, proliferating, or

et al., 2012; Shackleton et al., 2009). However, each view in isolation is insufficient to explain fully the diversity seen within cancers. The genetic model focuses on genetic heterogeneity without considering that individual cells within a genetically homogeneous subclone might still exhibit variation in function due to any of the nongenetic determinants described above. Similarly, a major limitation of the CSC model or hierarchical model is that it views the tumor as being genetically homogeneous and static, without accounting for the existence of genetically distinct subclones or tumor evolution. For example, a tumor might contain different subclones, some of which are virtually homogeneous in terms of T-IC activity because they are highly progressed and possess a high mutational burden, whereas other subclones with fewer oncogenic mutations might be almost devoid of T-IC. Such subclones could also possess differences in the cell surface markers used for sorting. Thus fractionation of the bulk tumor into T-IC and non-T-IC populations could simply be the result of segregating subclones with very low T-IC activity from those with high T-IC activity (Figure 4). In this scenario, sorting has simply segregated on the basis of genetic identity rather than providing the essential test of the CSC model, which requires testing the T-IC ability of genetically identical cells within a single subclone (Figure 5), as described recently in our study of the clonal dynamics of T-IC from human colon cancer (Kreso et al., 2013). Clearly the recent findings on subclonal diversity raise an important challenge to the validity of the broad literature on T-IC. Despite these concerns, as we will argue below, we propose that the genetic clonal evolution and CSC models can be unified into a comprehensive view of cancer heterogeneity.

entering dormancy. Thus, static genetic analysis of bulk tumor tissue or single-cell topological sampling of different tumor sites cannot formally prove that any particular genetically distinct cell or subclone is functionally important. For instance, a laser-captured cell used for genomic analysis might be on a trajectory toward death just before it is sampled and therefore is not relevant for tumor growth. Arguably, the only important cells in a tumor are the ones that are responsible for long-term clonal growth; any other cell ultimately leads to clonal exhaustion. As such, clone-propagating cells represent the unit of selection for the tumor (Greaves, 2013). As we argue above, the hallmark of such a cell is the capacity to self-renew, as without self-renewal clonal exhaustion is inevitable. Thus, a critical question that arises from the many sequencing studies that have described intra-tumoral subclonal diversity is whether diversity exists in long term propagating cancer cells. The best way to test this question is by combining cancer genetic analyses with functional T-IC assays of primary human cancers.

Three independent studies in human B-ALL and T-ALL have provided the essential evidence that subclonal genetic diversity exists within functionally defined L-IC (Anderson et al., 2011; Clappier et al., 2011; Notta et al., 2011). Elegant single cell FISH studies established that the diagnostic tumor contained genetic subclones and provided evidence for the evolutionary relationship between them. In all three studies, diagnostic samples were transplanted into xenograft recipients and since only L-IC are able to initiate leukemic propagation, the genetic makeup of the xenografts reflected the genotype of the L-IC(s) that were transplanted. Individual mice transplanted with cells from the same sample were shown to contain genetically distinct subclones, proving that genetic diversity exists among L-ICs.



**Figure 5. Experimental Approach to Investigate CSC Properties in the Context of Genetic Subclones**

Studying CSCs will require separation of distinct genetic subclones, because CSCs cannot be reliably identified in genetically heterogeneous tumors. One method by which subclones can be separated is by transplanting cancer cells at clonal cell doses over multiple recipients. Following engraftment of the human cancer cells, the hierarchical composition of a particular subclone can be studied using prospective purification of cells. Sequential transplantations of cancer cells allows for the tracking of further clonal evolution. For solid tumors, sampling different geographical regions from the primary tumor will be important for capturing distinct subclones. Although not depicted, deep sequencing and analyzing different metastatic sites can be used to analyze the genetic lineage relationships of subclones within a cancer.

Moreover, the different mutations that distinguished each subclone resulted in variability in functional properties including stemness, L-IC frequency, and aggressiveness of xenograft repopulation. Indeed, some subclones with less aggressive growth properties could only be detected in clonal assays when they were transplanted at limiting dilution, without competition from more aggressive subclones. Characterization of subclonal genetic diversity enabled reconstruction of the evolutionary process reaching back to the ancestral subclone, and demonstrated that functionally distinct subclones were related by branching evolution (Notta et al., 2011). Importantly, these studies showed that some xenografts were repopulated with a minor diagnostic subclone that was related to a paired relapse sample (Anderson et al., 2011; Clappier et al., 2011). Collectively, these xenograft studies provided functional proof for the prediction that some relapse cases arise from an undetected ancestral clone rather than through ongoing mutation of the dominant diagnostic clone (Inaba et al., 2013).

A key finding of these combined genetic and functional studies was that genotype influences L-IC frequency. B-ALL samples with *CDKN2A/B* mutation had an L-IC frequency that was on average 1,000-fold higher than that of samples without these mutations (Notta et al., 2011). Mouse models with defined genetic lesions also support the conclusion that genotype can influence T-IC properties. For example, three different mouse models of lung cancer showed variability with respect to the phenotype of the T-IC (Curtis et al., 2010). In tumors initiated with activation of oncogenic K-ras<sup>G12D</sup> and p53 deficiency, cells expressing Sca-1 could be prospectively purified as T-IC. However, in

adenocarcinomas driven by K-ras<sup>G12D</sup> alone, Sca-1 did not significantly enrich for T-IC activity. Furthermore, for adenocarcinomas expressing a mutant human EGFR transgene, only Sca-1-negative cells harbored T-IC activity (Curtis et al., 2010). This study predicts that the same markers may not identify CSCs in all patient samples of a specific tumor type. Importantly, while CSC markers varied between the different genetic tumor

models, a hierarchical organization was present within each model, supporting the notion that on top of genetic diversity, nongenetic functional variability governs tumor growth. Thus, these models highlight the influence that the genetic background of a cancer has on CSC properties.

Collectively, these data indicate that T-ICs are not static entities but can evolve. When they evolve and acquire additional mutations, the T-IC frequency can increase, indicating that the increasing genetic burden can lead to increased self-renewal as well as interfere with the malignant maturation process. Thus, a dynamic model emerges where early in tumor progression the tumor is a close caricature of the developmental hierarchy of the tissue from which it arose, with a minority T-IC and a high proportion of more differentiated non-T-IC (Figure 2). As T-IC accumulate advantageous mutations, these perturb differentiation processes further and increase self-renewal such that the T-IC expand in the subclone, reading out as having increased frequency in assays. As tumors progress, the mutational burden becomes high resulting in impairment of the remaining maturation programs and even higher capacity for self-renewal, and further expansion of cells possessing T-IC properties. In this model, as cancers progress, tumor hierarchies become shallower within genetic subclones. In some cases, once the clone has progressed to such an advanced state, the frequency of T-IC may be so high that the tumor subclone essentially becomes functionally homogeneous without evidence of a hierarchy. If this model is accurate, it could contribute a plausible explanation for why independent studies of the same tumor type (for example, melanoma) might yield different results between

labs with respect to T-IC phenotype, frequency, etc.; tumors at different stages of progression might not be comparable because their mutational burden differs.

Because mutations arise in both T-IC and non-T-IC populations, and because only T-IC contain long term propagating and self-renewing cells, it is likely that most T-IC arise from the genetic evolution of T-IC and not from the non-T-IC compartment, which lacks self-renewal. However, it is also possible that occasionally mutations will arise in non-T-IC that endow them with self-renewal capacity and convert them into T-IC. Therefore, during progression, the T-IC compartment might be a composite of T-IC generated from evolving T-IC, as well as some newly generated T-IC. Different tumors and mutations might have higher rates of non-T-IC conversion to T-IC than others. Computational simulations have shown that non-T-IC may also be important for the overall robustness of T-ICs, while T-ICs are the units of selection during the evolution of a cancer (Greaves, 2013; Sprouffske et al., 2013). We propose this revised model as a unification of the genetic evolution and developmental/CSC hierarchy models. Indeed, the other nongenetic determinants described above (noise, stochasticity, plasticity, TME) can also be accommodated as mechanisms that can convert non-T-IC into T-IC. Overall, this unified model provides a framework for future studies to determine which tumor types might follow these predictions.

### Challenge with Studying Solid Tumors

Solid tumor studies pose a particular challenge to capture the subclonal diversity present within the parent tumor. Several studies have shown that genetic subclones are topographically separate (Gerlinger et al., 2012; Yachida et al., 2010). As such, biopsy specimens that are used for research may not be representative of the entire parent tumor. For solid tumors, sampling multiple, different geographical regions from the primary tumor will be important for capturing distinct genetic subclones. This may not be possible for all solid tumors, but it will be instrumental in determining the extent of genetic subclonal variability and it will aid in interpreting concurrent CSC studies. Even if one biopsy is taken, it may represent several genetic subclones that will require separation into individual clones before CSC studies can be carried out on each subclone. While subclonal lineage relationships can be reconstructed using deep sequencing, another method by which subclones can be separated is by transplanting cancer cells at clonal cell doses over multiple recipients (Figure 5). Following engraftment of independent genetic subclones, CSC-related questions can then be addressed in each subclone. Thus, even if the entire tumor specimen is not sampled for genetic analysis, CSC studies will be carried out at least on some of the subclones present in the parent tumor. Studying CSCs in the context of independent subclones is important as all subclones and the corresponding CSCs need to be eradicated for successful therapy.

### Linking Stemness, Prognosis, and Therapy

If CSCs represent the unit of selection in tumors, as discussed above, then clinically relevant parameters including survival must be more closely related to the properties of CSCs than to non-CSCs. Testing this concept requires understanding the unique molecular circuitry of T-IC as compared to non-T-IC.

We have recently reported initial identification of an L-IC-specific transcriptional signature through gene-expression analysis of 16 AML samples that were each separated into four fractions and the L-IC activity of each fraction tested in optimized xenograft assays (Eppert et al., 2011). Functional L-IC testing was essential because the cell fraction(s) that contained L-IC were variable for each sample. The L-IC signature was prognostic for overall survival across a wide spectrum of AML patients, providing strong validation of the CSC model and of the clinical relevance of L-IC. In parallel, we carried out analysis of the transcriptional landscape of the entire spectrum of normal human HSC and progenitors. We found that HSC and L-IC gene-expression signatures converged on a shared stemness signature that was also highly prognostic (Eppert et al., 2011). We recently expanded this study and have completed gene-expression profiling of functionally defined L-IC and non-L-IC containing fractions from an additional 84 AML patient samples. Gene-expression analysis of this more representative data set generated a powerful signature that is highly prognostic when tested on approximately 1,000 AML patients in four independent cohorts. The fact that a single signature has such high prognostic power across a diverse spectrum of patients, each with distinct genetic mutations, establishes that stemness is a central biological property or process upon which many driver mutations coalesce. Recently, stemness signatures have been developed from normal stem cells in solid tissues including the intestine and breast (Merlos-Suárez et al., 2011; Pece et al., 2010). Similar to our leukemia results, the stem cell signatures were highly predictive of T-IC content and patient outcome. Overall, these early studies support a link between genetics and stemness and highlight the need to develop more stemness signatures from a wide distribution of tumor types to test the generalizability of this concept.

Strong evidence is emerging to support a link between stemness and therapy resistance in glioblastoma, colon cancer, breast cancer, and numerous other tumors, where studies show that T-IC fractions are more resistant to therapy compared to non-T-IC (Bao et al., 2006; Diehn et al., 2009; Ishikawa et al., 2007; Saito et al., 2010; Viale et al., 2009; Zhang et al., 2010). Indeed, T-IC possess as a number of biological properties that distinguish them from the remainder of tumor cells; not only resistance to treatment (Bao et al., 2006; Li et al., 2008; Tehrani et al., 2010) but also evasion of cell death (Majeti et al., 2009; Todaro et al., 2007) and dormancy (Kreso et al., 2013). While many of these papers involve *in vitro* or xenograft assays, patient data is also accumulating. In patients with 5q- myelodysplastic syndrome (MDS), complete remission can be achieved with lenalidomide treatment but patients invariably relapse. FISH analysis of bone-marrow specimens obtained prior to therapy showed that both progenitor (CD34<sup>+</sup>CD38<sup>+</sup>) and L-IC (CD34<sup>+</sup>CD38<sup>-</sup>Thy1<sup>+</sup>) compartments harbored the 5q- deletion (Tehrani et al., 2010). In most patients, the L-IC compartment was resistant to lenalidomide treatment while progenitor cells were eliminated. In one patient with clinically advanced disease, both L-IC and progenitor cell compartments were resistant to treatment, suggesting either that therapy selected for a genetic subclone with an L-IC population that is not marked by the CD34<sup>+</sup>CD38<sup>-</sup>Thy1<sup>+</sup> surface markers or that with additional mutations non-CSCs gained L-IC properties. Recent studies in

multiple myeloma, a lymphoid malignancy, have also shown that a newly identified L-IC population is resistant to proteasome inhibitor treatment compared to the bulk tumor cells (Leung-Hagesteijn et al., 2013). Collectively, these studies highlight the interplay between genetics and CSC properties that drive clinical parameters such as therapy response and ultimately survival.

The emerging evidence linking stemness to prognosis and therapy failure suggests that therapeutic targeting of determinants of stemness might be an effective means to eradicate T-IC and prevent recurrence. Although there is still considerable uncertainty as to how stemness is regulated, several regulators including Bmi-1 have been strongly linked to self-renewal and have been implicated in the maintenance of stem cells in several tissues (Lessard and Sauvageau, 2003; Molofsky et al., 2003; Park et al., 2003). We have found that human colorectal T-IC function is dependent on BMI-1. Downregulation of BMI-1 inhibits the ability of colorectal T-ICs to self-renew, resulting in abrogation of their tumorigenic potential (Kreso et al., 2014). Treatment of primary colorectal cancer xenografts with a small molecule BMI-1 inhibitor resulted in colorectal T-IC reduction with long-term and irreversible impairment of tumor growth. These studies point to the need to attempt clinically feasible targeting of this and other predicted components of the self-renewal machinery. Because stemness-associated factors are likely shared between normal stem cells and CSCs, successful eradication of CSCs will require understanding to what extent CSCs differ from normal stem cells to minimize the impact of therapies on normal stem cell function.

### The Road Ahead

Over the last several decades, there has been a revolution in our understanding of cancer growth. Advances in sequencing technologies have paved the way to deciphering the tumor genome. It is becoming increasingly clear that a tumor does not have one single tumor genome, but instead comprises multiple genomes that belong to distinct subclones. These subclones may evolve in parallel over the lifetime of a cancer and contribute to intratumoral heterogeneity. However, even within single genetic subclones, not all cells function equally: some cells retain capacity for self-renewal and long-term clonal maintenance, some lay dormant, some fuel tumor growth, and most tumor cells are postmitotic and destined for clearance.

Despite the apparent complexity, there are unifying principles rooted in developmental hierarchies that can guide our approach to targeting cancer. The litmus test for defining a dangerous cancer clone is whether the clone contains cells that exhibit unlimited growth potential. Unlimited growth potential is exhibited by the most primitive cells, which possess stemness properties such as self-renewal. Thus, by understanding stemness properties within tumors, we will be able to gain insight into the most important cells that can drive sequential rounds of tumor growth. Work in several tumor types has shown that cells with stem cell properties are equipped with innate machinery that protects them from radiation and chemotherapy. As well, stem cell gene-expression programs correlate with patient outcome, further supporting the relevance of stemness properties in cancer. By delineating genetic from nongenetic stemness influences, we will be able to tease apart the unique aspects of tumor

growth and ultimately gain a unified understanding of how diverse genetic subclones, each with their own superimposed developmental hierarchy, coordinate tumor maintenance.

In nature, evolution creates biodiversity and this in turn makes an entire ecosystem robust. In cancer, diversity within tumor cells at the genetic and functional level together with their coexistence with the microenvironment also increases tumor fitness, allowing tumor cells to offset survival pressures imposed by therapy. More effective therapies will require gaining insight into this diversity.

### ACKNOWLEDGMENTS

We would like to thank J. Wang, K. Eppert, S. Dobson, L. Shlush, F. Notta, and M. Anders for critical discussions and review of this review article.

### REFERENCES

- Abdel-Wahab, O., and Levine, R.L. (2010). EZH2 Mutations: Mutating the Epigenetic Machinery in Myeloid Malignancies. *Cancer Cell* 18, 105–107.
- Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* 100, 3983–3988.
- Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.L., et al.; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MML-Seq Consortium; ICGC PedBrain (2013). Signatures of mutational processes in human cancer. *Nature* 500, 415–421.
- Amendola, M., Giustacchini, A., Gentner, B., and Naldini, L. (2013). A double-switch vector system positively regulates transgene expression by endogenous microRNA expression (miR-ON vector). *Molecular therapy: the journal of the American Society of Gene Therapy* 21, 934–946.
- Anderson, K., Lutz, C., van Delft, F.W., Bateman, C.M., Guo, Y., Colman, S.M., Kempinski, H., Moorman, A.V., Tittle, I., Swansbury, J., et al. (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* 469, 356–361.
- Bao, S., Wu, Q., McLendon, R.E., Hao, Y., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D., and Rich, J.N. (2006). Glioma stem cells promote chemoresistance by preferential activation of the DNA damage response. *Nature* 444, 756–760.
- Barranco, S.C., Ho, D.H., Drewinko, B., Romsdahl, M.M., and Humphrey, R.M. (1972). Differential sensitivities of human melanoma cells grown in vitro to arabinosylcytosine. *Cancer Res.* 32, 2733–2736.
- Bartholdy, B., Christopheit, M., Will, B., Mo, Y., Barreyro, L., Yu, Y., Bhagat, T.D., Okoye-Okafor, U.C., Todorova, T.I., Greally, J.M., et al. (2014). HSC commitment-associated epigenetic signature is prognostic in acute myeloid leukemia. *J. Clin. Invest.* <http://dx.doi.org/10.1172/JCI71264>.
- Bateman, C.M., Colman, S.M., Chaplin, T., Young, B.D., Eden, T.O., Bhakta, M., Gratias, E.J., van Wering, E.R., Cazzaniga, G., Harrison, C.J., et al. (2010). Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood* 115, 3553–3558.
- Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome - biological and translational implications. *Nat. Rev. Cancer* 11, 726–734.
- Baylin, S.B., Weisburger, W.R., Eggleston, J.C., Mendelsohn, G., Beaven, M.A., Abeloff, M.D., and Ettinger, D.S. (1978). Variable content of histaminase, L-dopa decarboxylase and calcitonin in small-cell carcinoma of the lung. Biologic and clinical implications. *N. Engl. J. Med.* 299, 105–110.
- Bennett, D.C., Peachey, L.A., Durbin, H., and Rudland, P.S. (1978). A possible mammary stem cell line. *Cell* 15, 283–298.
- Boiko, A.D., Razorenova, O.V., van de Rijn, M., Swetter, S.M., Johnson, D.L., Ly, D.P., Butler, P.D., Yang, G.P., Joshua, B., Kaplan, M.J., et al. (2010). Human melanoma-initiating cells express neural crest nerve growth factor receptor CD271. *Nature* 466, 133–137.

- Bonner, W.A., Hulett, H.R., Sweet, R.G., and Herzenberg, L.A. (1972). Fluorescence activated cell sorting. *Rev. Sci. Instrum.* **43**, 404–409.
- Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730–737.
- Bossen, C., Ingold, K., Tardivel, A., Bodmer, J.L., Gaide, O., Hertig, S., Ambrose, C., Tschopp, J., and Schneider, P. (2006). Interactions of tumor necrosis factor (TNF) and TNF receptor family members in the mouse and human. *J. Biol. Chem.* **281**, 13964–13971.
- Bruce, W.R., and Van Der Gaag, H. (1963). A Quantitative Assay for the Number of Murine Lymphoma Cells Capable of Proliferation in Vivo. *Nature* **199**, 79–80.
- Buick, R.N., Till, J.E., and McCulloch, E.A. (1977). Colony assay for proliferative blast cells circulating in myeloblastic leukaemia. *Lancet* **1**, 862–863.
- Burrell, R.A., McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338–345.
- Cairns, J. (1975). Mutation selection and the natural history of cancer. *Nature* **255**, 197–200.
- Campbell, P.J., Yachida, S., Mudie, L.J., Stephens, P.J., Pleasance, E.D., Stebbings, L.A., Morsberger, L.A., Latimer, C., McLaren, S., Lin, M.-L., et al. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109–1113.
- Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* **154**, 61–74.
- Charles, N., Ozawa, T., Squatrito, M., Bleau, A.-M., Brennan, C.W., Hambardzumyan, D., and Holland, E.C. (2010). Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. *Cell Stem Cell* **6**, 141–152.
- Cheung, A.M., Nguyen, L.V., Carles, A., Beer, P., Miller, P.H., Knapp, D.J., Dhillon, K., Hirst, M., and Eaves, C.J. (2013). Analysis of the clonal growth and differentiation dynamics of primitive barcoded human cord blood cells in NSG mice. *Blood* **122**, 3129–3137.
- Clappier, E., Gerby, B., Sigaux, F., Delord, M., Touzri, F., Hernandez, L., Balerini, P., Baruchel, A., Pflumio, F., and Soulier, J. (2011). Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse. *J. Exp. Med.* **208**, 653–661.
- Clarke, M.F., Dick, J.E., Dirks, P.B., Eaves, C.J., Jamieson, C.H., Jones, D.L., Visvader, J., Weissman, I.L., and Wahl, G.M. (2006). Cancer stem cells—perspectives on current status and future directions: AACR Workshop on cancer stem cells. *Cancer Res.* **66**, 9339–9344.
- Clarkson, B.D. (1969). Review of recent studies of cellular proliferation in acute leukemia. *Natl. Cancer Inst. Monogr.* **30**, 81–120.
- Clarkson, B. (1974). The survival value of the dormant state in neoplastic and normal populations. In *Control of proliferation in animal cells*, B. Clarkson and R. Baserga, eds. (New York: Cold Spring Harbor Laboratory), pp. 945–972.
- Clarkson, B., Ohkita, T., Ota, K., and Oconnor, A. (1965). Studies of Cellular Proliferation in Acute Leukemia. *J. Clin. Invest.* **44**, 1035.
- Clarkson, B., Ohkita, T., Ota, K., and Fried, J. (1967). Studies of cellular proliferation in human leukemia. I. Estimation of growth rates of leukemic and normal hematopoietic cells in two adults with acute leukemia given single injections of tritiated thymidine. *J. Clin. Invest.* **46**, 506–529.
- Clarkson, B., Fried, J., Strife, A., Sakai, Y., Ota, K., and Okita, T. (1970). Studies of cellular proliferation in human leukemia. 3. Behavior of leukemic cells in three adults with acute leukemia given continuous infusions of 3H-thymidine for 8 or 10 days. *Cancer* **25**, 1237–1260.
- Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. *Nat. Med.* **17**, 313–319.
- Cohen, A.A., Geva-Zatorsky, N., Eden, E., Frenkel-Morgenstern, M., Issaeva, I., Sigal, A., Milo, R., Cohen-Saidon, C., Liron, Y., Kam, Z., et al. (2008). Dynamic proteomics of individual cancer cells in response to a drug. *Science* **322**, 1511–1516.
- Cohnheim, J. (1875). Congenitales, quergestreiftes Muskelsarkom der Nieren. *Virchows Arch.* **65**, 64–69.
- Collins, A.T., Berry, P.A., Hyde, C., Stower, M.J., and Maitland, N.J. (2005). Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res.* **65**, 10946–10951.
- Curtis, S.J., Sinkevicius, K.W., Li, D., Lau, A.N., Roach, R.R., Zamponi, R., Woolfenden, A.E., Kirsch, D.G., Wong, K.K., and Kim, C.F. (2010). Primary tumor genotype is an important determinant in identification of lung cancer propagating cells. *Cell Stem Cell* **7**, 127–133.
- Dalerba, P., Kalisky, T., Sahoo, D., Rajendran, P.S., Rothenberg, M.E., Leyrat, A.A., Sim, S., Okamoto, J., Johnston, D.M., Qian, D., et al. (2011). Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29**, 1120–1127.
- Danielson, K.G., Anderson, L.W., and Hosick, H.L. (1980). Selection and characterization in culture of mammary tumor cells with distinctive growth properties in vivo. *Cancer Res.* **40**, 1812–1819.
- Darwin, C. (1859). *On the origin of species by means of natural selection.* (London: J. Murray).
- Dexter, D.L., Kowalski, H.M., Blazar, B.A., Fligiel, Z., Vogel, R., and Heppner, G.H. (1978). Heterogeneity of tumor cells from a single mouse mammary tumor. *Cancer Res.* **38**, 3174–3181.
- Dexter, D.L., Fligiel, Z., Vogel, R., and Calabresi, P. (1979). Heterogeneity of Neoplastic-Cells from a Single Human-Colon Carcinoma. *Proc. Am. Assoc. Cancer Res.* **20**, 199–199.
- Dick, J.E. (2008). Stem cell concepts renew cancer research. *Blood* **112**, 4793–4807.
- Dick, J.E., Lapidot, T., and Pflumio, F. (1991). Transplantation of normal and leukemic human bone marrow into immune-deficient mice: development of animal models for human hematopoiesis. *Immunol. Rev.* **124**, 25–43.
- Diehn, M., Cho, R.W., Lobo, N.A., Kalisky, T., Dorie, M.J., Kulp, A.N., Qian, D., Lam, J.S., Ailles, L.E., Wong, M., et al. (2009). Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* **458**, 780–783.
- Eppert, K., Takenaka, K., Lechman, E.R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K.H., Poepl, A., Ling, V., Beyene, J., et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat. Med.* **17**, 1086–1093.
- Eramo, A., Lotti, F., Sette, G., Pilozi, E., Biffoni, M., Di Virgilio, A., Conticello, C., Ruco, L., Peschle, C., and De Maria, R. (2008). Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ.* **15**, 504–514.
- Fidler, I.J., and Kripke, M.L. (1977). Metastasis results from preexisting variant cells within a malignant tumor. *Science* **197**, 893–895.
- Garraway, L.A., and Lander, E.S. (2013). Lessons from the cancer genome. *Cell* **153**, 17–37.
- Gascoigne, K.E., and Taylor, S.S. (2008). Cancer cells display profound intra- and interline variation following prolonged exposure to antimetabolic drugs. *Cancer Cell* **14**, 111–122.
- Gavosto, F., Pileri, A., Gabutti, V., and Masera, P. (1967). Non-self-maintaining kinetics of proliferating blasts in human acute leukaemia. *Nature* **216**, 188–189.
- Gentles, A.J., Plevritis, S.K., Majeti, R., and Alizadeh, A.A. (2010). Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. *J. Am. Med. Assoc.: the journal of the American Medical Association* **304**, 2706–2715.
- Gentner, B., Visigalli, I., Hiramatsu, H., Lechman, E., Ungari, S., Giustacchini, A., Schira, G., Amendola, M., Quattrini, A., Martino, S., et al. (2010). Identification of hematopoietic stem cell-specific miRNAs enables gene therapy of globoid cell leukodystrophy. *Sci. Transl. Med.* **2**, 58ra84.
- Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892.

- Glasspool, R.M., Teodoridis, J.M., and Brown, R. (2006). Epigenetics as a mechanism driving polygenic clinical drug resistance. *Br. J. Cancer* 94, 1087–1092.
- Gray, J.M., and Pierce, G.B., Jr. (1964). Relationship between Growth Rate + Differentiation of Melanoma in Vivo. *J. Natl. Cancer Inst.* 32, 1201–1210.
- Greaves, M. (2009). Darwin and evolutionary tales in leukemia. The Ham-Wasserman Lecture. *Hematology (Am Soc Hematol Educ Program)* 2009, 3–12.
- Greaves, M. (2010). Cancer stem cells: back to Darwin? *Semin. Cancer Biol.* 20, 65–70.
- Greaves, M. (2013). Cancer stem cells as ‘units of selection’. *Evolutionary applications* 6, 102–108.
- Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. *Nature* 487, 306–313.
- Greenman, C.D., Pleasance, E.D., Newman, S., Yang, F., Fu, B., Nik-Zainal, S., Jones, D., Lau, K.W., Carter, N., Edwards, P.A.W., et al. (2012). Estimation of rearrangement phylogeny for cancer genomes. *Genome Res.* 22, 346–361.
- Guo, G., Luc, S., Marco, E., Lin, T.W., Peng, C., Kerenyi, M.A., Beyaz, S., Kim, W., Xu, J., Das, P.P., et al. (2013). Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. *Cell Stem Cell* 13, 492–505.
- Gupta, P.B., Chaffer, C.L., and Weinberg, R.A. (2009). Cancer stem cells: mirage or reality? *Nat. Med.* 15, 1010–1012.
- Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C., and Lander, E.S. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* 146, 633–644.
- Hager, J.C., Fligiel, S., Stanley, W., Richardson, A.M., and Heppner, G.H. (1981). Characterization of a variant-producing tumor cell line from a heterogeneous strain BALB/cfC3H mouse mammary tumor. *Cancer Res.* 41, 1293–1300.
- Hanahan, D., and Coussens, L.M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 21, 309–322.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Henderson, J.S., and Rous, P. (1962). The plating of tumor components on the subcutaneous expanses of young mice. Findings with benign and malignant epidermal growths and with mammary carcinomas. *J. Exp. Med.* 115, 1211–1230.
- Heppner, G.H. (1984). Tumor heterogeneity. *Cancer Res.* 44, 2259–2265.
- Heppner, G.H., Dexter, D.L., DeNucci, T., Miller, F.R., and Calabresi, P. (1978). Heterogeneity in drug sensitivity among tumor cell subpopulations of a single mammary tumor. *Cancer Res.* 38, 3758–3763.
- Hermann, P.C., Huber, S.L., Herrler, T., Aicher, A., Ellwart, J.W., Guba, M., Bruns, C.J., and Heeschen, C. (2007). Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* 1, 313–323.
- Hewitt, H.B. (1958). Studies of the dissemination and quantitative transplantation of a lymphocytic leukaemia of CBA mice. *Br. J. Cancer* 12, 378–401.
- Hong, D., Gupta, R., Ancliff, P., Atzberger, A., Brown, J., Soneji, S., Green, J., Colman, S., Piacibello, W., Buckle, V., et al. (2008). Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science* 319, 336–339.
- Inaba, H., Greaves, M., and Mullighan, C.G. (2013). Acute lymphoblastic leukaemia. *Lancet* 381, 1943–1955.
- Iorio, M.V., and Croce, C.M. (2012). MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO molecular medicine* 4, 143–159.
- Ishikawa, F., Yoshida, S., Saito, Y., Hijikata, A., Kitamura, H., Tanaka, S., Nakamura, R., Tanaka, T., Tomiyama, H., Saito, N., et al. (2007). Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. *Nat. Biotechnol.* 25, 1315–1321.
- Ishizawa, K., Rasheed, Z.A., Karisch, R., Wang, Q., Kowalski, J., Susky, E., Pereira, K., Karamboulas, C., Moghal, N., Rajeshkumar, N.V., et al. (2010). Tumor-initiating cells are rare in many human tumors. *Cell Stem Cell* 7, 279–282.
- Jan, M., Snyder, T.M., Corces-Zimmerman, M.R., Vyas, P., Weissman, I.L., Quake, S.R., and Majeti, R. (2012). Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci. Transl. Med.* 4, ra118.
- Junttila, M.R., and de Sauvage, F.J. (2013). Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* 501, 346–354.
- Kamel-Reid, S., and Dick, J.E. (1988). Engraftment of immune-deficient mice with human hematopoietic stem cells. *Science* 242, 1706–1709.
- Kamel-Reid, S., Letarte, M., Sirard, C., Doedens, M., Grunberger, T., Fulop, G., Freedman, M.H., Phillips, R.A., and Dick, J.E. (1989). A model of human acute lymphoblastic leukemia in immune-deficient SCID mice. *Science* 246, 1597–1600.
- Kamel-Reid, S., Letarte, M., Doedens, M., Greaves, A., Murdoch, B., Grunberger, T., Lapidot, T., Thorner, P., Freedman, M.H., Phillips, R.A., et al. (1991). Bone marrow from children in relapse with pre-B acute lymphoblastic leukemia proliferates and disseminates rapidly in scid mice. *Blood* 78, 2973–2981.
- Kandatho, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. *Nature* 502, 333–339.
- Killmann, S.A., Cronkite, E.P., Robertson, J.S., Flidner, T.M., and Bond, V.P. (1963). Estimation of phases of the life cycle of leukemic cells from labeling in human beings in vivo with tritiated thymidine. *Lab. Invest.* 12, 671–684.
- Klein, G., and Klein, E. (1956). Conversion of solid neoplasms into ascites tumors. *Ann. N Y Acad. Sci.* 63, 640–661.
- Kreso, A., O’Brien, C.A., van Galen, P., Gan, O.I., Notta, F., Brown, A.M., Ng, K., Ma, J., Wienholds, E., Dunant, C., et al. (2013). Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* 339, 543–548.
- Kreso, A., van Galen, P., Pedley, N.M., Lima-Fernandes, E., Frelin, C., Davis, T., Cao, L., Baiazitov, R., Du, W., Sydorenko, N., et al. (2014). Self-renewal as a therapeutic target in human colorectal cancer. *Nat. Med.* 20, 29–36.
- Kuperwasser, C., Chavarria, T., Wu, M., Magrane, G., Gray, J.W., Carey, L., Richardson, A., and Weinberg, R.A. (2004). Reconstruction of functionally normal and malignant human breast tissues in mice. *Proc. Natl. Acad. Sci. USA* 101, 4966–4971.
- Lapidot, T., Pflumio, F., Doedens, M., Murdoch, B., Williams, D.E., and Dick, J.E. (1992). Cytokine stimulation of multilineage hematopoiesis from immature human cells engrafted in SCID mice. *Science* 255, 1137–1141.
- Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367, 645–648.
- Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499, 214–218.
- Lechman, E.R., Gentner, B., van Galen, P., Giustacchini, A., Saini, M., Boccalatte, F.E., Hiramatsu, H., Restuccia, U., Bachi, A., Voisin, V., et al. (2012). Attenuation of miR-126 activity expands HSC in vivo without exhaustion. *Cell Stem Cell* 11, 799–811.
- Lessard, J., and Sauvageau, G. (2003). Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature* 423, 255–260.
- Leung-Hageteijn, C., Erdmann, N., Cheung, G., Keats, J.J., Stewart, A.K., Reece, D.E., Chung, K.C., and Tiedemann, R.E. (2013). Xbp1s-negative tumor B cells and pre-plasmablasts mediate therapeutic proteasome inhibitor resistance in multiple myeloma. *Cancer Cell* 24, 289–304.
- Ley, T.J., Ding, L., Walter, M.J., McLellan, M.D., Lamprecht, T., Larson, D.E., Kandatho, C., Payton, J.E., Baty, J., Welch, J., et al. (2010). DNMT3A mutations in acute myeloid leukemia. *N. Engl. J. Med.* 363, 2424–2433.

- Li, A., Zhou, J., Zuckerman, D., Rue, M., Dalton, V., Lyons, C., Silverman, L.B., Sallan, S.E., and Gribben, J.G. (2003). Sequence analysis of clonal immunoglobulin and T-cell receptor gene rearrangements in children with acute lymphoblastic leukemia at diagnosis and at relapse: implications for pathogenesis and for the clinical utility of PCR-based methods of minimal residual disease detection. *Blood* 102, 4520–4526.
- Li, C., Heidt, D.G., Dalerba, P., Burant, C.F., Zhang, L., Adsay, V., Wicha, M., Clarke, M.F., and Simeone, D.M. (2007). Identification of pancreatic cancer stem cells. *Cancer Res.* 67, 1030–1037.
- Li, X., Lewis, M.T., Huang, J., Gutierrez, C., Osborne, C.K., Wu, M.F., Hilsenbeck, S.G., Pavlick, A., Zhang, X., Chamness, G.C., et al. (2008). Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J. Natl. Cancer Inst.* 100, 672–679.
- Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J.P., Hector, A., Hooper, D.U., Huston, M.A., Raffaelli, D., Schmid, B., et al. (2001). Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* 294, 804–808.
- Losick, R., and Desplan, C. (2008). Stochasticity and cell fate. *Science* 320, 65–68.
- Magee, J.A., Piskounova, E., and Morrison, S.J. (2012). Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer Cell* 21, 283–296.
- Majeti, R., Chao, M.P., Alizadeh, A.A., Pang, W.W., Jaiswal, S., Gibbs, K.D., Jr., van Rooijen, N., and Weissman, I.L. (2009). CD47 is an adverse prognostic factor and therapeutic antibody target on human acute myeloid leukemia stem cells. *Cell* 138, 286–299.
- Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715.
- Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer* 12, 323–334.
- McCulloch, E.A. (1983). Stem cells in normal and leukemic hemopoiesis (Henry Stratton Lecture, 1982). *Blood* 62, 1–13.
- Meacham, C.E., and Morrison, S.J. (2013). Tumour heterogeneity and cancer cell plasticity. *Nature* 501, 328–337.
- Melchor, L., Brioli, A., Wardell, C.P., Murison, A., Potter, N.E., Kaiser, M.F., Fryer, R.A., Johnson, D.C., Begum, D.B., Hulkki Wilson, S., et al. (2014). Single-cell genetic analysis reveals the composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma. *Leukemia*. <http://dx.doi.org/10.1038/leu.2014.1013>.
- Merlos-Suárez, A., Barriga, F.M., Jung, P., Iglesias, M., Céspedes, M.V., Rossell, D., Sevillano, M., Hernando-Mombolona, X., da Silva-Diz, V., Muñoz, P., et al. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* 8, 511–524.
- Metcalfe, D., Moore, M.A., and Warner, N.L. (1969). Colony formation in vitro by myelomonocytic leukemic cells. *J. Natl. Cancer Inst.* 43, 983–1001.
- Mitelman, F. (1971). Chromosomes of 50 Primary Rous Rat Sarcomas. *Hereditas-Genetiskt Arkiv* 69, 155.
- Mitelman, F., Mark, J., Levan, G., and Levan, A. (1972). Tumor etiology and chromosome pattern. *Science* 176, 1340–1341.
- Molofsky, A.V., Pardoll, R., Iwashita, T., Park, I.K., Clarke, M.F., and Morrison, S.J. (2003). Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* 425, 962–967.
- Moore, M.A., Williams, N., and Metcalfe, D. (1973). In vitro colony formation by normal and leukemic human hematopoietic cells: characterization of the colony-forming cells. *J. Natl. Cancer Inst.* 50, 603–623.
- Mullighan, C.G., Phillips, L.A., Su, X., Ma, J., Miller, C.B., Shurtleff, S.A., and Downing, J.R. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* 322, 1377–1380.
- Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94.
- Nguyen, L.V., Vanner, R., Dirks, P., and Eaves, C.J. (2012). Cancer stem cells: an evolving concept. *Nat. Rev. Cancer* 12, 133–143.
- Nik-Zainal, S., Van Loo, P., Wedge, D.C., Alexandrov, L.B., Greenman, C.D., Lau, K.W., Raine, K., Jones, D., Marshall, J., Ramakrishna, M., et al.; Breast Cancer Working Group of the International Cancer Genome Consortium (2012). The life history of 21 breast cancers. *Cell* 149, 994–1007.
- Notta, F., Mullighan, C.G., Wang, J.C., Poepl, A., Doulatov, S., Phillips, L.A., Ma, J., Minden, M.D., Downing, J.R., and Dick, J.E. (2011). Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. *Nature* 469, 362–367.
- Nowell, P.C. (1976). The clonal evolution of tumor cell populations. *Science* 194, 23–28.
- O'Brien, C.A., Pollett, A., Gallinger, S., and Dick, J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* 445, 106–110.
- O'Brien, C.A., Kreso, A., Ryan, P., Hermans, K.G., Gibson, L., Wang, Y., Tsatsanis, A., Gallinger, S., and Dick, J.E. (2012). ID1 and ID3 regulate the self-renewal capacity of human colon cancer-initiating cells through p21. *Cancer Cell* 21, 777–792.
- Park, I.K., Qian, D., Kiel, M., Becker, M.W., Pihalja, M., Weissman, I.L., Morrison, S.J., and Clarke, M.F. (2003). Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* 423, 302–305.
- Patrawala, L., Calhoun, T., Schneider-Broussard, R., Li, H., Bhatia, B., Tang, S., Reilly, J.G., Chandra, D., Zhou, J., Claypool, K., et al. (2006). Highly purified CD44+ prostate cancer cells from xenograft human tumors are enriched in tumorigenic and metastatic progenitor cells. *Oncogene* 25, 1696–1708.
- Pece, S., Tosoni, D., Confalonieri, S., Mazzarol, G., Vecchi, M., Ronzoni, S., Bernard, L., Viale, G., Pelicci, P.G., and Di Fiore, P.P. (2010). Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell* 140, 62–73.
- Pertschuk, L.P., Tobin, E.H., Brigati, D.J., Kim, D.S., Bloom, N.D., Gaetjens, E., Berman, P.J., Carter, A.C., and Degenshein, G.A. (1978). Immunofluorescent detection of estrogen receptors in breast cancer. Comparison with dextran-coated charcoal and sucrose gradient assays. *Cancer* 41, 907–911.
- Pierce, G.B., and Cox, W.F. (1978). Neoplasms as caricatures of tissue renewal. In *Cell differentiation and neoplasia*, G.F. Saunders, ed. (New York: Rave Press), pp. 1996–2004.
- 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.
- Pierce, G.B., Jr., Dixon, F.J., Jr., and Verney, E.L. (1960). Teratocarcinogenic and tissue-forming potentials of the cell types comprising neoplastic embryoid bodies. *Lab. Invest.* 9, 583–602.
- Poste, G., Doll, J., Hart, I.R., and Fidler, I.J. (1980). In vitro selection of murine B16 melanoma variants with enhanced tissue-invasive properties. *Cancer Res.* 40, 1636–1644.
- Potter, N.E., Ermini, L., Papaemmanuil, E., Cazzaniga, G., Vijayaraghavan, G., Titley, I., Ford, A., Campbell, P., Kearney, L., and Greaves, M. (2013). Single-cell mutational profiling and clonal phylogeny in cancer. *Genome Res.* 23, 2115–2125.
- Prehn, R.T. (1970). Analysis of antigenic heterogeneity within individual 3-methylcholanthrene-induced mouse sarcomas. *J. Natl. Cancer Inst.* 45, 1039–1045.
- Prince, M.E., Sivanandan, R., Kaczorowski, A., Wolf, G.T., Kaplan, M.J., Dalerba, P., Weissman, I.L., Clarke, M.F., and Aliles, L.E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc. Natl. Acad. Sci. USA* 104, 973–978.
- Quintana, E., Shackleton, M., Foster, H.R., Fullen, D.R., Sabel, M.S., Johnson, T.M., and Morrison, S.J. (2010). Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* 18, 510–523.
- Raz, A., McLellan, W.L., Hart, I.R., Bucana, C.D., Hoyer, L.C., Sela, B.A., Dragsten, P., and Fidler, I.J. (1980). Cell surface properties of B16 melanoma variants with differing metastatic potential. *Cancer Res.* 40, 1645–1651.

- Ricci-Vitiani, L., Lombardi, D.G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C., and De Maria, R. (2007). Identification and expansion of human colon-cancer-initiating cells. *Nature* *445*, 111–115.
- Rinkevich, Y., Lindau, P., Ueno, H., Longaker, M.T., and Weissman, I.L. (2011). Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. *Nature* *476*, 409–413.
- Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., Zabierowski, S.E., Brafford, P.A., Vultur, A., Basu, D., Gimotty, P., Vogt, T., and Herlyn, M. (2010). A temporally distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* *141*, 583–594.
- Rongvaux, A., Takizawa, H., Strowig, T., Willinger, T., Eynon, E.E., Flavell, R.A., and Manz, M.G. (2013). Human hemato-lymphoid system mice: current use and future potential for medicine. *Annu. Rev. Immunol.* *31*, 635–674.
- Saito, Y., Uchida, N., Tanaka, S., Suzuki, N., Tomizawa-Murasawa, M., Sone, A., Najima, Y., Takagi, S., Aoki, Y., Wake, A., et al. (2010). Induction of cell cycle entry eliminates human leukemia stem cells in a mouse model of AML. *Nat. Biotechnol.* *28*, 275–280.
- Schwitalla, S., Fingerle, A.A., Cammareri, P., Nebelsiek, T., Göktuna, S.I., Ziegler, P.K., Canli, O., Heijmans, J., Huels, D.J., Moreaux, G., et al. (2013). Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* *152*, 25–38.
- Shackleton, M., Quintana, E., Fearon, E.R., and Morrison, S.J. (2009). Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* *138*, 822–829.
- Shah, M.Y., and Licht, J.D. (2011). DNMT3A mutations in acute myeloid leukemia. *Nat. Genet.* *43*, 289–290.
- Shapiro, J.R., Yung, W.K., and Shapiro, W.R. (1981). Isolation, karyotype, and clonal growth of heterogeneous subpopulations of human malignant gliomas. *Cancer Res.* *41*, 2349–2359.
- Sharma, S.V., Lee, D.Y., Li, B., Quinlan, M.P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M.A., et al. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* *141*, 69–80.
- Shlush, L.I., Chapal-Illani, N., Adar, R., Pery, N., Maruvka, Y., Spiro, A., Shouval, R., Rowe, J.M., Tzukerman, M., Bercovich, D., et al. (2012). Cell lineage analysis of acute leukemia relapse uncovers the role of replication-rate heterogeneity and microsatellite instability. *Blood* *120*, 603–612.
- Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. *Nature* *432*, 396–401.
- Southam, C., Brunschwig, A., and Dizon, Q. (1962). Autologous and homologous transplantation of human cancer. In *Biological interactions in normal and neoplastic growth: a contribution to the tumor-host problem*, M.J. Brennan and W.L. Simpson, eds. (Boston: Little, Brown), pp. 723–738.
- Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nat. Rev. Cancer* *6*, 846–856.
- Spencer, S.L., Gaudet, S., Albeck, J.G., Burke, J.M., and Sorger, P.K. (2009). Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* *459*, 428–432.
- Sproufske, K., Athena Aktipis, C., Radich, J.P., Carroll, M., Nedelcu, A.M., and Maley, C.C. (2013). An evolutionary explanation for the presence of cancer nonstem cells in neoplasms. *Evolutionary applications* *6*, 92–101.
- Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., Greenman, C., Wedge, D.C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G.R., et al.; Oslo Breast Cancer Consortium (OSBREAC) (2012). The landscape of cancer genes and mutational processes in breast cancer. *Nature* *486*, 400–404.
- Süel, G.M., Kulkarni, R.P., Dworkin, J., Garcia-Ojalvo, J., and Elowitz, M.B. (2007). Tunability and noise dependence in differentiation dynamics. *Science* *315*, 1716–1719.
- Tata, P.R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B.M., Vinnarsky, V., Cho, J.L., Breton, S., Sahay, A., et al. (2013). Dedifferentiation of committed epithelial cells into stem cells in vivo. *Nature* *503*, 218–223.
- Taussig, D.C., Miraki-Moud, F., Anjos-Afonso, F., Pearce, D.J., Allen, K., Ridler, C., Lillington, D., Oakervee, H., Cavenagh, J., Agrawal, S.G., et al. (2008). Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. *Blood* *112*, 568–575.
- Tehranchi, R., Woll, P.S., Anderson, K., Buza-Vidas, N., Mizukami, T., Mead, A.J., Astrand-Grundström, I., Strömbeck, B., Horvat, A., Ferry, H., et al. (2010). Persistent malignant stem cells in del(5q) myelodysplasia in remission. *N. Engl. J. Med.* *363*, 1025–1037.
- Tilman, D., Reich, P.B., and Knops, J.M.H. (2006). Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* *441*, 629–632.
- Todaro, M., Alea, M.P., Di Stefano, A.B., Cammareri, P., Vermeulen, L., Iovino, F., Tripodo, C., Russo, A., Gulotta, G., Medema, J.P., and Stassi, G. (2007). Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4. *Cell Stem Cell* *1*, 389–402.
- 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.
- Vermeulen, L., De Sousa E Melo, F., van der Heijden, M., Cameron, K., de Jong, J.H., Borovski, T., Tuynman, J.B., Todaro, M., Merz, C., Rodemond, H., et al. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nat. Cell Biol.* *12*, 468–476.
- Viale, A., De Franco, F., Orleth, A., Cambiaghi, V., Giuliani, V., Bossi, D., Ronchini, C., Ronzoni, S., Muradore, I., Monestiroli, S., et al. (2009). Cell-cycle restriction limits DNA damage and maintains self-renewal of leukaemia stem cells. *Nature* *457*, 51–56.
- Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M., and Bos, J.L. (1988). Genetic alterations during colorectal-tumor development. *N. Engl. J. Med.* *319*, 525–532.
- Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., et al. (2012). The origin and evolution of mutations in acute myeloid leukemia. *Cell* *150*, 264–278.
- Wernet, M.F., Mazzoni, E.O., Celik, A., Duncan, D.M., Duncan, I., and Desplan, C. (2006). Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature* *440*, 174–180.
- Wu, C., Wei, Q., Utomo, V., Nadesan, P., Whetstone, H., Kandel, R., Wunder, J.S., and Alman, B.A. (2007). Side population cells isolated from mesenchymal neoplasms have tumor initiating potential. *Cancer Res.* *67*, 8216–8222.
- Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R.H., Eshleman, J.R., Nowak, M.A., et al. (2010). Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* *467*, 1114–1117.
- 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.
- Yung, W.K.A., Shapiro, J.R., and Shapiro, W.R. (1982). Heterogeneous chemosensitivities of subpopulations of human glioma cells in culture. *Cancer Res.* *42*, 992–998.
- Zhang, M., Atkinson, R.L., and Rosen, J.M. (2010). Selective targeting of radiation-resistant tumor-initiating cells. *Proc. Natl. Acad. Sci. USA* *107*, 3522–3527.
- Zuna, J., Ford, A.M., Peham, M., Patel, N., Saha, V., Eckert, C., Kochling, J., Panzer-Grumayer, R., Trka, J., and Greaves, M. (2004). TEL deletion analysis supports a novel view of relapse in childhood acute lymphoblastic leukemia. *Clinical cancer research: an official journal of the American Association for Cancer Research* *10*, 5355–5360.

# Leptin-Receptor-Expressing Mesenchymal Stromal Cells Represent the Main Source of Bone Formed by Adult Bone Marrow

Bo O. Zhou,<sup>1</sup> Rui Yue,<sup>1</sup> Malea M. Murphy,<sup>1</sup> James G. Peyer,<sup>1</sup> and Sean J. Morrison<sup>1,2,\*</sup>

<sup>1</sup>Children's Research Institute and Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

<sup>2</sup>Howard Hughes Medical Institute

\*Correspondence: sean.morrison@utsouthwestern.edu

<http://dx.doi.org/10.1016/j.stem.2014.06.008>

## SUMMARY

Studies of the identity and physiological function of mesenchymal stromal cells (MSCs) have been hampered by a lack of markers that permit both prospective identification and fate mapping *in vivo*. We found that Leptin Receptor (LepR) is a marker that highly enriches bone marrow MSCs. Approximately 0.3% of bone marrow cells were LepR<sup>+</sup>, 10% of which were CFU-Fs, accounting for 94% of bone marrow CFU-Fs. LepR<sup>+</sup> cells formed bone, cartilage, and adipocytes in culture and upon transplantation *in vivo*. LepR<sup>+</sup> cells were *Scf*-GFP<sup>+</sup>, *Cxcl12*-DsRed<sup>high</sup>, and *Nestin*-GFP<sup>low</sup>, markers which also highly enriched CFU-Fs, but negative for *Nestin*-CreER and *NG2*-CreER, markers which were unlikely to be found in CFU-Fs. Fate-mapping showed that LepR<sup>+</sup> cells arose postnatally and gave rise to most bone and adipocytes formed in adult bone marrow, including bone regenerated after irradiation or fracture. LepR<sup>+</sup> cells were quiescent, but they proliferated after injury. Therefore, LepR<sup>+</sup> cells are the major source of bone and adipocytes in adult bone marrow.

## INTRODUCTION

Multipotent mesenchymal stromal cells (MSCs) have been defined in culture as nonhematopoietic, plastic-adherent, colony-forming cells, referred to as colony-forming unit-fibroblasts (CFU-Fs), that can differentiate into osteogenic, chondrogenic, and adipogenic progeny (Bianco et al., 2008; Friedenstein et al., 1970; Horwitz et al., 2005; Pittenger et al., 1999). A fundamental question concerns the identity of the cells that form CFU-Fs in culture and their physiological function *in vivo*.

*In vivo*, MSCs are often perivascular (Crisan et al., 2008; Sacchetti et al., 2007). In human tissues, MSCs have been prospectively identified based on the lack of expression of endothelial and hematopoietic markers along with positive expression for CD146 (Sacchetti et al., 2007) or CD146 with platelet-derived growth factor receptor  $\beta$  (PDGFR $\beta$ ), CD106, NG2, or some combination thereof (Chou et al., 2012; Crisan et al., 2008; Schwab

and Gargett, 2007). CD146<sup>+</sup> MSCs from human bone marrow form osteogenic, chondrogenic, and adipogenic cells in culture and give rise to bone upon transplantation *in vivo*, forming bony ossicles that become invested with hematopoietic bone marrow (Sacchetti et al., 2007). The CD146<sup>+</sup> cells persist around sinusoidal blood vessels in the ossicles and express HSC niche factors. Ectopic bones that become invested with bone marrow can also be formed by CD105<sup>+</sup>Thy1<sup>-</sup> mesenchymal cells from fetal mouse bones (Chan et al., 2009). Although much has been learned about the localization and developmental potential of MSCs, limitations in the ability to fate-map these cells *in vivo* have hindered our understanding of their normal physiological function.

Mouse MSCs have been prospectively identified based on the lack of expression of hematopoietic and endothelial markers and positive expression of PDGFR $\alpha$  (Morikawa et al., 2009; Omatsu et al., 2010; Park et al., 2012). The PDGFR $\alpha$ <sup>+</sup>Sca-1<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> subset of cells appears to reside primarily around arterioles but does not express the hematopoietic stem cell (HSC) niche factor *Cxcl12*, while PDGFR $\alpha$ <sup>+</sup>Sca-1<sup>-</sup>CD45<sup>-</sup>Ter119<sup>-</sup> cells that express high levels of *Cxcl12*, also known as CXCL12-abundant reticular (CAR) cells, reside primarily around sinusoids (Morikawa et al., 2009; Omatsu et al., 2010). CAR cells include bipotent progenitors of osteoblasts and adipocytes (Omatsu et al., 2010). Ablation of CAR cells with diphtheria toxin (DT) depletes not only adipogenic and osteogenic progenitors, but also HSCs (Omatsu et al., 2010). These data suggest that MSCs within the bone marrow are a cellular component of the HSC niche.

Cells that express the *Nestin*-GFP transgene contain all of the CFU-Fs in mouse bone marrow (Kunisaki et al., 2013; Méndez-Ferrer et al., 2010). These cells also express high levels of the HSC niche factors *Cxcl12* and *Stem cell factor* (*Scf*) (Kunisaki et al., 2013; Méndez-Ferrer et al., 2010). *Nestin*-CreER-expressing cells in adult bone marrow can contribute to osteoblasts and chondrocytes, though no quantification was provided regarding their CFU-F content or their level of contribution to new bone (Méndez-Ferrer et al., 2010). *Nestin*-GFP is widely expressed by perivascular stromal cells in the bone marrow, but they express little endogenous *Nestin*, *Nestin*-CreER, or other *Nestin* transgenes (Ding et al., 2012). Furthermore, *Nestin*-GFP<sup>+</sup> cells are heterogeneous, including both *Nestin*-GFP<sup>high</sup> cells that localize mainly around arterioles and *Nestin*-GFP<sup>low</sup> cells that localize mainly around sinusoids, raising questions about the distribution of HSC niche factors and MSCs among these cell populations (Kunisaki et al., 2013).

During fetal development, Osterix (*Osx*) is expressed by cells that form bone marrow stroma, including perivascular, osteogenic, and adipogenic cells (Liu et al., 2013; Maes et al., 2010). In adult bone marrow, *Osx* is expressed by a subset of CAR cells (Omatsu et al., 2010), and fate mapping of these cells using *Osx*-CreER revealed a transient contribution to osteoblasts (Park et al., 2012). In contrast, *Mx-1*-Cre recombined widely among hematopoietic cells, PDGFR $\alpha$ <sup>+</sup> stromal cells, and CFU-Fs in the bone marrow, and these cells gave rise to most of the osteoblasts formed in adult bone marrow (Park et al., 2012). The widespread marking of both hematopoietic and stromal cells by *Mx-1*-Cre meant that this marker could not be used to prospectively identify osteogenic progenitors.

Recently, we identified Leptin Receptor<sup>+</sup> (LepR<sup>+</sup>) perivascular stromal cells that are the major source of *Scf* and *Cxcl12* in the bone marrow (Ding and Morrison, 2013; Ding et al., 2012). Conditional deletion of *Scf* with *Lepr*-Cre led to the depletion of quiescent HSCs (Ding et al., 2012; Oguro et al., 2013) and conditional deletion of *Cxcl12* with *Lepr*-Cre led to HSC mobilization (Ding and Morrison, 2013). Here we show that LepR<sup>+</sup> cells are highly enriched for CFU-F and uniformly express *Prx1*-Cre, PDGFR $\alpha$ , and CD51, markers expressed by bone marrow MSCs. LepR<sup>+</sup> cells were the main source of new osteoblasts and adipocytes in adult bone marrow and could form bony ossicles that support hematopoiesis in vivo. In contrast, *Wnt1*-Cre-expressing neural-crest-derived cells and *Nestin*-CreER-expressing cells included few CFU-Fs and made little contribution to adult osteogenesis.

## RESULTS

### LepR<sup>+</sup> Bone Marrow Stromal Cells Are Around Sinusoids and Arterioles

Sections from wild-type mice exhibited perivascular LepR staining throughout the bone marrow, around both sinusoids and arterioles (Figures 1A, 1C, and 1D). The staining was in perivascular stromal cells that expressed the HSC niche factor *Scf* (Figure 1C). An antibody against the LepR extracellular domain stained in a pattern very similar to that of Tomato expression in *Lepr-cre*; *tdTomato* conditional reporter mice (Figure 1D). *Ubc-creER*; *Lep<sup>fl/fl</sup>* mice that had been treated with tamoxifen for 1 month to conditionally delete *Lepr* had little staining with the antibody in sections (Figure 1B) or in PDGFR $\alpha$ <sup>+</sup> CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> bone marrow stromal cells analyzed by flow cytometry (Figure S1A available online).

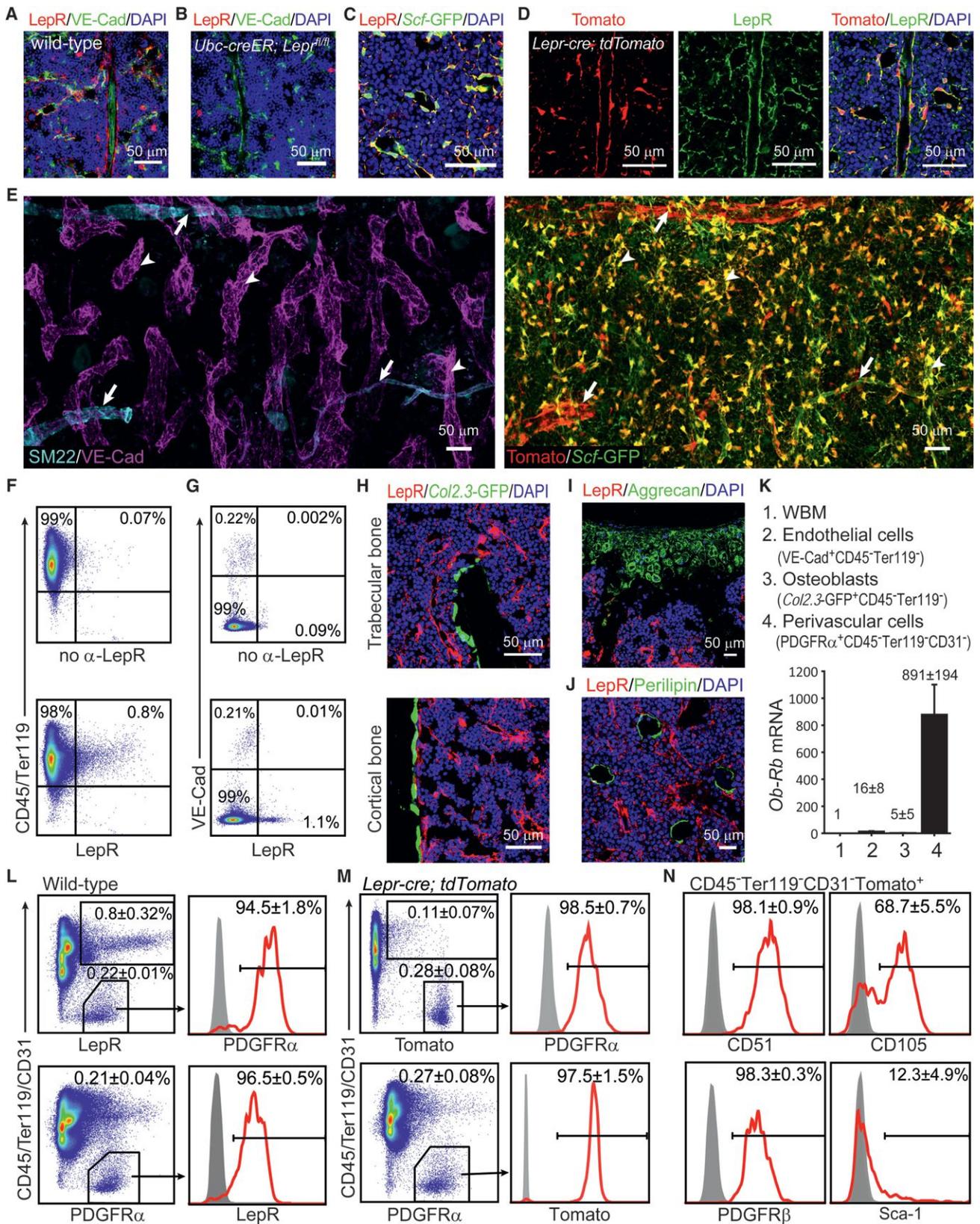
We identified sinusoids and arterioles based on VE-Cadherin (VE-Cad) staining, which bound endothelial cells in both sinusoids and arterioles, and SM22 staining, which specifically marked vascular smooth muscle around arterioles. Sinusoids were typically larger in diameter, less uniform, and thinner walled as compared to arterioles (Figure 1E). We observed LepR<sup>+</sup> cells around both sinusoids and arterioles throughout the bone marrow, though LepR<sup>+</sup> cells were much more prominent around some arterioles than others (Figure 1E). Nearly all the perisinusoidal LepR<sup>+</sup> cells were *Scf*-GFP<sup>+</sup>; however, the periarteriolar LepR<sup>+</sup> cells, especially those densely surrounding larger arterioles, expressed less *Scf*-GFP (Figure 1E). The LepR<sup>+</sup> cells around arterioles were negative for SM22 or  $\alpha$ SMA (Figures 1E and S1C).

We observed little LepR antibody staining in CD45<sup>+</sup> or Ter119<sup>+</sup> hematopoietic cells (Figure 1F) or in VE-Cad<sup>+</sup> bone marrow endothelial cells (Figures 1A, 1E, and 1G). The rare LepR antibody staining that was observed in these cells may reflect the expression of short isoforms of *Lepr* that lack the intracellular signaling domain (*Ob-Ra*, *Ob-Rc*, *Ob-Rd*, and *Ob-Re*). These isoforms are somewhat more broadly expressed than the *Ob-Rb* isoform, which encodes full-length LepR, including the intracellular signaling domain. It is this full-length isoform whose expression is marked by *Lepr-cre*. Little recombination was observed in macrophages (Figure S1D) or in other hematopoietic cells (Figure 1M) using *Lepr-cre*.

Consistent with the lack of EYFP expression in bone-lining cells from 2-month-old *Lepr-cre*; *EYFP* reporter mice (Ding et al., 2012), we were unable to detect LepR antibody staining in *Col2.3*-GFP<sup>+</sup> osteoblastic cells from cortical or trabecular bone in 2-month-old mice (Figure 1H). Neither Aggrecan<sup>+</sup> articular cartilage cells in the femur (knee joint; Figure 1I) nor Perilipin<sup>+</sup> fat cells in the bone marrow (Figure 1J) exhibited LepR antibody staining. Thus, at 2 months of age, LepR expression was largely restricted to perivascular stromal cells and not to more differentiated mesenchymal derivatives in bone marrow.

Consistent with the LepR antibody staining and *Lepr-cre*; *tdTomato* conditional reporter expression pattern, quantitative real-time PCR (qPCR) showed that full-length *Ob-Rb* transcripts were at 100- to 1,000-fold higher levels in PDGFR $\alpha$ <sup>+</sup> CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> perivascular stromal cells as compared to unfractionated bone marrow cells, *Col2.3*-GFP<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> osteoblastic cells, and VE-Cad<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow endothelial cells (Figure 1K). Virtually all LepR<sup>+</sup> cells expressed *Scf*-GFP and nearly all *Scf*-GFP<sup>+</sup> cells expressed LepR (Figure S1E). LepR<sup>+</sup> *Scf*-GFP<sup>-</sup> cells around certain arterioles appeared to represent <1% of LepR<sup>+</sup> cells in bone marrow (Figure 1E). Nearly all LepR<sup>+</sup> cells expressed high levels of *Cxcl12*-DsRed and nearly all cells that expressed high levels of *Cxcl12*-DsRed expressed LepR (Figure S1F). Most LepR<sup>+</sup> cells expressed low levels of *Nestin*-GFP (Figures S1G and S1H), consistent with a recent report (Kunisaki et al., 2013). Cells that expressed high levels of *Nestin*-GFP did not stain with either LepR or PDGFR $\alpha$  (Figure S1H). The vast majority of bone marrow cells that express high levels of HSC niche factors and the bone marrow MSC marker PDGFR $\alpha$  are thus LepR<sup>+</sup>.

LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells accounted for 0.2% to 0.3% of enzymatically dissociated bone marrow cells, irrespective of whether these cells were identified by LepR antibody staining (Figure 1L) or Tomato expression in *Lepr-cre*; *tdTomato* conditional reporter mice (Figure 1M). Nearly all LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> bone marrow stromal cells were positive for PDGFR $\alpha$  and nearly all PDGFR $\alpha$ <sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> bone marrow cells were LepR<sup>+</sup> (Figures 1L and 1M). These data suggested that LepR<sup>+</sup> bone marrow stromal cells might be highly enriched for MSCs. Consistent with this possibility, we found that LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> bone marrow stromal cells were uniformly positive for the MSC markers CD51 (Pinho et al., 2013) and PDGFR $\beta$  (Komada et al., 2012) (Figure 1N). Approximately 68% of LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> cells were positive for the MSC marker CD105 (Chan et al., 2009; Park et al., 2012) (Figure 1N). LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> cells were heterogeneous for Sca-1 (Figure 1N), which is



(legend on next page)

expressed by a subset of MSCs (Morikawa et al., 2009; Omatsu et al., 2010).

### LepR<sup>+</sup> Cells Are the Main Source of CFU-Fs in Bone Marrow

To assess CFU-F activity we enzymatically dissociated bone marrow cells and added them to adherent cultures at clonal density. Figure 2B shows the percentage of cells in each cell population sorted from unfractionated bone marrow cells (including both hematopoietic and stromal elements) that formed CFU-F colonies in culture. Figure 2C shows the percentage of cells in each cell population sorted from nonhematopoietic (CD45<sup>-</sup>Ter119<sup>-</sup>) bone marrow stromal cells that formed CFU-F colonies in culture. In our experiments, 0.012% ± 0.002% of all enzymatically dissociated bone marrow cells (Figure 2B) or 1.3% ± 0.5% of CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells formed CFU-F colonies (Figure 2C).

Consistent with a prior study (Morikawa et al., 2009), 9.8% ± 5.0% of PDGFRα<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells, 15.6% ± 3.0% of PDGFRα<sup>+</sup>Sca-1<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells, and 8.3% ± 5.2% of PDGFRα<sup>+</sup>Sca-1<sup>-</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells formed CFU-F colonies (Figure 2C). Although PDGFRα<sup>+</sup>Sca-1<sup>+</sup> stromal cells were more highly enriched for CFU-F activity than PDGFRα<sup>+</sup>Sca-1<sup>-</sup> cells (Figure 2C), PDGFRα<sup>+</sup>Sca-1<sup>-</sup> stromal cells contained most of the CFU-F activity in the bone marrow (Figure S2A) because they were much more abundant than PDGFRα<sup>+</sup>Sca-1<sup>+</sup> cells (0.22% ± 0.08% versus 0.03% ± 0.02% of bone marrow cells; Figures 1N and S1J). PDGFRα<sup>+</sup> bone marrow cells are thus highly enriched for CFU-F, but we could not fate map PDGFRα<sup>+</sup> cells because *Pdgfra*-CreER (Rivers et al., 2008) recombined poorly in bone marrow PDGFRα<sup>+</sup> cells (data not shown).

The vast majority of bone marrow CFU-Fs derive from LepR<sup>+</sup> cells. In CFU-F colonies formed by bone marrow cells from *Lepr-cre; tdTomato* mice, 94% ± 4% were Tomato<sup>+</sup> (Figures 2A and S2B–S2E). In 2- to 4-month-old *Lepr-cre; tdTomato*

mice, 8.5% ± 2.5% of all Tomato<sup>+</sup> bone marrow cells (Figure 2B) and 11% ± 3.5% of Tomato<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells formed CFU-F colonies (Figure 2C), whereas only 0.10% ± 0.08% of Tomato<sup>-</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells formed CFU-F colonies (Figure 2C). Consistent with a previous study (Park et al., 2012), almost all of the Tomato<sup>+</sup> CFU-Fs in *Lepr-cre; tdTomato* bone marrow were CD105<sup>+</sup> (Figure 2C). Cells derived from full-length LepR<sup>+</sup> cells in the bone marrow were thus as highly enriched for CFU-Fs as PDGFRα<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow cells.

We split cells obtained from individual CFU-F colonies formed by LepR<sup>+</sup> cells into three aliquots and subcloned them into cultures that promoted bone, cartilage, or fat cell differentiation. In CFU-F colonies formed by Tomato<sup>+</sup> bone marrow cells from *Lepr-cre; tdTomato* mice, 8.9 ± 4.5% underwent multilineage differentiation (Figure 2D), giving rise to Alizarin-red-S-stained osteoblastic cells (Figures S2F and S2G), Oil-red-O-stained adipocytes (Figures S2H and S2I), and Toluidine-blue-stained chondrocytic cells (Figures S2J and S2K). Most of the remaining CFU-F colonies formed by Tomato<sup>+</sup> cells differentiated to osteoblastic cells with or without chondrocytic cells or fat cells. Overall, 58% ± 17% of all CFU-F colonies formed by Tomato<sup>+</sup> cells formed osteoblastic cells in culture.

We also sorted individual LepR<sup>+</sup> cells from *Lepr-cre; tdTomato; Col2.3-GFP* mice into culture, allowed them to form CFU-F colonies, then expanded individual colonies, implanted the cells into denatured collagen sponges, transplanted them subcutaneously, and assessed the development of bony ossicles and hematopoiesis 8 weeks later. The presence of hematopoiesis was determined based on the presence of undifferentiated and differentiated erythroid, myeloid, and lymphoid cells in ossicle sections. Fourteen of twenty-five sponges did not form bone. Of the 11 that did form bone, we detected hematopoiesis in 8 (Figures 2D and S2L–S2N), 3 of which had abundant hematopoiesis that resembled bone marrow and 5 of which had smaller foci of hematopoietic cells distributed throughout each ossicle. In each case, we observed donor-derived

### Figure 1. LepR and Scf-GFP-Expressing Cells Are Abundant around Sinusoids throughout the Bone Marrow and Are Distinct from Other Stromal Cells

(A and B) Representative femur sections from 3- to 4-month-old wild-type (A) and *Ubc-creER; LepR<sup>fl/fl</sup>* mice (B). The anti-LepR antibody stained perivascular cells in wild-type (A) but not *Ubc-creER; LepR<sup>fl/fl</sup>* (B) bone marrow (unless otherwise indicated, each panel reflects data from three mice/genotype from three independent experiments).

(C) Staining with anti-LepR antibody and Scf-GFP.

(D) Staining with anti-LepR antibody strongly overlapped with Tomato expression around sinusoids and arterioles in the bone marrow of *Lepr-cre; tdTomato* mice.

(E) 3D reconstruction of a Z stack of tiled confocal images of femur bone marrow from a *Lepr-cre; tdTomato; Scf-GFP* mouse. Anti-VE-Cad staining marked sinusoids (arrowheads, left panel) and arterioles while anti-SM22 staining specifically marked arterioles (arrows, left panel). Left and right panels represent images from the same field of view. LepR was expressed by perivascular cells around sinusoids and arterioles, but LepR<sup>+</sup>Scf-GFP<sup>+</sup> cells were most abundant around sinusoids. Note that most Scf-GFP staining that did not overlap with Tomato staining represented the processes of perivascular cells that had Tomato staining in their cell body (see Figure S1B). The frequency of Scf-GFP<sup>+</sup> cells appears high in this image because it represents a Z stack of images from a thick section, not a single optical section.

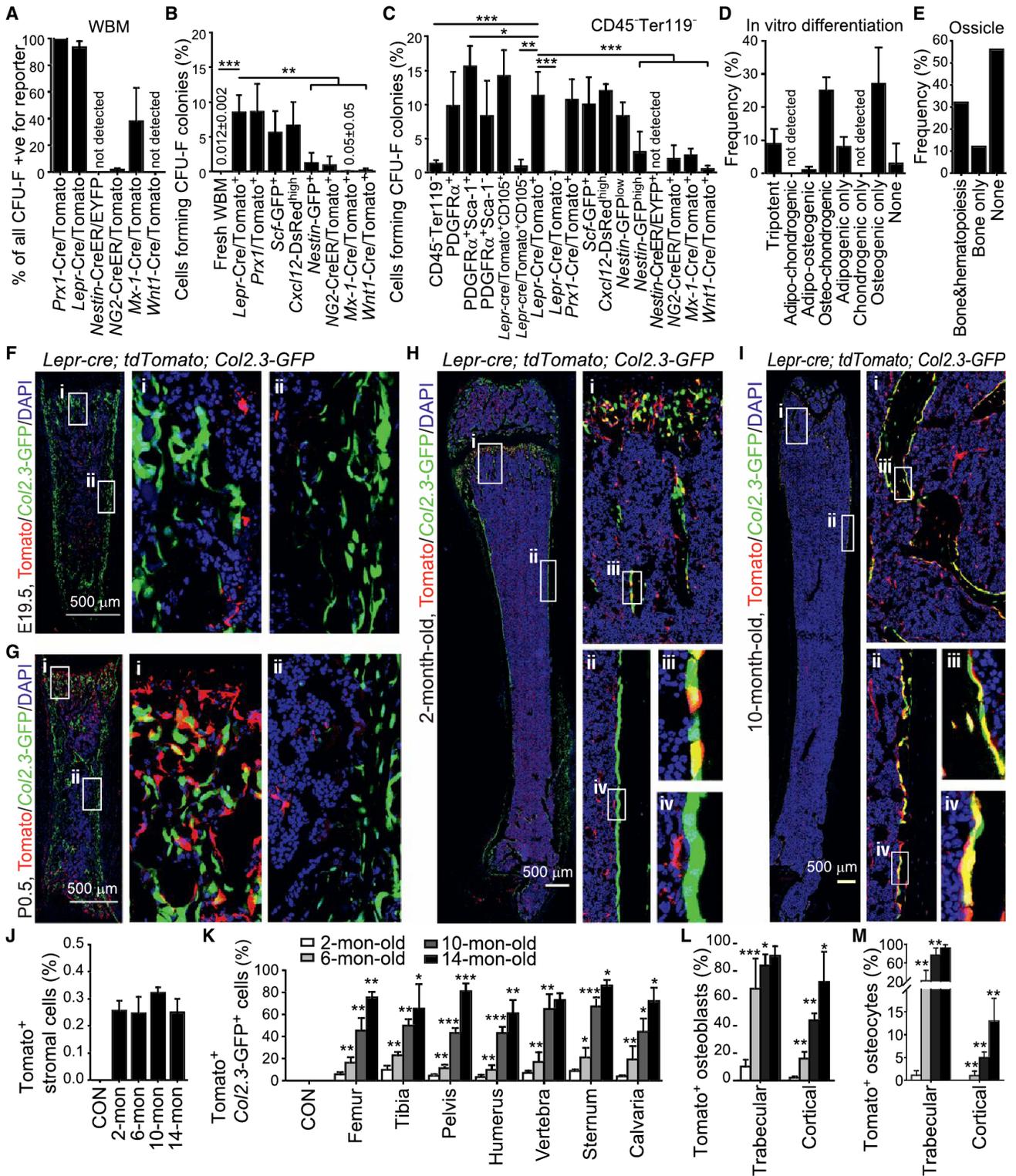
(F and G) Flow cytometry analysis showed that CD45<sup>+</sup>Ter119<sup>+</sup> hematopoietic cells (F) and VE-Cad<sup>+</sup> endothelial cells (G) rarely stained positively for LepR. The bone marrow was dissociated mechanically in (F) (thus lacking stroma) or enzymatically in (G) (including stroma).

(H–J) *Col2.3-GFP<sup>+</sup>* osteoblasts (H), Aggrecan<sup>+</sup> chondrocytes (I), and Perilipin<sup>+</sup> adipocytes (J) did not stain with anti-LepR antibody. n = 3–5 mice from at least three independent experiments.

(K) Quantitative RT-PCR of *Ob-Rb* transcript levels (normalized to *β-Actin*). Data represent mean ± SD (standard deviation) from four independent experiments.

(L and M) In 2- to 4-month-old mice, nearly all LepR<sup>+</sup> bone marrow cells stained positively for PDGFRα, and vice versa, irrespective of whether the LepR<sup>+</sup> cells were identified by antibody staining (L) or Tomato expression in *Lepr-cre; tdTomato* mice (M). The data represent mean ± SD from 3–5 mice from at least three independent experiments.

(N) Marker expression by Tomato<sup>+</sup> bone marrow cells from *Lepr-cre; tdTomato* mice.



**Figure 2. *Lepr*<sup>+</sup> Cells Contain Most of the CFU-Fs in Adult Bone Marrow and Are the Major Source of New Bone in Adult Mice**

(A) Percentage of all CFU-F colonies that were labeled by conditional reporter expression when cultured from enzymatically dissociated bone marrow of the indicated genotypes. Macrophage colonies were excluded by staining with anti-CD45 antibody in all experiments. n = 3–11 mice/genotype from at least three independent experiments.

(B) Percentage of bone marrow cells expressing each marker that formed CFU-F colonies in culture (n = 3–5 mice/genotype from at least three independent experiments).

(legend continued on next page)

(Tomato<sup>+</sup>) bone, stromal cells, and adipocytes in addition to host-derived CD45<sup>+</sup> hematopoietic cells (Figures S2L and S2M). Cells from the three ossicles that contained the most hematopoietic cells gave multilineage reconstitution of irradiated mice, demonstrating the presence of primitive hematopoietic progenitors (data not shown).

When bone marrow cells were obtained from the femurs and tibias of *Prx1-cre; tdTomato* conditional reporter mice, all CFU-F colonies formed by unfractionated bone marrow cells were Tomato<sup>+</sup> (Figure 2A). In 2- to 4-month-old *Prx1-cre; tdTomato* mice, 8.6% ± 4.0% of all Tomato<sup>+</sup> bone marrow cells (Figure 2B) and 10% ± 3% of Tomato<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> bone marrow stromal cells formed CFU-F colonies (Figure 2C). Nearly all LepR<sup>+</sup> antibody stained cells were Tomato<sup>+</sup> in *Prx1-cre; tdTomato* mice and nearly all Tomato<sup>+</sup> cells were LepR<sup>+</sup> (Figures S2P and S2Q). Thus, LepR<sup>+</sup> cells in the marrow of limb bones uniformly express the MSC marker *Prx1*.

Consistent with the conclusion that MSCs secrete factors that promote HSC maintenance (Méndez-Ferrer et al., 2010; Omatsu et al., 2010; Sacchetti et al., 2007), both *Scf-GFP*<sup>+</sup> bone marrow cells and *Cxcl12-DsRed*<sup>high</sup> bone marrow cells were highly enriched for CFU-Fs (Figures 2B and 2C). The observation that these cells are similarly enriched for CFU-F and LepR<sup>+</sup> cells is consistent with our data demonstrating that nearly all cells that express *Scf-GFP* or high levels of *Cxcl12-DsRed* are LepR<sup>+</sup> (Figures S1D and S1E). These data suggest a strong overlap between LepR<sup>+</sup> stromal cells and CAR cells in the bone marrow.

CFU-Fs often expressed *Mx1-Cre*, consistent with a recent study (Park et al., 2012). Bone marrow sections from *Mx1-cre; tdTomato* mice showed widespread Tomato expression among hematopoietic cells, vascular cells, perivascular cells, and bone-lining osteoblastic cells (Figures S4B–S4D). We found that 37% ± 25% of all CFU-F colonies formed by *Mx1-cre; tdTomato* bone marrow cells were Tomato<sup>+</sup> (Figure 2A) and 46% ± 23% of all CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup>PDGFR $\alpha$ <sup>+</sup> stromal cells were Tomato<sup>+</sup> (Figure S4A). Only rare Tomato<sup>+</sup> bone marrow cells formed CFU-F colonies (Figures 2B and 2C) due to widespread recombination among hematopoietic cells and other stromal cells (Figure S4A).

CFU-Fs rarely expressed *NG2-CreER*. Only 1.8% ± 1.0% of CFU-F colonies formed by bone marrow cells from *NG2-creER; tdTomato* reporter mice were Tomato<sup>+</sup> (Figure 2A). Tomato<sup>+</sup> cells accounted for 0.0026% ± 0.007% of all bone marrow cells in *NG2-creER; tdTomato* mice and 8.8% ± 3.6% of these cells

stained positively for the MSC marker PDGFR $\alpha$  (Figure S3A). Only 9.8% ± 3.4% of Tomato<sup>+</sup> bone marrow cells in *NG2-creER; tdTomato; Scf-GFP* mice were positive for *Scf-GFP* (Figure S3A). We observed almost no *NG2-CreER*-expressing cells among PDGFR $\alpha$ <sup>+</sup> cells or *Scf-GFP*<sup>+</sup> cells in the bone marrow (Figures S3B–S3D).

In the bone marrow, *NG2-creER* labeled vascular smooth muscle cells (Figure S3E) and GFAP<sup>+</sup> glia associated with nerve fibers (Figure S3F), chondrocytes, osteocytes, and rare osteoblasts (data not shown). *NG2*-antibody-stained cells also included Aggrecan<sup>+</sup> chondrocytes (Figure S3G), *Col2.3-GFP*<sup>+</sup> osteoblasts, and osteocytes (Figure S3H) in 2-month-old mice. Smooth muscle cells (Murfee et al., 2005), peripheral nerve Schwann cells (Schneider et al., 2001), cartilaginous cells, and osteoblasts (Fukushi et al., 2003) have all been previously reported to express *NG2*.

CFU-Fs also rarely expressed *Nestin-creER*. We did not detect any EYFP<sup>+</sup> CFU-F colonies formed by bone marrow cells from *Nestin-creER; loxp-EYFP* mice (Figure 2A). We observed labeling of rare cells associated with some arterioles in the bone marrow of *Nestin-creER; loxp-EYFP* mice (Figure S4F) consistent with previously published images (Méndez-Ferrer et al., 2010). EYFP<sup>+</sup> cells accounted for 0.0012% ± 0.0005% of bone marrow cells and were nearly uniformly negative for PDGFR $\alpha$  (Figure S4E). When we aged *Nestin-creER; loxp-EYFP* mice for 11 months, we only detected rare EYFP<sup>+</sup> osteoblasts (Figure S4F), demonstrating that *Nestin-CreER*-expressing cells are not a significant source of bone in vivo.

The observation that *Nestin-GFP*<sup>+</sup> cells included CFU-Fs, but that *Nestin-CreER*-expressing cells did not (Figures 2C and S1H), is consistent with our observation that different *Nestin* transgenes exhibit different expression patterns in the bone marrow (Ding et al., 2012). Neither *Nestin-GFP*<sup>low</sup> nor *Nestin-GFP*<sup>high</sup> bone marrow cells appear to express endogenous *Nestin* (see microarray data in Méndez-Ferrer et al., 2010 and RNaseq data in Kunisaki et al., 2013).

CFU-Fs were not neural crest-derived based on fate mapping with *Wnt1-Cre* (Chai et al., 2000; Echelard et al., 1994; Joseph et al., 2004). We observed Tomato<sup>+</sup> nerve fibers and glia in the bone marrow of *Wnt1-cre; tdTomato* mice (Figure S4H), confirming that neural-crest-derived cells were marked in the bone marrow. In 2- to 4-month-old *Wnt1-cre; tdTomato* mice, 0.17% ± 0.29% of all Tomato<sup>+</sup> bone marrow cells (Figure 2B) and 0.5% ± 0.5% of Tomato<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup> cells (Figure 2C)

(C) Percentage of nonhematopoietic (CD45<sup>-</sup>Ter119<sup>-</sup>) bone marrow cells expressing each marker that formed CFU-F colonies in culture (n = 3–11 mice/genotype from at least three independent experiments). Two-tailed Student's t tests were used to assess statistical significance. \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001.

(D) The percentage of CFU-F colonies that arose from Tomato<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> cells from *LepR-cre; tdTomato* mice that gave rise to Oil red O<sup>+</sup> adipocytes, Toluidine blue<sup>+</sup> chondrocytes, and/or Alizarin red S<sup>+</sup> osteoblasts (n = 3 mice from three independent experiments).

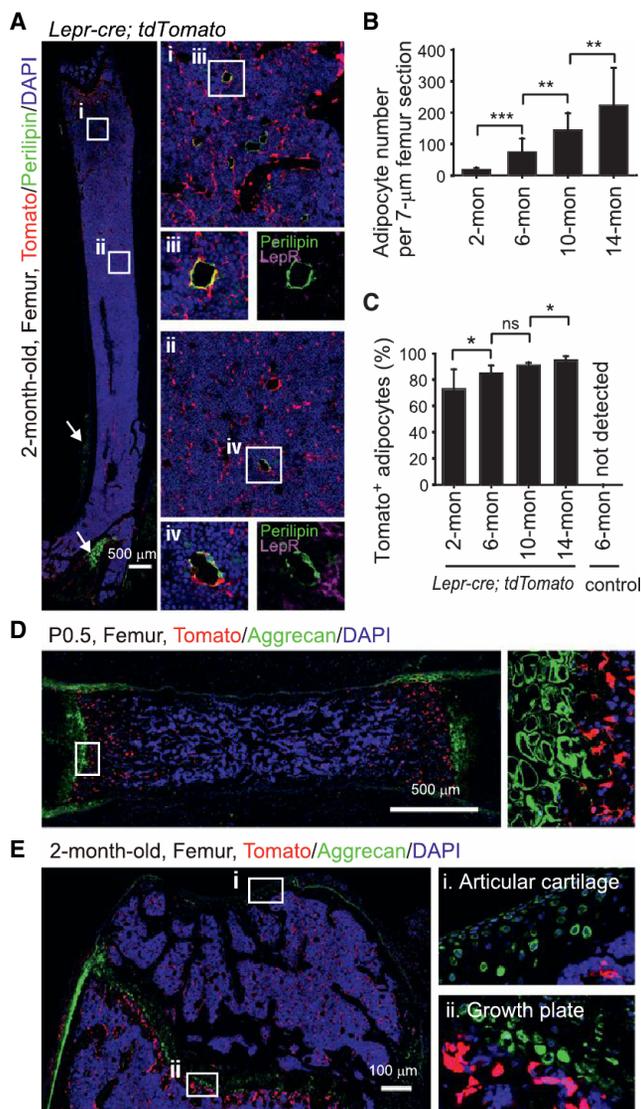
(E) Development of bone and hematopoiesis in ossicles formed by individual CFU-F colonies that arose from LepR<sup>+</sup> stromal cells from four *Col2.3-GFP; LepR-cre; tdTomato* mice.

(F–I) Representative femur sections from *LepR-cre; tdTomato; Col2.3-GFP* mice of different ages showing the increasing generation of Tomato<sup>+</sup>*Col2.3-GFP*<sup>+</sup> osteoblasts with age (three to five mice/age from at least four independent experiments).

(J) The frequency of Tomato<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup>*Col2.3-GFP*<sup>-</sup> bone marrow stromal cells in the femurs of *LepR-cre; tdTomato; Col2.3-GFP* mice did not change with age. Cells from ~6-month-old *Col2.3-GFP; tdTomato* mice were negative controls (CON). Data in all remaining panels represent mean ± SD from three to five mice/age from at least three independent experiments.

(K) Percentage of *Col2.3-GFP*<sup>+</sup> osteoblasts that were also Tomato<sup>+</sup> in enzymatically dissociated bone from *LepR-cre; tdTomato; Col2.3-GFP* mice of different ages. Osteoblasts from age-matched *Col2.3-GFP* or *Col2.3-GFP; tdTomato* mice were used as negative controls in each experiment (CON). Two-tailed Student's t tests were used to assess statistical significance among consecutive ages. \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001.

(L and M) Percentage of osteoblasts (L) and osteocytes (M) that were Tomato<sup>+</sup> in bone sections from *LepR-cre; tdTomato; Col2.3-GFP* mice.



**Figure 3. LepR<sup>+</sup> Cells Give Rise to Most Bone Marrow Adipocytes but Few Chondrocytes**

(A) Representative femur sections from a 2-month-old *Lepr-cre; tdTomato* mouse. Perilipin<sup>+</sup> adipocytes in the bone marrow did not stain with an anti-LepR antibody (purple) but were Tomato<sup>+</sup> (red). Periosteal adipocytes (arrows) were uniformly Tomato<sup>-</sup> (representative of four mice from four independent experiments).

(B and C) Quantification of adipocyte number per 7 μm femur section (B) and the percentage of adipocytes that were Tomato<sup>+</sup> at each age (C) in *Lepr-cre; tdTomato* mice. Two-tailed Student's *t* tests were used to assess statistical significance. ns, not significant; \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001. *n* = 3–5 mice/age from at least three independent experiments. Data represent mean ± SD. (D and E) Aggrecan<sup>+</sup> chondrocytes were not Tomato<sup>+</sup> in P0.5 (D) or 2-month-old (E) *Lepr-cre; tdTomato* mice (*n* = 3–5 mice/age from at least three independent experiments).

formed CFU-F colonies. However, none of the CFU-F colonies formed by unfractionated bone marrow cells from *Wnt1-cre; tdTomato* reporter mice were Tomato<sup>+</sup> (Figure 2A). Tomato<sup>+</sup> cells in the bone marrow of 2-month-old *Wnt1-cre; tdTomato* reporter mice accounted for only 0.0017% ± 0.001% of bone marrow

cells and were mostly negative for the MSC marker PDGFRα (Figure S4G). We aged *Wnt1-cre; tdTomato* mice for 5 months and observed rare Tomato<sup>+</sup> osteocytes (Figure S4H) but no Tomato<sup>+</sup> osteoblasts, indicating that these cells are not a significant source of bone-forming progenitors in young adult mice.

### LepR<sup>+</sup> Cells Are a Major Source of Bone in Adult Mice

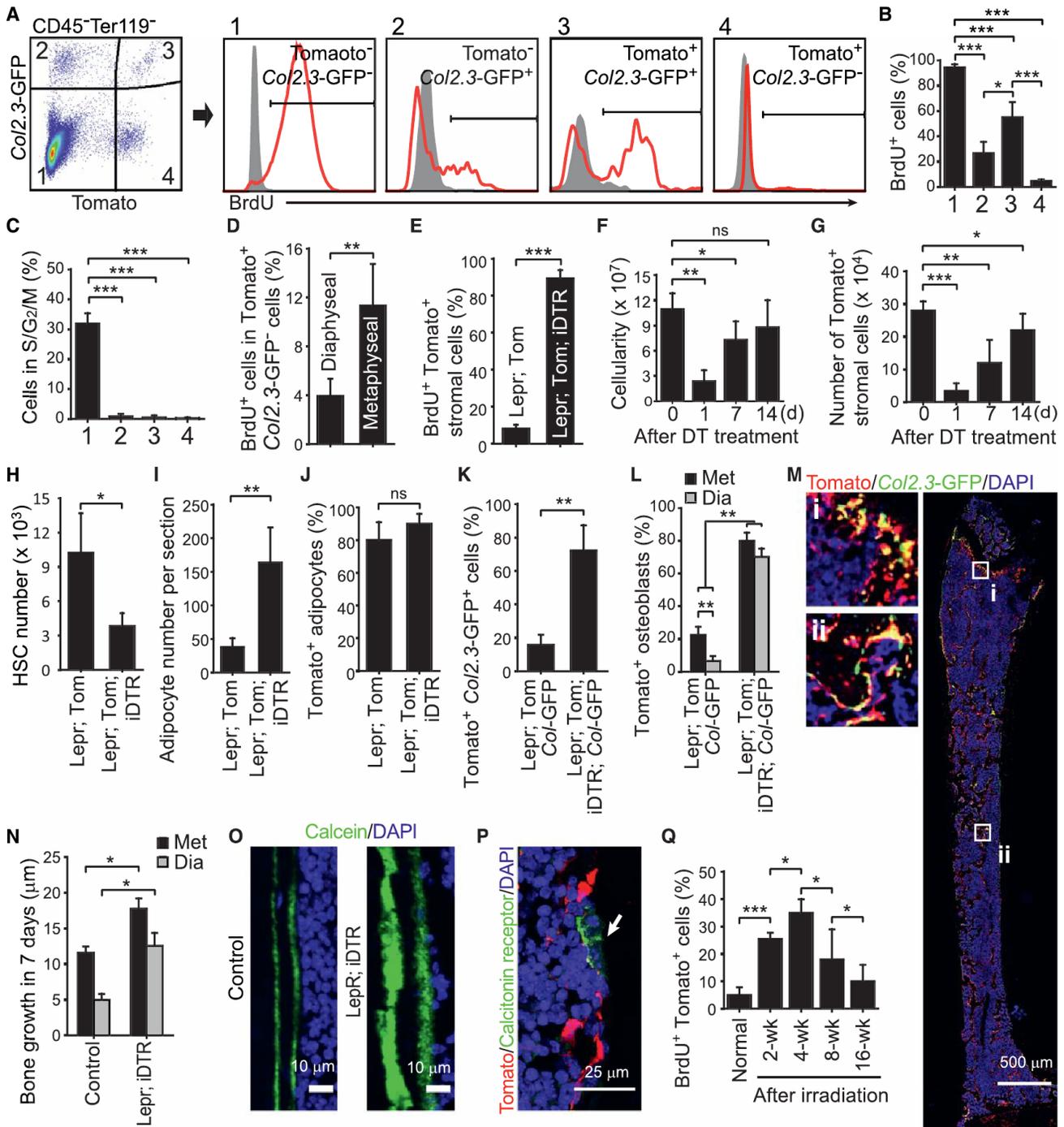
LepR<sup>+</sup> cells arose perinatally in bone marrow and made little bone before 2 months of age. We fate mapped the LepR<sup>+</sup> cells in vivo using *Lepr-cre; tdTomato; Col2.3-GFP* mice in which osteoblastic bone-lining cells can be unambiguously identified based on *Col2.3-GFP* expression (Kalajzic et al., 2002). At embryonic day (E) 19.5, Tomato<sup>+</sup> cells were rare in the bone marrow of *Lepr-cre; tdTomato; Col2.3-GFP* mice and we observed no contribution of these cells to bone (Figure 2F). By postnatal day (P) 0.5, there was a sharp increase in the number of Tomato<sup>+</sup> cells in metaphyseal bone marrow, though Tomato<sup>+</sup> cells remained rare in the diaphyseal bone marrow and only rare Tomato<sup>+</sup>Col2.3-GFP<sup>+</sup> osteoblasts were observed in trabecular bone (Figure 2G). By 2 months of age, Tomato<sup>+</sup> cells were visible throughout the bone marrow in both metaphysis and diaphysis, but Tomato<sup>+</sup>Col2.3-GFP<sup>+</sup> cells remained infrequent (Figure 2H), accounting for 3%–10% of Col2.3-GFP<sup>+</sup> cells in several bones (Figure 2K). However, the contribution of LepR<sup>+</sup> cells to bone increased sharply with age. Tomato<sup>+</sup> cells accounted for 10%–23% of Col2.3-GFP<sup>+</sup> cells in several bones at 6 months of age, 43%–67% of Col2.3-GFP<sup>+</sup> cells at 10 months of age, and 61%–81% of Col2.3-GFP<sup>+</sup> cells at 14 months of age (Figures 2K and S2R). The increased frequency of Tomato<sup>+</sup>Col2.3-GFP<sup>+</sup> cells was not caused by the induction of *Lepr* expression within Col2.3-GFP<sup>+</sup> osteoblastic cells (Figures S2S and S2T).

By 10 months of age, Tomato<sup>+</sup> cells contributed not only to bone-lining osteoblastic cells but also to osteocytes within the bone matrix (Figure 2Iiii). The percentage of osteocytes that were Tomato<sup>+</sup> increased significantly with age, but much more rapidly within trabecular bone (Figure 2M). Tomato<sup>+</sup> cells represented 92% ± 7% of osteocytes in trabecular bone and 13% ± 5% of osteocytes in cortical bone at 14 months of age.

### LepR<sup>+</sup> Cells Are a Major Source of Adipocytes in Adult Bone Marrow

Although LepR antibody did not detectably stain Perilipin<sup>+</sup> adipocytes (Figures 1J and 3A), most Perilipin<sup>+</sup> adipocytes in the bone marrow of 2- to 14-month-old *Lepr-cre; tdTomato* mice were Tomato<sup>+</sup> (Figures 3A–3C). The number of adipocytes in the bone marrow increased dramatically with age (Figure 3B). In contrast to those in the bone marrow, periosteal adipocytes (outside of the marrow cavity) were all negative for Tomato (Figure 3A), suggesting a distinct cellular origin. LepR<sup>+</sup> cells are thus the major source of adipocytes in adult bone marrow.

Chondrogenesis is active during fetal development but largely inactive throughout adulthood (Raghunath et al., 2005). In mice at P0.5 and at 2 months of age, we detected no Tomato expression among Aggrecan<sup>+</sup> chondrocytes in the femurs of *Lepr-cre; tdTomato* mice (Figure 3D) in spite of Tomato<sup>+</sup> osteoblasts and perivascular stromal cells adjacent to the growth plate (Figure 3E). In 6-, 10-, and 14-month-old mice, we remained unable to detect Tomato<sup>+</sup> chondrocytes in articular cartilage or growth plate cartilage associated with femurs and tibias (Figure S5A



**Figure 4. *Lepr*<sup>+</sup>*Col2.3-GFP*<sup>-</sup> Cells Are Quiescent under Normal Physiological Conditions in Adult Bone Marrow but Go into Cycle to Regenerate Bone after Injury**

(A–C) BrdU incorporation (14 day pulse) (A and B) or Hoechst staining (C) by various stromal cell fractions from enzymatically dissociated bone and bone marrow obtained from 2-month-old *Lepr-cre; tdTomato; Col2.3-GFP* mice. Unless otherwise indicated, data in all remaining panels represent mean ± SD from three or four mice in three independent experiments, with statistical significance assessed by two-tailed Student's t tests. ns, not significant; \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001.

(D) Percentage of *Tomato*<sup>+</sup>*Col2.3-GFP*<sup>-</sup>*CD45*<sup>-</sup>*Ter119*<sup>-</sup>*CD31*<sup>-</sup> bone marrow stromal cells that incorporated a 14 day pulse of BrdU.

(E) Percentage of *Tomato*<sup>+</sup>*CD45*<sup>-</sup>*Ter119*<sup>-</sup>*CD31*<sup>-</sup> bone marrow stromal cells that incorporated a 14 day pulse of BrdU in femurs and tibias from *Lepr-cre; tdTomato* mice and *Lepr-cre; tdTomato; iDTR* mice 2 weeks after diphtheria toxin (DT) treatment.

(F) Bone marrow cellularity in the femurs and tibias of *Lepr-cre; tdTomato; iDTR* mice at the indicated time points after DT treatment.

(G) Number of *Tomato*<sup>+</sup>*CD45*<sup>-</sup>*Ter119*<sup>-</sup>*CD31*<sup>-</sup> stromal cells in bone marrow from *Lepr-cre; tdTomato; iDTR* mice after DT treatment.

(H) Number of *CD150*<sup>+</sup>*CD48*<sup>-</sup> *Lineage*<sup>-</sup> *Sca-1*<sup>+</sup> *c-kit*<sup>+</sup> HSCs in the femurs and tibias 7 days after DT treatment.

(legend continued on next page)

and data not shown). We did observe an increasing contribution with age of Tomato<sup>+</sup> cells to a layer of Aggrecan<sup>-</sup> cells on the cartilage surface that appeared to be part of the synovial membrane (Figures S5Ai and S5B). We do not know whether Tomato<sup>+</sup> synovial cells are lineally related to LepR<sup>+</sup> bone marrow stromal cells. Overall, we observed little contribution of LepR<sup>+</sup> cells to cartilage.

### LepR<sup>+</sup> Bone Marrow Stromal Cells Are Quiescent

We administered bromo-deoxyuridine (BrdU) to 2-month-old *Lepr-cre; tdTomato; Col2.3-GFP* mice for 14 days. We found that 94% ± 2.4% of Tomato<sup>-</sup>Col2.3-GFP<sup>-</sup> stromal cells were BrdU<sup>+</sup>, 27% ± 8.9% of Tomato<sup>-</sup>Col2.3-GFP<sup>+</sup> osteoblasts were BrdU<sup>+</sup>, and 55% ± 12% of Tomato<sup>+</sup>Col2.3-GFP<sup>+</sup> osteoblasts were BrdU<sup>+</sup>, but only 4.8% ± 1.1% of Tomato<sup>+</sup>Col2.3-GFP<sup>-</sup> stromal cells were BrdU<sup>+</sup> (Figures 4A and 4B). Consistent with this, only 0.23% ± 0.3% of Tomato<sup>+</sup>Col2.3-GFP<sup>-</sup> stromal cells had greater than 2N DNA content by Hoechst staining (Figure 4C). LepR<sup>+</sup> stromal cells are thus largely quiescent in normal adult bone marrow.

At 2 months of age, the LepR<sup>+</sup> cells gave rise to osteoblasts and adipocytes to a much greater extent in metaphyseal bone marrow than in diaphyseal bone marrow (Figures 2H, 2L, and 3A). Tomato<sup>+</sup>Col2.3-GFP<sup>-</sup> cells from the metaphysis incorporated a 14-day pulse of BrdU at nearly three times the rate of Tomato<sup>+</sup>Col2.3-GFP<sup>-</sup> cells from the diaphysis (Figure 4D). Thus, there is regional regulation of LepR<sup>+</sup> stromal cells in the bone marrow.

We ablated LepR<sup>+</sup> bone marrow cells by administering DT to *Lepr-cre; tdTomato; Rosa26-loxP-stop-loxP-iDTR* (inducible DT receptor) mice and littermate controls. In contrast to DT-treated controls (lacking *iDTR*), in which only 8.2% ± 2.0% of Tomato<sup>+</sup>Col2.3-GFP<sup>-</sup> bone marrow stromal cells incorporated a 14-day pulse of BrdU, 89% ± 4.5% of Tomato<sup>+</sup>CD45<sup>+</sup>Ter119<sup>-</sup>CD31<sup>-</sup> cells from *Lepr-cre; tdTomato; iDTR* mice incorporated BrdU (Figure 4E). Bone marrow cellularity and the number of Tomato<sup>+</sup> stromal cells in the bone marrow declined significantly 1 day after DT treatment but rebounded within 2 weeks (Figures 4F and 4G). Consistent with our demonstration that LepR<sup>+</sup> stromal cells are critical for the maintenance of quiescent HSCs (Ding and Morrison, 2013; Ding et al., 2012; Oguro et al., 2013), DT treatment depleted CD150<sup>+</sup>CD48<sup>-</sup>LSK HSCs (Figure 4H).

Within 14 days of DT treatment, adipogenesis (Figure 4J) and osteogenesis (Figure 4K) increased profoundly in the bone marrow. Trabecular bone began filling up the marrow cavity, including the diaphysis (Figure 4M). Virtually all of the bone marrow adipocytes as well as the new Col2.3-GFP<sup>+</sup> osteoblasts were Tomato<sup>+</sup> (Figures 4J–4L). The rate of bone formation, as

measured by calcein labeling, was significantly higher in *Lepr-cre; tdTomato; iDTR* mice as compared to controls in both the metaphysis and the diaphysis (Figures 4N and 4O). LepR<sup>+</sup> cells did not give rise to Calcitonin<sup>+</sup> osteoclasts (Figure 4P). The regeneration of LepR<sup>+</sup> cells after ablation is thus associated with increased adipogenesis and osteogenesis.

### LepR<sup>+</sup> Cells Are Activated by Irradiation to Form Osteoblasts and Adipocytes

We found that 25% ± 2.3% of the Tomato<sup>+</sup> stromal cells incorporated BrdU over a 2 week period after irradiation of *Lepr-cre; tdTomato* mice (Figure 4Q). In the next 2 weeks after irradiation, the percentage of Tomato<sup>+</sup> stromal cells that incorporated BrdU significantly increased to 35% ± 5.0% (Figure 4Q). At subsequent time points the percentage of Tomato<sup>+</sup> stromal cells that incorporated BrdU significantly declined (Figure 4Q). LepR<sup>+</sup> cells therefore divide transiently after irradiation. We observed a substantial increase in adipocyte frequency within the bone marrow after irradiation (Figure 5A) and 95% ± 3.0% of the Perilipin<sup>+</sup> adipocytes were Tomato<sup>+</sup> (Figure 5B). The frequency of osteoblasts undergoing cell death significantly increased 2 days after irradiation (Figure 5C). By 16 weeks after irradiation, most (65% ± 11%) osteoblasts derived from LepR<sup>+</sup> cells (Figure 5D and Figure S6A). Irradiation did not activate LepR expression in osteoblasts (Figures S6B and S6C). We remained unable to detect any contribution of LepR<sup>+</sup> cells to chondrocytes after irradiation (Figure S6D and data not shown).

### Regeneration of Fractured Bone by LepR<sup>+</sup> Cells

Two weeks after a break was created in the tibia of two-month-old *Lepr-cre; tdTomato; Col2.3-GFP* mice (Figure 5E), a substantial increase of Tomato<sup>+</sup> stromal cells was observed in the marrow cavity adjacent to the fracture site (Figure 5F). Col2.3-GFP<sup>+</sup> cells close to periosteum were mostly negative for Tomato expression, but much of the callus was composed of Tomato<sup>+</sup> cells (Figure 5F). Tomato<sup>+</sup> cells accounted for 46% ± 13% of Aggrecan<sup>+</sup> chondrocytes in the soft callus 2 weeks following the fracture (Figures 5I and 5J). At 2 and 8 weeks after the fracture, 45% ± 5% and 85% ± 10% of Col2.3-GFP<sup>+</sup> osteoblasts were Tomato<sup>+</sup> (Figures 5F–5H). In contrast, only 9%–14% of Col2.3-GFP<sup>+</sup> osteoblasts were Tomato<sup>+</sup> in the undamaged tibia of the same mice (Figure 5H). We confirmed by qRT-PCR that *Ob-Rb* mRNA was not expressed by newly generated osteoblasts at the tibia fracture site (Figure S6E). Most of the osteoblasts and osteocytes that persisted long-term at the fracture site thus arose from LepR<sup>+</sup> cells.

To investigate the contribution of LepR<sup>+</sup> cells to the repair of perforated cartilage, we performed a subchondral perforation in articular cartilage associated with the femur in the knee of

(I and J) Quantification of adipocyte number per 7 μm femur section (I) and the percentage of adipocytes that were Tomato<sup>+</sup> (J) 14 days after DT treatment.

(K) Percentage of Col2.3-GFP<sup>+</sup> osteoblasts that were also Tomato<sup>+</sup> in enzymatically dissociated bone 2 weeks after DT treatment.

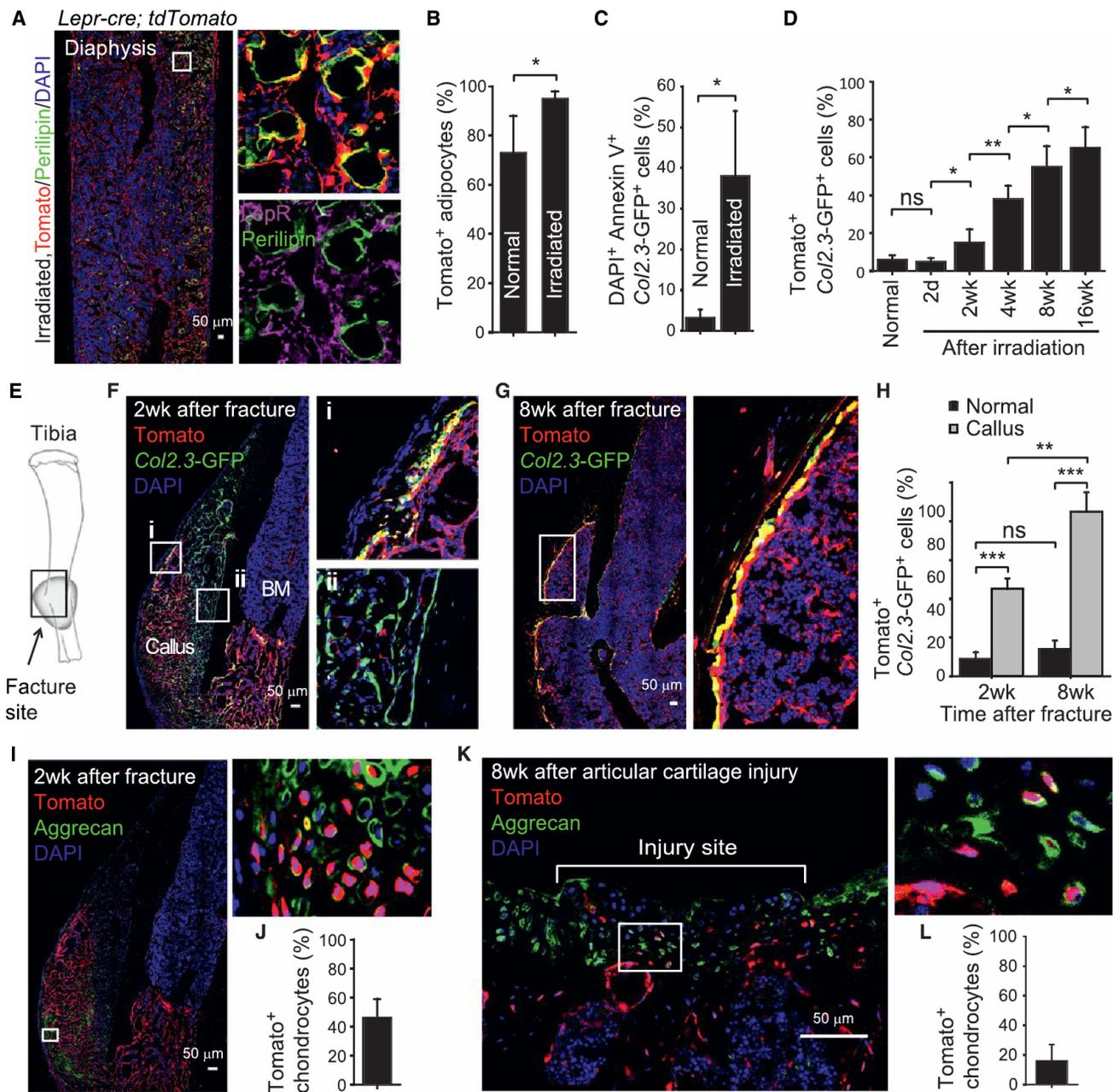
(L) Percentage of Col2.3-GFP<sup>+</sup> osteoblasts that were also Tomato<sup>+</sup> at metaphyseal (Met) and diaphyseal (Dia) bones 2 weeks after diphtheria toxin treatment.

(M) Representative femur sections from *Lepr-cre; tdTomato; Col2.3-GFP; iDTR* mice at 2 weeks after DT treatment. Note the formation of ectopic trabecular bone by Tomato<sup>+</sup>Col2.3-GFP<sup>+</sup> cells in the diaphyseal bone marrow cavity (Mii).

(N and O) Bone formation rate in *Lepr-cre; iDTR* and control mice after DT treatment. Two doses of calcein were injected at day 0 and 7 after DT treatment and the distance between calcein bands was measured at 14 days after DT treatment.

(P) Osteoclasts (arrow) were not labeled by Tomato in *Lepr-cre; tdTomato* mice.

(Q) Percentage of Tomato<sup>+</sup>CD45<sup>+</sup>Ter119<sup>-</sup>CD31<sup>-</sup> bone marrow stromal cells that incorporated a 14 day pulse of BrdU in *Lepr-cre; tdTomato* mice at various times after irradiation. n = 3–5 mice/time point from three independent experiments.



**Figure 5. *Lepr*<sup>+</sup> Cells Are the Major Source of New Osteoblasts and Adipocytes during Tissue Regeneration and Can also Form Chondrocytes after Subchondral Perforation**

(A and B) Representative femur section from a 2-month-old *Lepr-cre; tdTomato* mouse 14 days after lethal irradiation and transplantation of wild-type bone marrow cells. Perilipin<sup>+</sup> adipocytes in the bone marrow did not stain with an anti-*Lepr* antibody (purple) but were Tomato<sup>+</sup> (red), demonstrating that they derived from endogenous radio-resistant *Lepr*<sup>+</sup> cells (mean  $\pm$  SD from five mice in four independent experiments).

(C) Percentage of *Col2.3-GFP*<sup>+</sup> osteoblasts that were also DAPI<sup>+</sup> Annexin V<sup>+</sup> in enzymatically dissociated bone 2 days after irradiation. Data in all remaining panels represent mean  $\pm$  SD from three or four mice (per time point) in three independent experiments.

(D) Percentage of *Col2.3-GFP*<sup>+</sup> osteoblasts that were also Tomato<sup>+</sup> in enzymatically dissociated bone from *Lepr-cre; tdTomato; Col2.3-GFP* mice at various time points after irradiation.

(E) Schematic of experimental fracture site. The black rectangle depicts the region shown in images in (G), (H), and (J).

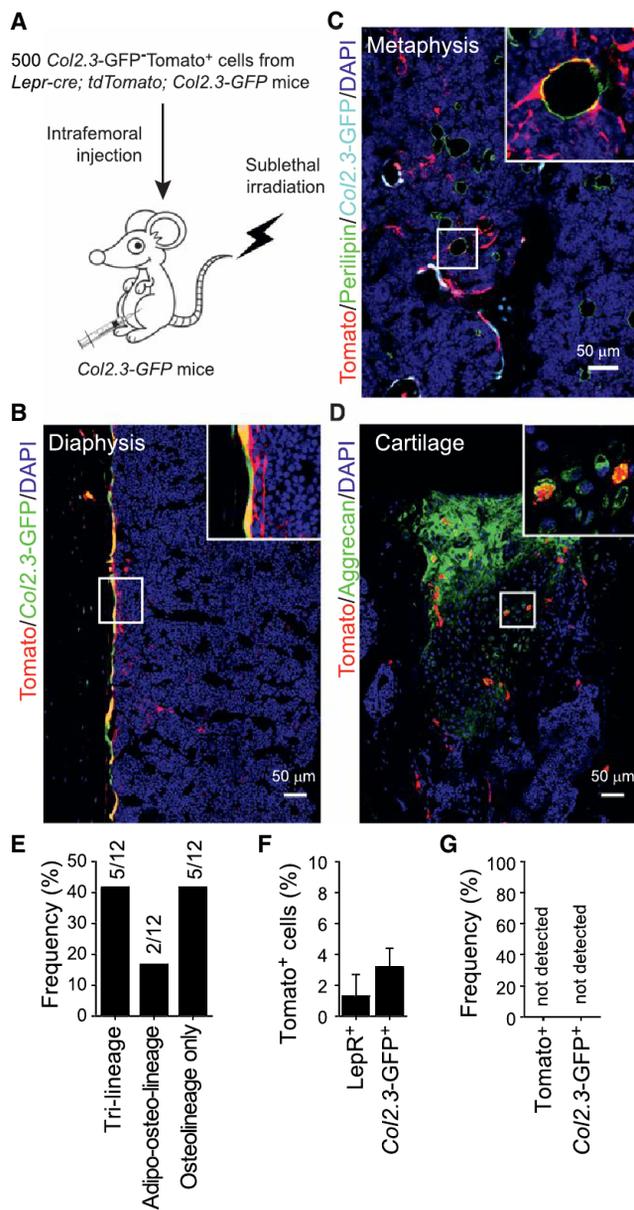
(F and G) Tomato expression by *Col2.3-GFP*<sup>+</sup> osteoblasts at the fracture site in *Lepr-cre; tdTomato; Col2.3-GFP* mice at 2 weeks (F) or 8 weeks (G) after fracture.

(H) Percentage of *Col2.3-GFP*<sup>+</sup> osteoblasts that were also Tomato<sup>+</sup> in unfractured tibias from control mice (normal) and bone callus from *Lepr-cre; tdTomato; Col2.3-GFP* mice.

(I and J) Percentage of Aggrecan<sup>+</sup> chondrocytes that were also Tomato<sup>+</sup> at the fracture site 2 weeks after the fracture.

(K and L) Percentage of Aggrecan<sup>+</sup> chondrocytes that were also Tomato<sup>+</sup> 8 weeks after subchondral perforation of articular cartilage in *Lepr-cre; tdTomato* mice.

Statistical significance was always assessed using two-tailed Student's *t* tests. ns, not significant; \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001.



**Figure 6. LepR<sup>+</sup> Cells Give Rise to Osteoblasts, Adipocytes, and Chondrocytes after Intrafemoral Transplantation**

(A) Experimental design.

(B–D) Representative femur sections from *Col2.3-GFP* mice transplanted with 500 *Tomato*<sup>+</sup>*Col2.3-GFP*<sup>−</sup> cells as described in (A) (n = 5). Note that the transplanted *Tomato*<sup>+</sup>*Col2.3-GFP*<sup>−</sup> cells gave rise to *Col2.3-GFP*<sup>+</sup> osteoblasts (B), *Perilipin*<sup>+</sup> adipocytes (C), and *Aggrecan*<sup>+</sup> cartilage cells (at the injection site, D).

(E) Fraction of recipient mice in which *Tomato*<sup>+</sup> cells were observed to contribute to each of the indicated mesenchymal lineages (n = 12 mice).

(F) The percentage of *LepR*<sup>+</sup> bone marrow stromal cells or *Col2.3-GFP*<sup>+</sup> osteoblasts that were also *Tomato*<sup>+</sup> (donor-derived) in the femurs of recipient mice (mean ± SD from three mice in three independent experiments).

(G) No *Tomato*<sup>+</sup>*LepR*<sup>+</sup> cells or *Tomato*<sup>+</sup>*Col2.3-GFP*<sup>+</sup> cells were observed in the femurs of mice transplanted with 10<sup>5</sup> nonhematopoietic *Col2.3-GFP*<sup>−</sup> *Tomato*<sup>−</sup> cells from *Lepr-cre; tdTomato; Col2.3-GFP* mice (n = 3).

*Lepr-cre; tdTomato* mice. Eight weeks following perforation, the injured region of the cartilage was covered by *Aggrecan*<sup>+</sup> fibrocartilaginous tissues (Figure 5K) and 16% ± 11% of the *Aggrecan*<sup>+</sup> cells were positive for *Tomato* expression (Figures 5K and 5L). Thus, *LepR*<sup>+</sup> cells form cartilage after injury even though they contribute little to the formation of cartilage during development.

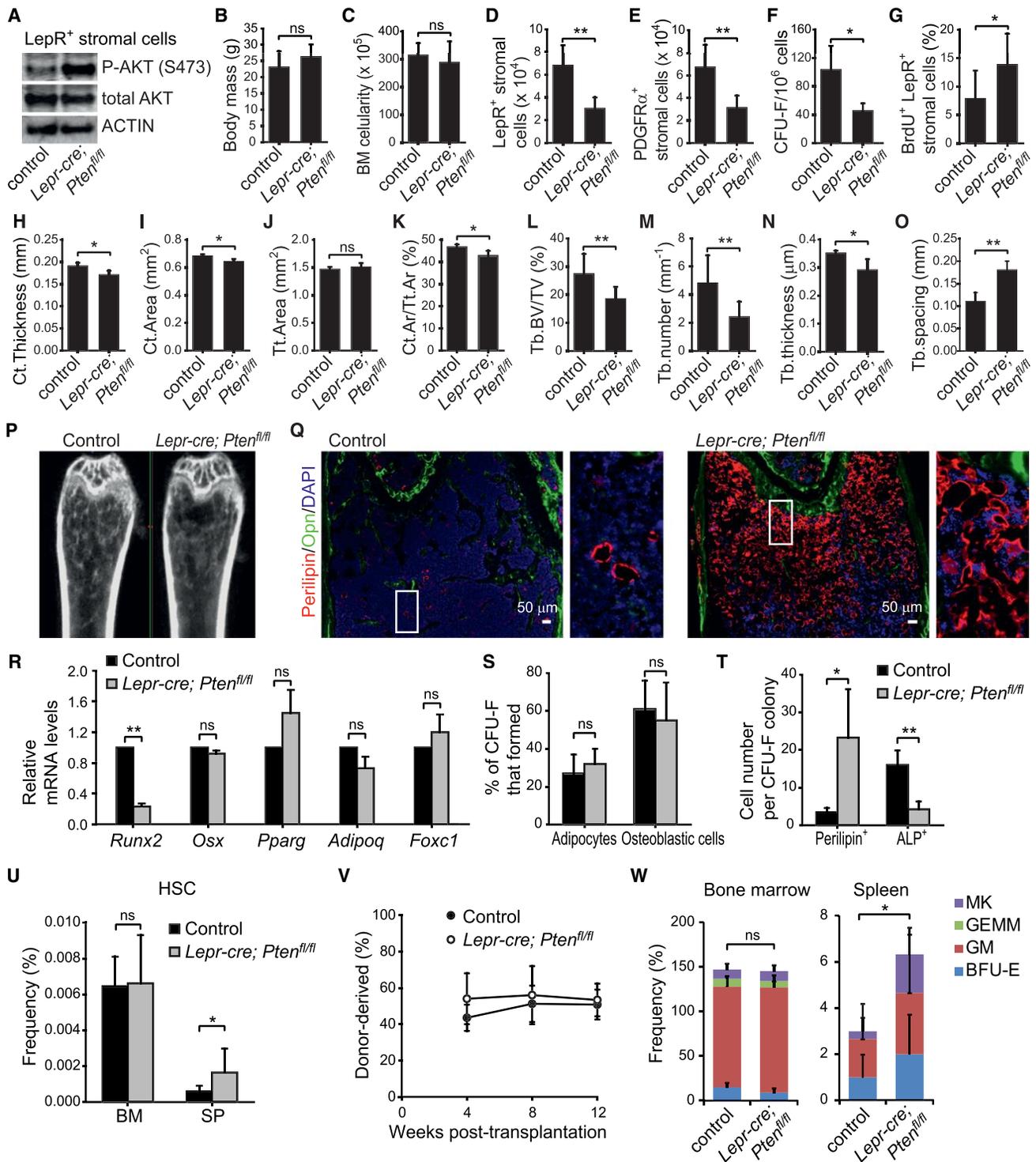
### LepR<sup>+</sup> Cells Form Bone, Cartilage, and Adipocytes after Transplantation

We transplanted 500 *Tomato*<sup>+</sup>*Col2.3-GFP*<sup>−</sup> cells from *Lepr-cre; tdTomato; Col2.3-GFP* mice into sublethally irradiated *Col2.3-GFP* mice by intrafemoral injection (Figure 6A). Four weeks later, *Tomato* expression was observed in *Col2.3-GFP*<sup>+</sup> osteoblasts (Figure 6B), *Perilipin*<sup>+</sup> adipocytes (Figure 6C), and *Aggrecan*<sup>+</sup> chondrocytes in articular cartilage (which had been perforated by the injection; Figure 6D). Of 12 recipient femurs, 5 contained trilineage reconstitution by transplanted *Tomato*<sup>+</sup> cells, 2 contained *Tomato*<sup>+</sup> osteolineage cells and adipocytes, but not cartilage, and 5 contained only *Tomato*<sup>+</sup> osteolineage cells (Figure 6E). In an independent experiment in which engraftment was analyzed by flow cytometry, *Tomato*<sup>+</sup> cells accounted for 1.3% ± 1.4% of *LepR*<sup>+</sup> cells and 3.2% ± 1.2% of *Col2.3-GFP*<sup>+</sup> cells in recipient femurs (Figure 6F). In contrast, transplantation of 10<sup>5</sup> *Tomato*<sup>−</sup>*Col2.3-GFP*<sup>−</sup>*CD45*<sup>−</sup>*Ter119*<sup>−</sup> stromal cells did not generate any *Tomato*<sup>+</sup> or *Col2.3-GFP*<sup>+</sup> cells (Figure 6G), suggesting that *LepR*<sup>−</sup> bone marrow cells have little capacity to form mesenchymal derivatives.

### PTEN Regulates Quiescence, Maintenance, and Differentiation of LepR<sup>+</sup> Cells

*Pten* is cell-autonomously required for the maintenance of HSCs (Kalaitzidis et al., 2012; Lee et al., 2010; Magee et al., 2012; Yilmaz et al., 2006; Zhang et al., 2006) and neural stem cells (Bonaguidi et al., 2011). To assess the consequences of *Pten* deletion from MSCs, we generated *Lepr-cre; Pten*<sup>fl/fl</sup> mice. As expected, AKT (S473) phosphorylation increased in *LepR*<sup>+</sup> stromal cells isolated from *Lepr-cre; Pten*<sup>fl/fl</sup> mice (Figure 7A). Body mass and bone marrow cellularity were normal in *Lepr-cre; Pten*<sup>fl/fl</sup> mice (Figures 7B and 7C). However, we observed a more than 2-fold reduction in the frequencies of *LepR*<sup>+</sup> *CD45*<sup>−</sup>*Ter119*<sup>−</sup>*CD31*<sup>−</sup> stromal cells (Figure 7D), *PDGFR* $\alpha$ <sup>+</sup> *CD45*<sup>−</sup>*Ter119*<sup>−</sup>*CD31*<sup>−</sup> stromal cells (Figure 7E), and CFU-Fs (Figure 7F) in *Lepr-cre; Pten*<sup>fl/fl</sup> mice relative to littermate controls. Based on BrdU incorporation, *LepR*<sup>+</sup> cells divided more frequently in *Lepr-cre; Pten*<sup>fl/fl</sup> mice relative to littermate controls (Figure 7G). *Pten* is thus cell-autonomously required to negatively regulate AKT activation and to maintain normal numbers of quiescent *LepR*<sup>+</sup> MSCs in adult bone marrow.

Four-month-old *Lepr-cre; Pten*<sup>fl/fl</sup> mice had a decreased volume of trabecular and cortical bone (Figures 7H–7P) but increased adipogenesis in the metaphysis relative to littermate controls (Figure 7Q). The expression of *Runx2*, a transcription factor required for osteogenesis (Komori et al., 1997; Otto et al., 1997), was markedly reduced in *LepR*<sup>+</sup> stromal cells from *Lepr-cre; Pten*<sup>fl/fl</sup> as compared to control mice (Figure 7R). *Pten* deletion did not affect the percentage of CFU-F colonies that contained adipocytes versus osteoblastic cells (Figure 7S), but it did significantly increase the number of



**Figure 7. PTEN Regulates Quiescence, Maintenance, and Differentiation in LepR<sup>+</sup> Stromal Cells**

(A) Representative western blots of flow cytometrically isolated LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> stromal cells from 4-month-old *Lepr-cre; Pten<sup>fl/fl</sup>* mice and littermate controls.

(B) Body mass. (Unless otherwise indicated, all remaining panels show mean  $\pm$  SD from three to five mice in three to four independent experiments, with statistical significance assessed by two-tailed Student's t tests; ns, not significant; \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001.)

(C) Bone marrow cellularity in the femurs of 4-month-old *Lepr-cre; Pten<sup>fl/fl</sup>* mice and littermate controls.

(D) Number of LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> cells.

(E) Number of PDGFR $\alpha$ <sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> cells.

(legend continued on next page)

Perilipin<sup>+</sup> adipocytes and significantly reduce the number of alkaline phosphatase<sup>+</sup> osteogenic cells that spontaneously differentiated within these colonies (Figure 7T). *Pten* is thus required by LepR<sup>+</sup> cells to promote osteogenesis and to restrain adipogenesis.

We did not detect a decline in the frequency of HSCs (Figure 7U) or colony-forming hematopoietic progenitors (Figure 7W and Figures S7B–S7F) in the bone marrow of *Lepr-cre; Pten<sup>fl/fl</sup>* mice. Bone marrow cells from *Lepr-cre; Pten<sup>fl/fl</sup>* and control mice gave similar levels of donor cell reconstitution upon transplantation into irradiated mice (Figures 7V and S7A). Nonetheless, we did observe significant increases in the frequencies of HSCs (Figure 7U) and colony-forming hematopoietic progenitors (Figure 7W) in the spleen. These data suggest that *Pten* deletion from LepR<sup>+</sup> bone marrow stromal cells changed the bone marrow niche in a way that led to the mobilization of HSCs and colony-forming progenitors.

## DISCUSSION

Our data support the conclusion that MSCs are an important component of the HSC niche (Kunisaki et al., 2013; Méndez-Ferrer et al., 2010; Omatsu et al., 2010; Sacchetti et al., 2007) because LepR<sup>+</sup> stromal cells are a major source of HSC niche factors in addition to MSC activity. Nearly all *Scf*-GFP<sup>+</sup> bone marrow stromal cells and *Cxcl12*-DsRed<sup>high</sup> bone marrow stromal cells were LepR<sup>+</sup> (Figures 1C, 1E, S1E, and S1F). The high expression of *Cxcl12* (Figure S1E) and PDGFR $\alpha$  (Figures 1L and 1M) by LepR<sup>+</sup> stromal cells indicates that these cells overlap strongly with CAR cells, which are also found primarily around sinusoids throughout the bone marrow (Omatsu et al., 2010; Sugiyama et al., 2006). However, our data indicate that NG2-CreER-expressing cells are not a significant source of MSCs in the bone marrow. NG2-CreER-expressing cells were much more rare than *Scf*-GFP<sup>+</sup> cells or *Cxcl12*-DsRed<sup>+</sup> cells (Figures S1E, S1F, and S3A–S3C) and we observed little PDGFR $\alpha$  or *Scf*-GFP expression by these cells (Figures S3A–S3D). While Kunisaki et al. (2013) concluded that NG2<sup>+</sup>Nestin<sup>high</sup>LepR<sup>-</sup> peria-arteriolar cells express high levels of *Scf* and *Cxcl12*, their RNaseq analysis showed that the “Nestin<sup>high</sup>LepR<sup>-</sup>” cells they analyzed were negative for *Nestin* and positive for *Lepr* expression (see GSE48764 in the Gene Expression Omnibus, referenced by Kunisaki et al., 2013). Thus, the data from Kunisaki

et al. are consistent with our data indicating that cells with high levels of *Scf* and *Cxcl12* in the bone marrow are marked by LepR.

## EXPERIMENTAL PROCEDURES

### Mice

All mice were maintained in a C57BL/6 background, including *Lepr-cre* (DeFalco et al., 2001), *Lepr<sup>fl/fl</sup>* (Cohen et al., 2001), *Pten<sup>fl/fl</sup>* (Groszer et al., 2001), *Scf-GFP* (Ding et al., 2012), *Cxcl12*-DsRed (Ding and Morrison, 2013), *Ubc-creER* (Ruzankina et al., 2007), *Rosa26-CAG-loxp-stop-loxp-tdTomato* (Madisen et al., 2010), *Rosa26-loxp-stop-loxp-EYFP* (Srinivas et al., 2001), *Rosa26-loxp-stop-loxp-iDTR* (Buch et al., 2005), *NG2-creER<sup>TM</sup>* (Zhu et al., 2011), *Wnt-1-cre* (Daniellian and McMahon, 1996), *Nestin-GFP* (Mignone et al., 2004), *Nestin-creER* (Balordi and Fishell, 2007), *Mx-1-cre* (Kühn et al., 1995), and *Col2.3-GFP* (Kalajzic et al., 2002) mice. To induce CreER activity, male mice (>2-month-old) were injected with 1 mg tamoxifen (Sigma) daily for 5 consecutive days followed by being fed ab libidum with chow containing 400 mg/kg tamoxifen for at least 2 weeks. To induce *Mx-1*-Cre expression, 2-month-old mice were injected with 10  $\mu$ g poly-inosine:poly-cytosine (pIpC; Amersham)/20 g body mass every other day for 10 days. To treat mice with DT, we injected mice intraperitoneally with 100 ng of DT for 7 consecutive days. All mice were housed in the Animal Resource Center at the University of Texas Southwestern Medical Center (UTSW). All procedures were approved by the UTSW Institutional Animal Care and Use Committee.

For methods related to flow cytometry, bone sectioning, qPCR, culture assay conditions, ossicle formation, irradiation, cell cycle analysis, calcein labeling, micro CT, western analysis, bone fracturing, intrafemoral injection, and subchondral perforation, see the Supplemental Experimental Procedures.

## SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes seven figures and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2014.06.008>.

## AUTHOR CONTRIBUTIONS

B.O.Z. performed most of the experiments. R.Y. analyzed the hematopoietic phenotype of *Lepr-cre; Pten<sup>fl/fl</sup>* mice. M.M.M. obtained most of the 3D images of bone marrow. J.P. analyzed the bone formation rate of *Lepr-cre; iDTR* mice. B.O.Z. and S.J.M. designed and interpreted all experiments and wrote the manuscript.

## ACKNOWLEDGMENTS

S.J.M. is a Howard Hughes Medical Institute Investigator, the Mary McDermott Cook Chair in Pediatric Genetics, and the director of the Hamon Laboratory for Stem Cells and Cancer. This work was supported by the National Heart, Lung and Blood Institute (HL097760) and the Cancer Prevention and Research

(F) CFU-Fs per million bone marrow cells.

(G) Percentage of LepR<sup>+</sup>CD45<sup>-</sup>Ter119<sup>-</sup>CD31<sup>-</sup> bone marrow stromal cells that incorporated a 14 day pulse of BrdU.

(H–O) MicroCT measurement of cortical (Ct) thickness (H), cortical area (I), total (Tt) area (J), cortical area/total area (Ct.Ar/Tt.Ar) (K), trabecular (Tb) bone volume/total volume (BV/TV) (L), trabecular number (M), trabecular thickness (N), and trabecular spacing (O) of femurs.

(P) Representative microCT images of femurs.

(Q) Representative femur sections showing excessive adipogenesis at metaphyseal bone marrow in *Lepr-cre; Pten<sup>fl/fl</sup>* mice.

(R) qPCR analysis of transcript levels for genes associated with osteogenic or adipogenic differentiation in LepR<sup>+</sup> stromal cells. Transcript levels were normalized based on  $\beta$ -actin amplification then set to 1 in control samples for comparison purposes.

(S) The percentage of CFU-F colonies that contained adipocytes (Oil red O<sup>+</sup>) and/or osteoblastic cells (Alizarin red S<sup>+</sup>) after culture in differentiation medium for 3 weeks.

(T) The average number of Perilipin<sup>+</sup> adipocytes or Alkaline phosphatase<sup>+</sup> (ALP<sup>+</sup>) osteogenic cells that spontaneously differentiated per CFU-F colony after culture for 1 week in standard medium.

(U) Number of CD150<sup>+</sup>CD48<sup>-</sup>Sca-1<sup>+</sup>c-kit<sup>+</sup> HSCs in the bone marrow and spleen.

(V) Donor-cell engraftment when  $3 \times 10^5$  donor bone marrow cells were transplanted along with  $3 \times 10^5$  recipient bone marrow cells into irradiated recipient mice (n = 11–12 recipient mice per genotype in three experiments).

(W) Frequencies of myeloerythroid-colony-forming progenitors in the bone marrow and spleen.

Institute of Texas. B.O.Z. was supported by a fellowship from the Leukemia and Lymphoma Society. R.Y. was supported by a fellowship from the Damon Runyon Foundation. J.G.P. was supported by a fellowship from the National Science Foundation. We thank Dr. Hung Luu in Children's Medical Center for his expert analysis of hematopoiesis in ossicles and Dr. Aktar Ali in the Mouse Metabolic Phenotyping Core at UT Southwestern for microCT analysis of bone.

Received: February 27, 2014

Revised: May 18, 2014

Accepted: June 6, 2014

Published: June 19, 2014

## REFERENCES

- Balordi, F., and Fishell, G. (2007). Hedgehog signaling in the subventricular zone is required for both the maintenance of stem cells and the migration of newborn neurons. *J. Neurosci.* *27*, 5936–5947.
- Bianco, P., Robey, P.G., and Simmons, P.J. (2008). Mesenchymal stem cells: revisiting history, concepts, and assays. *Cell Stem Cell* *2*, 313–319.
- Bonaguidi, M.A., Wheeler, M.A., Shapiro, J.S., Stadel, R.P., Sun, G.J., Ming, G.L., and Song, H. (2011). In vivo clonal analysis reveals self-renewing and multipotent adult neural stem cell characteristics. *Cell* *145*, 1142–1155.
- Buch, T., Heppner, F.L., Tertilt, C., Heinen, T.J., Kremer, M., Wunderlich, F.T., Jung, S., and Waisman, A. (2005). A Cre-inducible diphtheria toxin receptor mediates cell lineage ablation after toxin administration. *Nat. Methods* *2*, 419–426.
- Chai, Y., Jiang, X., Ito, Y., Bringas, P., Jr., Han, J., Rowitch, D.H., Soriano, P., McMahon, A.P., and Sucov, H.M. (2000). Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development* *127*, 1671–1679.
- Chan, C.K., Chen, C.C., Luppen, C.A., Kim, J.B., DeBoer, A.T., Wei, K., Helms, J.A., Kuo, C.J., Kraft, D.L., and Weissman, I.L. (2009). Endochondral ossification is required for haematopoietic stem-cell niche formation. *Nature* *457*, 490–494.
- Chou, D.B., Swarder, B., Bouladoux, N., Roy, C.N., Uchida, A.M., Grigg, M., Robey, P.G., and Belkaid, Y. (2012). Stromal-derived IL-6 alters the balance of myeloerythroid progenitors during *Toxoplasma gondii* infection. *J. Leukoc. Biol.* *92*, 123–131.
- Cohen, P., Zhao, C., Cai, X., Montez, J.M., Rohani, S.C., Feinstein, P., Mombaerts, P., and Friedman, J.M. (2001). Selective deletion of leptin receptor in neurons leads to obesity. *J. Clin. Invest.* *108*, 1113–1121.
- Crisan, M., Yap, S., Castella, L., Chen, C.W., Corselli, M., Park, T.S., Andriolo, G., Sun, B., Zheng, B., Zhang, L., et al. (2008). A perivascular origin for mesenchymal stem cells in multiple human organs. *Cell Stem Cell* *3*, 301–313.
- Danielian, P.S., and McMahon, A.P. (1996). Engrailed-1 as a target of the Wnt-1 signalling pathway in vertebrate midbrain development. *Nature* *383*, 332–334.
- DeFalco, J., Tomishima, M., Liu, H., Zhao, C., Cai, X., Marth, J.D., Enquist, L., and Friedman, J.M. (2001). Virus-assisted mapping of neural inputs to a feeding center in the hypothalamus. *Science* *291*, 2608–2613.
- Ding, L., and Morrison, S.J. (2013). Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* *495*, 231–235.
- Ding, L., Saunders, T.L., Enikolopov, G., and Morrison, S.J. (2012). Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature* *487*, 457–462.
- Echelard, Y., Vassileva, G., and McMahon, A.P. (1994). Cis-acting regulatory sequences governing Wnt-1 expression in the developing mouse CNS. *Development* *120*, 2213–2224.
- Friedenstein, A.J., Chailakhjan, R.K., and Lalykina, K.S. (1970). The development of fibroblast colonies in monolayer cultures of guinea-pig bone marrow and spleen cells. *Cell Tissue Kinet.* *3*, 393–403.
- Fukushi, J., Inatani, M., Yamaguchi, Y., and Stallcup, W.B. (2003). Expression of NG2 proteoglycan during endochondral and intramembranous ossification. *Dev. Dyn.* *228*, 143–148.
- Groszer, M., Erickson, R., Scripture-Adams, D.D., Lesche, R., Trumpp, A., Zack, J.A., Kornblum, H.I., Liu, X., and Wu, H. (2001). Negative regulation of neural stem/progenitor cell proliferation by the Pten tumor suppressor gene in vivo. *Science* *294*, 2186–2189.
- Horwitz, E.M., Le Blanc, K., Dominici, M., Mueller, I., Slaper-Cortenbach, I., Marini, F.C., Deans, R.J., Krause, D.S., and Keating, A.; International Society for Cellular Therapy (2005). Clarification of the nomenclature for MSC: The International Society for Cellular Therapy position statement. *Cytotherapy* *7*, 393–395.
- Joseph, N.M., Mukoyama, Y.S., Mosher, J.T., Jaegle, M., Crone, S.A., Dormand, E.L., Lee, K.F., Meijer, D., Anderson, D.J., and Morrison, S.J. (2004). Neural crest stem cells undergo multilineage differentiation in developing peripheral nerves to generate endoneurial fibroblasts in addition to Schwann cells. *Development* *131*, 5599–5612.
- Kalaitzidis, D., Sykes, S.M., Wang, Z., Punt, N., Tang, Y., Ragu, C., Sinha, A.U., Lane, S.W., Souza, A.L., Clish, C.B., et al. (2012). mTOR complex 1 plays critical roles in hematopoiesis and Pten-loss-evoked leukemogenesis. *Cell Stem Cell* *11*, 429–439.
- Kaljajic, Z., Liu, P., Kaljajic, I., Du, Z., Braut, A., Mina, M., Canalis, E., and Rowe, D.W. (2002). Directing the expression of a green fluorescent protein transgene in differentiated osteoblasts: comparison between rat type I collagen and rat osteocalcin promoters. *Bone* *31*, 654–660.
- Komada, Y., Yamane, T., Kadota, D., Isono, K., Takakura, N., Hayashi, S., and Yamazaki, H. (2012). Origins and properties of dental, thymic, and bone marrow mesenchymal cells and their stem cells. *PLoS ONE* *7*, e46436.
- Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R.T., Gao, Y.H., Inada, M., et al. (1997). Targeted disruption of *Cbfa1* results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* *89*, 755–764.
- Kühn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995). Inducible gene targeting in mice. *Science* *269*, 1427–1429.
- Kunisaki, Y., Bruns, I., Scheiermann, C., Ahmed, J., Pinho, S., Zhang, D., Mizoguchi, T., Wei, Q., Lucas, D., Ito, K., et al. (2013). Arteriolar niches maintain haematopoietic stem cell quiescence. *Nature* *502*, 637–643.
- Lee, J.Y., Nakada, D., Yilmaz, O.H., Tothova, Z., Joseph, N.M., Lim, M.S., Gilliland, D.G., and Morrison, S.J. (2010). mTOR activation induces tumor suppressors that inhibit leukemogenesis and deplete hematopoietic stem cells after Pten deletion. *Cell Stem Cell* *7*, 593–605.
- Liu, Y., Strecker, S., Wang, L., Kronenberg, M.S., Wang, W., Rowe, D.W., and Maye, P. (2013). Osterix-cre labeled progenitor cells contribute to the formation and maintenance of the bone marrow stroma. *PLoS ONE* *8*, e71318.
- Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., et al. (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* *13*, 133–140.
- Maes, C., Kobayashi, T., Selig, M.K., Torrekens, S., Roth, S.I., Mackem, S., Carmeliet, G., and Kronenberg, H.M. (2010). Osteoblast precursors, but not mature osteoblasts, move into developing and fractured bones along with invading blood vessels. *Dev. Cell* *19*, 329–344.
- Magee, J.A., Ikenoue, T., Nakada, D., Lee, J.Y., Guan, K.L., and Morrison, S.J. (2012). Temporal changes in PTEN and mTORC2 regulation of hematopoietic stem cell self-renewal and leukemia suppression. *Cell Stem Cell* *11*, 415–428.
- Méndez-Ferrer, S., Michurina, T.V., Ferraro, F., Mazloom, A.R., Macarthur, B.D., Lira, S.A., Scadden, D.T., Ma'ayan, A., Enikolopov, G.N., and Frenette, P.S. (2010). Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature* *466*, 829–834.
- Mignone, J.L., Kukekov, V., Chiang, A.S., Steindler, D., and Enikolopov, G. (2004). Neural stem and progenitor cells in nestin-GFP transgenic mice. *J. Comp. Neurol.* *469*, 311–324.
- Morikawa, S., Mabuchi, Y., Kubota, Y., Nagai, Y., Niibe, K., Hiratsu, E., Suzuki, S., Miyauchi-Hara, C., Nagoshi, N., Sunabori, T., et al. (2009). Prospective

- identification, isolation, and systemic transplantation of multipotent mesenchymal stem cells in murine bone marrow. *J. Exp. Med.* **206**, 2483–2496.
- Murfee, W.L., Skalak, T.C., and Peirce, S.M. (2005). Differential arterial/venous expression of NG2 proteoglycan in perivascular cells along microvessels: identifying a venule-specific phenotype. *Microcirculation* **12**, 151–160.
- Oguro, H., Ding, L., and Morrison, S.J. (2013). SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* **13**, 102–116.
- Omatsu, Y., Sugiyama, T., Kohara, H., Kondoh, G., Fujii, N., Kohno, K., and Nagasawa, T. (2010). The essential functions of adipo-osteogenic progenitors as the hematopoietic stem and progenitor cell niche. *Immunity* **33**, 387–399.
- Otto, F., Thornell, A.P., Crompton, T., Denzel, A., Gilmour, K.C., Rosewell, I.R., Stamp, G.W., Beddington, R.S., Mundlos, S., Olsen, B.R., et al. (1997). *Cbfa1*, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell* **89**, 765–771.
- Park, D., Spencer, J.A., Koh, B.I., Kobayashi, T., Fujisaki, J., Clemens, T.L., Lin, C.P., Kronenberg, H.M., and Scadden, D.T. (2012). Endogenous bone marrow MSCs are dynamic, fate-restricted participants in bone maintenance and regeneration. *Cell Stem Cell* **10**, 259–272.
- Pinho, S., Lacombe, J., Hanoun, M., Mizoguchi, T., Bruns, I., Kunisaki, Y., and Frenette, P.S. (2013). PDGFR $\alpha$  and CD51 mark human nestin+ sphere-forming mesenchymal stem cells capable of hematopoietic progenitor cell expansion. *J. Exp. Med.* **210**, 1351–1367.
- Pittenger, M.F., Mackay, A.M., Beck, S.C., Jaiswal, R.K., Douglas, R., Mosca, J.D., Moorman, M.A., Simonetti, D.W., Craig, S., and Marshak, D.R. (1999). Multilineage potential of adult human mesenchymal stem cells. *Science* **284**, 143–147.
- Raghunath, J., Salacinski, H.J., Sales, K.M., Butler, P.E., and Seifalian, A.M. (2005). Advancing cartilage tissue engineering: the application of stem cell technology. *Curr. Opin. Biotechnol.* **16**, 503–509.
- Rivers, L.E., Young, K.M., Rizzi, M., Jamen, F., Psachoulia, K., Wade, A., Kassarlis, N., and Richardson, W.D. (2008). PDGFRA/NG2 glia generate myelinating oligodendrocytes and piriform projection neurons in adult mice. *Nat. Neurosci.* **11**, 1392–1401.
- Ruzankina, Y., Pinzon-Guzman, C., Asare, A., Ong, T., Pontano, L., Cotsarelis, G., Zediak, V.P., Velez, M., Bhandoola, A., and Brown, E.J. (2007). Deletion of the developmentally essential gene *ATR* in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell* **1**, 113–126.
- Sacchetti, B., Funari, A., Michienzi, S., Di Cesare, S., Piersanti, S., Saggio, I., Tagliafico, E., Ferrari, S., Robey, P.G., Riminucci, M., and Bianco, P. (2007). Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment. *Cell* **131**, 324–336.
- Schneider, S., Bosse, F., D’Urso, D., Muller, H., Sereda, M.W., Nave, K., Niehaus, A., Kempf, T., Schnolzer, M., and Trotter, J. (2001). The AN2 protein is a novel marker for the Schwann cell lineage expressed by immature and nonmyelinating Schwann cells. *J. Neurosci.* **21**, 920–933.
- Schwab, K.E., and Gargett, C.E. (2007). Co-expression of two perivascular cell markers isolates mesenchymal stem-like cells from human endometrium. *Hum. Reprod.* **22**, 2903–2911.
- Srinivas, S., Watanabe, T., Lin, C.S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 4.
- Sugiyama, T., Kohara, H., Noda, M., and Nagasawa, T. (2006). Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity* **25**, 977–988.
- Yilmaz, O.H., Valdez, R., Theisen, B.K., Guo, W., Ferguson, D.O., Wu, H., and Morrison, S.J. (2006). Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* **441**, 475–482.
- Zhang, J., Grindley, J.C., Yin, T., Jayasinghe, S., He, X.C., Ross, J.T., Haug, J.S., Rupp, D., Porter-Westpfahl, K.S., Wiedemann, L.M., et al. (2006). PTEN maintains haematopoietic stem cells and acts in lineage choice and leukaemia prevention. *Nature* **441**, 518–522.
- Zhu, X., Hill, R.A., Dietrich, D., Komitova, M., Suzuki, R., and Nishiyama, A. (2011). Age-dependent fate and lineage restriction of single NG2 cells. *Development* **138**, 745–753.

# Conversion of Danger Signals into Cytokine Signals by Hematopoietic Stem and Progenitor Cells for Regulation of Stress-Induced Hematopoiesis

Jimmy L. Zhao,<sup>1,5,\*</sup> Chao Ma,<sup>2,5</sup> Ryan M. O'Connell,<sup>4</sup> Arnav Mehta,<sup>1</sup> Race DiLoreto,<sup>2</sup> James R. Heath,<sup>2,3,\*</sup> and David Baltimore<sup>1,\*</sup>

<sup>1</sup>Division of Biology and Biological Engineering, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA

<sup>2</sup>NanoSystems Biology Cancer Center, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA

<sup>3</sup>Kavli Nanoscience Institute, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA

<sup>4</sup>Department of Pathology, University of Utah, 15 North Medical Drive East, Salt Lake City, UT 84112, USA

<sup>5</sup>These authors contribute equally to this work

\*Correspondence: jzhao@caltech.edu (J.L.Z.), heath@caltech.edu (J.R.H.), baltimo@caltech.edu (D.B.)

<http://dx.doi.org/10.1016/j.stem.2014.01.007>

## SUMMARY

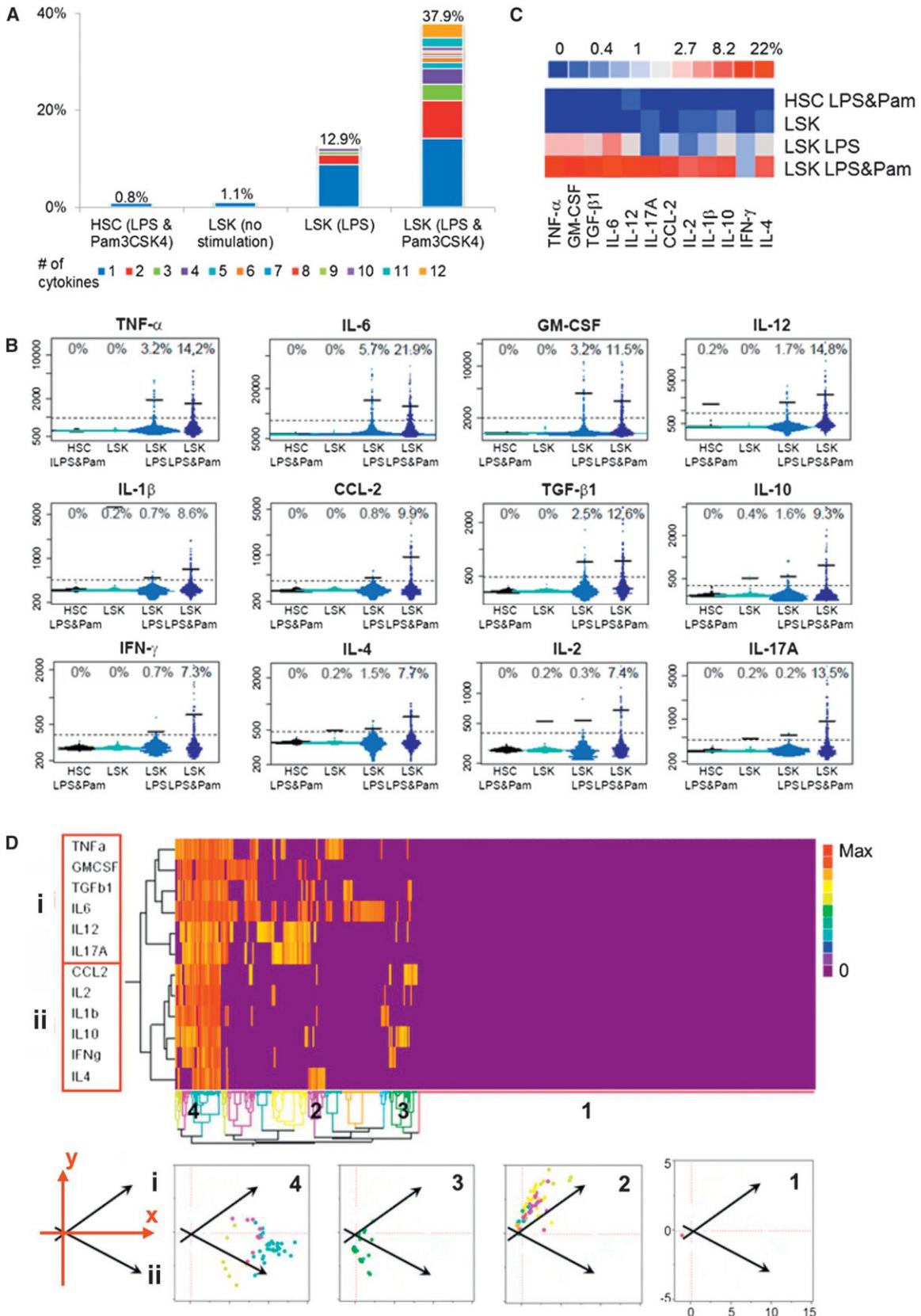
During an infection, the body increases the output of mature immune cells in order to fight off the pathogen. Despite convincing evidence that hematopoietic stem and progenitor cells (HSPCs) can sense pathogens directly, how this contributes to hematopoietic cell output remains unknown. Here, we have combined mouse models with a single-cell proteomics platform to show that, in response to Toll-like receptor stimulation, short-term HSCs and multipotent progenitor cells produce copious amounts of diverse cytokines through nuclear factor  $\kappa$ B (NF- $\kappa$ B) signaling. Interestingly, the cytokine production ability of HSPCs trumps mature immune cells in both magnitude and breadth. Among cytokines produced by HSPCs, IL-6 is a particularly important regulator of myeloid differentiation and HSPC proliferation in a paracrine manner and in mediating rapid myeloid cell recovery during neutropenia. This study has uncovered an important property of HSPCs that enables them to convert danger signals into versatile cytokine signals for the regulation of stress hematopoiesis.

## INTRODUCTION

Immune cells of the myeloid lineage are often considered the first responders of host defense against bacterial infection; meanwhile, hematopoietic stem and progenitor cells (HSPCs) may respond in a delayed fashion to ensure sufficient production of myeloid cells consumed during an infection. The response by HSPCs was originally thought to be of a passive response to the depletion of downstream immune cells, but more recent evidence suggests that HSPCs may participate directly by sensing systemically elevated cytokines through cytokine receptors and bacterial and viral components through Toll-like receptors (TLRs) (King and Goodell, 2011; Nagai et al., 2006).

It is well known that immune cells are potent cytokine producers upon encountering bacteria and viruses. When cytokines produced by immune cells and nonhematopoietic tissues accumulate to sufficient quantity, they circulate back to the bone marrow (BM) niche via blood circulation to activate HSPCs. Numerous cytokines, including IL-6, TNF- $\alpha$ , IFN- $\alpha$ , IFN- $\gamma$ , TGF- $\beta$ , and M-CSF, with the ability to regulate proliferation and differentiation of HSPCs have been identified (Baldrige et al., 2010, 2011; Challen et al., 2010; Essers et al., 2009; Maeda et al., 2009; Mossadegh-Keller et al., 2013; Pronk et al., 2011). On the other hand, it is clear now that HSPCs can also respond to TLR stimulation directly, leading to accelerated myeloid cell production in vitro (Nagai et al., 2006) and most likely in vivo as well (Megías et al., 2012). However, it remains unclear how direct pathogen sensing by HSPCs translates into signals directing myeloid differentiation under the stressed conditions. Conventional wisdom would suggest that TLR signaling activates lineage-specific transcriptional factors that can directly regulate differentiation within HSPCs. Currently, little is known about what transcription factors downstream of TLR activation might mediate this process. An alternative, but not mutually exclusive, hypothesis is that TLR stimulation activates a general proinflammatory program within HSPCs in order to induce cytokine production, which can act in an autocrine or paracrine manner to regulate differentiation.

In this study, we have combined extensive mouse genetics and a microfluidic single-cell proteomics platform to show that HSPCs can directly respond to bacterial components via the TLR/necrosis nuclear factor  $\kappa$ B (NF- $\kappa$ B) axis, and, in response, HSPCs, specifically short-term HSCs (ST-HSCs) and multipotent progenitor cells (MPPs), produce copious amounts of cytokines. In addition, single-cell analysis shows that HSPCs contain heterogeneous subsets based on their different cytokine production profiles. The cytokine production ability of HSPCs is shown to be regulated by NF- $\kappa$ B activity, given that p50-deficient HSPCs have significantly attenuated cytokine production, whereas miR-146a-deficient HSPCs display significantly enhanced cytokine production. Interestingly, HSPCs are significantly more potent cytokine producers in both breadth and quantity than the conventional known cytokine producers of the immune system, such as myeloid cells and lymphocytes.



(legend on next page)

Furthermore, we have shown that HSPCs possess TLRs, functional NF- $\kappa$ B signaling, and cytokine receptors—an entire cascade of molecules necessary for translating danger signals into cytokine signals. Lastly, we have demonstrated the functional significance of HSPC-produced cytokines, especially IL-6, in promoting myelopoiesis *in vitro* and *in vivo* in neutropenic mice after chemotherapeutic treatment or BM transplant. We believe that this represents a previously underappreciated mechanism by which HSPCs convert danger signals encountered during an infection into a range of versatile cytokine signals in order to ensure efficient stress-induced hematopoiesis. This circumvents both the delay associated with having to wait for systemic cytokine accumulation and the need to “reinvent” the molecular circuitry within HSPCs in order to convert TLR activation into specific differentiation signals.

## RESULTS

### Heterogeneity in Cytokine Production Profile among Purified HSPCs

To test whether any of the HSPC populations have the capability of cytokine production, we adapted a high-throughput, microfluidic-based technology to quantify a panel of up to 15 secreted proteins at the single-cell level (Ma et al., 2011). HSPCs are rare cells in BM, with LSK cells (defined as Lin<sup>-</sup>Sca1<sup>+</sup>cKit<sup>+</sup>), a mixed population of long-term HSCs (LT-HSCs), ST-HSCs, MPPs, and lymphoid-biased MPPs (LMPPs) accounting for less than 1% and LT-HSCs accounting for less than 0.1% of the total nucleated BM cells. This microfluidic platform allowed us to simultaneously measure a large number of secreted proteins at the single-cell level from thousands of phenotypically defined cells. After sorting LSK cells and LT-HSCs (defined as LSK CD150<sup>+</sup>CD48<sup>-</sup>) with stringent gating criteria and reanalyzing the sorted fraction for purity (Figure S1A available online), we loaded the cells onto chips containing several thousand microchambers. After 12 hr of incubation with culture medium containing either lipopolysaccharide (LPS), a TLR-4 ligand, alone or with Pam3CSK4, a TLR-2 ligand, secreted proteins were quantified by an ELISA-based method (Ma et al., 2011). At this time, nearly all LSK cells remained undifferentiated by fluorescence-activated cell sorting (FACS) analysis (Figure S2). Interestingly,

although very few stimulated LT-HSCs produced cytokines, a significant fraction of LSK cells produced a wide range of cytokines upon TLR stimulation (Figures 1A–1C and S3). Cytokine production by LSK cells was stimulation dependent, given that few LSK cells had detectable secretion in the absence of stimulation. Furthermore, although LPS alone was sufficient to stimulate 12.9% of LSK cells to produce cytokines, a combination of LPS and Pam3CSK4 boosted the percentage to 37.9%, and many more cells secreted multiple cytokines, suggesting an additive effect of simultaneously stimulating multiple TLRs (Figures 1A–1C). Among the 12 cytokines studied, IL-6 was the most prominently induced and was secreted by 21.9% of LPS and Pam3CSK4-stimulated LSK cells, whereas the production of the other cytokines ranged from 7%–15% (Figures 1B and 1C).

Among the 12 cytokines in the panel, most are produced by multiple immune cell types. However, some cytokines, such as IL-2, IL-4, IL-17, and IFN- $\gamma$  are mainly produced by lymphocytes, whereas IL-1 $\beta$ , IL-6, IL-12, TNF- $\alpha$ , and GM-CSF are more abundantly produced by cells of the myeloid lineage (Janeway et al., 2001). Interestingly, when unsupervised clustering analysis was performed on LSK cells, it identified two main cytokine clusters: a cytokine cluster (group i) including IL-6, TNF- $\alpha$ , IL-12, and GM-CSF that resembles the production profile of myeloid cells, and a cytokine cluster (group ii) including IL-2, IL-4, IL-10, and IFN- $\gamma$  that resembles the production profile of lymphocytes (Figure 1D). LSK cells were divided into multiple functional subsets differing in their ability to produce group i and group ii cytokines, ranging from nonproducers (Figure 1D, subset 1), group i producers (Figure 1D, subset 2), group ii producers (Figure 1D, subset 3), to superproducers of all cytokines (Figure 1D, subset 4). When the cells from the four subsets were plotted onto the 2D principal component plane with the two vectors representing the two cytokine groups, subset 2 cells fell into the group i direction, and subset 3 cells fell into the group ii direction, whereas subset 4 cells fell between the two vector directions (Figure 1D, bottom), demonstrating once again the differential cytokine secretion profiles among LSK cells.

LSK cells represent a heterogeneous population comprising LT-HSCs, ST-HSCs, MPPs, and LMPPs. They represent cells along a differentiation tree with successive loss of self-renewal ability while still retaining the full potential to replenish most, if

### Figure 1. Single-Cell Analysis of Cytokine Production by WT HSPCs in Response to TLR Stimulation

(A–D) Data from single-cell cytokine chip analysis. The four cell groups analyzed are HSCs with LPS and Pam3CSK4, LSK cells with medium only, LSK cells with LPS, and LSK cells with LPS and Pam3CSK4 stimulation for 12 hr. LSK is defined by Lin<sup>-</sup>Sca1<sup>+</sup>cKit<sup>+</sup>, and HSC is defined by LSK CD150<sup>+</sup>CD48<sup>-</sup>.

(A) Polyfunctionality and population-level statistics of the cell types analyzed. The percentage of cells secreting different number of cytokines is shown for each cell type studied. Different colors represent cells producing different number of cytokines from 1 to 12 (labeled with different colors); the number on top of each cell type shows the total percentage of cells secreting detectable amount of any of the 12 cytokines.

(B) Comparison of HSCs and LSK cells by individual cytokine. Each plot is composed of several thousand individual dots from several thousand single cells. The four cell groups arranged from left to right are HSCs under LPS and Pam3CSK4 stimulation (black), LSK cells with no stimulation (cyan), LSK cells with LPS stimulation only (light blue), and LSK cells with LPS and Pam3CSK4 stimulation (dark blue). The numbers on top represent the percentage of cytokine-producing cells identified by the gate (the dotted line), and the bars represent the mean intensity of only the cytokine-producing cells (average intensity of the cells above the dotted line).

(C) A summary heat map showing the percentage of HSCs and LSK cells that secrete any individual cytokine under different stimulations.

(D) Clustering analysis of WT LSK cells under LPS and Pam3CSK4 stimulation with results presented as heat maps and scatter plots. The data are analyzed by two-way hierarchical clustering that groups similar proteins and cells together. The grouping is shown by the tree structure for both proteins and cells. Each row represents an individual cytokine, and each column represents an individual cell. The result is presented by a heat map with colors representing the amount of cytokine secretion from purple (undetectable) to yellow (intermediate) to red (maximum). Two major groups of proteins (i and ii) and four main groups of cells (1 to 4) are identified. Each cell group is replotted onto the 2D principal component plane. This reduced space is the one that can explain the largest fraction (65%) of the information from the data. The arrows indicate the directions of the two cytokine functional groups aligned to the x and y axes.

See also Figures S1–S3.

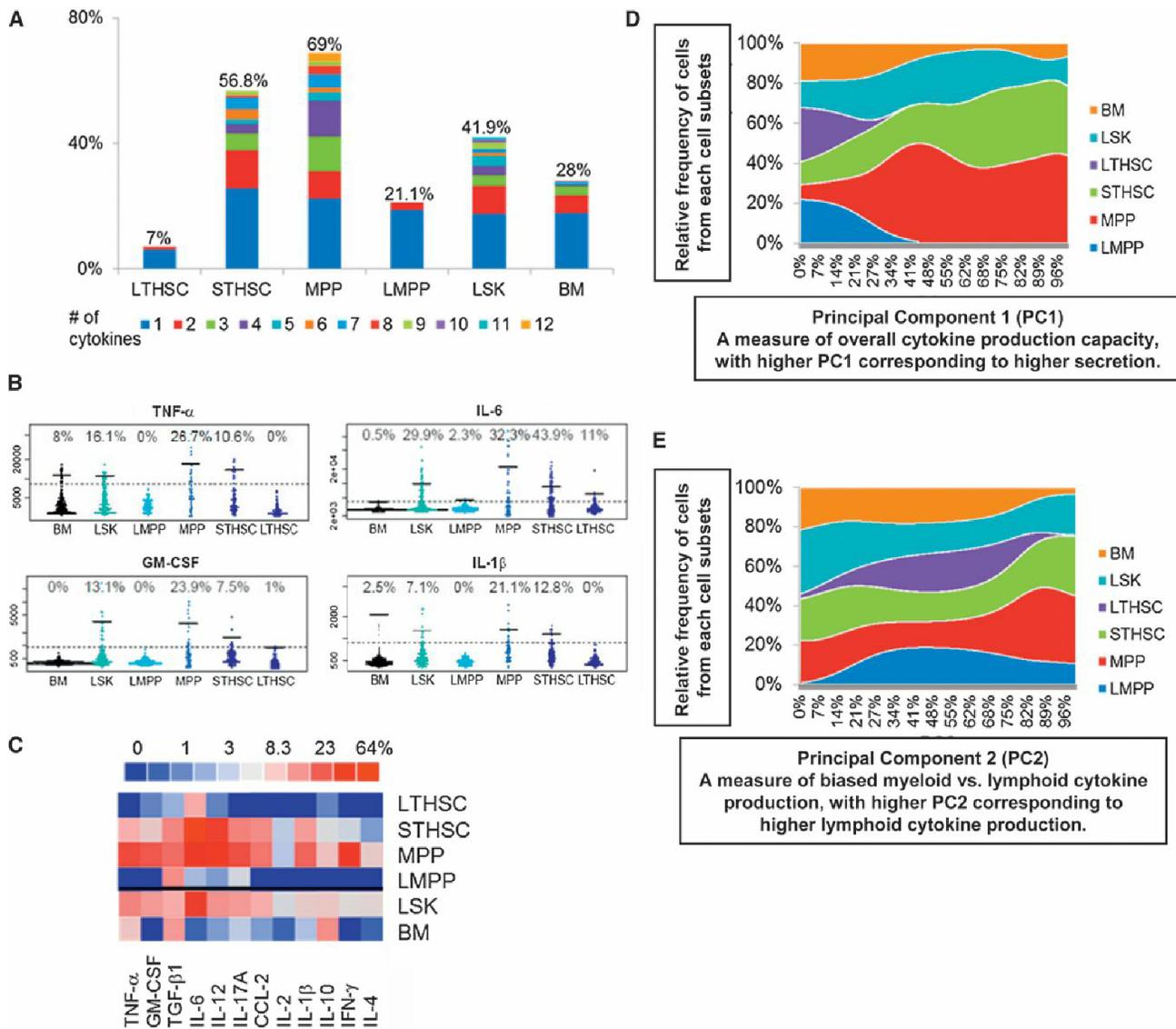
not all, hematopoietic cells (Adolfsson et al., 2005; Forsberg et al., 2006). Next, we asked whether the LSK heterogeneity based on cytokine production profile correlates with known HSPC subsets. To study this, we separately purified LT-HSCs (LSK Flt3<sup>-</sup>CD34<sup>-</sup>), ST-HSCs (LSK Flt3<sup>-</sup>CD34<sup>+</sup>), MPPs (LSK Flt3<sup>int</sup>CD34<sup>+</sup>), and LMPPs (LSK Flt3<sup>hi</sup>CD34<sup>+</sup>) from the LSK population and analyzed their cytokine secretion with the single-cell microfluidic chips (Figures 2 and S4). We confirmed that LT-HSCs lacked cytokine production with the use of a set of stem cell markers different from those used in Figure 1 because of controversy over what set of markers best defines HSCs. In comparison, ST-HSCs and MPPs were potent cytokine producers, whereas more differentiated LMPPs produced rather modest levels of cytokines (Figures 2A–2C and S4). Furthermore, similar to unfractionated LSK cells, ST-HSCs and MPPs remained heterogeneous in terms of cytokine production profile (Figure S4A). These results were succinctly summarized with principal component analysis in order to reduce the 12-dimensional cytokine intensity data set from six different BM subsets down to two principal components (Figures 2D and 2E). Principal component 1 (PC1) represents a measure of the overall cytokine secretion ability of a cell and is positively correlated with the intensity of each of the 12 cytokines (Figure S4C). When LT-HSCs, ST-HSCs, MPPs, and LMPPs along with unfractionated LSK cells and total BM were compared along PC1, a majority of LT-HSCs and LMPPs were clustered at the lower end of PC1, indicating very low overall cytokine secretions, whereas ST-HSCs and MPPs contained a large fraction of high cytokine secretors (Figure 2D). Principal component 2 (PC2) is a measure of cytokine secretion bias and is positively correlated with lymphoid cytokines (Figure S4D) but negatively correlated with myeloid cytokines (Figure S4E). When the six cell types were compared along PC2, LSK cells, ST-HSCs, and MPPs were heterogeneous with relatively even distribution across PC2, indicating that they contained a mixture of nonproducers, myeloid cytokine-biased, lymphoid cytokine-biased, and balanced producers. In comparison, unfractionated BM cells showed moderate myeloid-cytokine bias, and LMPPs showed moderate lymphoid-cytokine bias (Figure 2E). It is known that HSCs defined by current best cell-surface markers represent a heterogeneous population with biased myeloid or lymphoid lineage potential (Beerman et al., 2010; Challen et al., 2010; Copley et al., 2012). A recent study using a single-cell barcode technology also demonstrated the highly heterogeneous nature of the LMPP population (Naik et al., 2013). To add to the complexity, we have now shown that ST-HSCs and MPPs contain heterogeneous subsets based on their cytokine secretion profile. It is reasonable to speculate that the skewed cytokine production ability of ST-HSCs and MPPs may also reflect a biased lineage potential. Unfortunately, the current chip design does not allow us to recover the various cell subsets on the basis of their cytokine secretion profile for functional and lineage potential analysis. The possible relationship between the heterogeneous cytokine production ability and their lineage potential and functional capacity represents a future direction for research.

### Regulation of HSPC Cytokine Production by NF- $\kappa$ B

NF- $\kappa$ B is a family of transcription factors central to the regulation of inflammation and immune cell activation. Almost three de-

acades of research has elucidated many key physiological functions of NF- $\kappa$ B in innate and adaptive immune cells as well as its involvement in the pathogenesis of many immunological diseases (Baltimore, 2011). However, much less is known about the physiological function of NF- $\kappa$ B in HSPCs during inflammation. Because both LPS and Pam3CSK4 can activate NF- $\kappa$ B, which in turn can upregulate the transcription of many inflammatory cytokines in mature immune cells, we suspected that NF- $\kappa$ B might also regulate cytokine production in HSPCs. To test this, we first determined whether the important components of the TLR/NF- $\kappa$ B pathway were expressed at the protein level in HSPCs. Using a transgenic mouse that has a knockin of a RELA-GFP fusion protein at the endogenous locus (De Lorenzi et al., 2009), we found that all LT-HSCs and LSK cells were GFP<sup>+</sup>, indicating that all HSPCs express RELA (also known as p65) protein, a key subunit of NF- $\kappa$ B (Figure 3A). Then, we showed that both TLR-2 and TLR-4 receptors were expressed on the surface of LSK cells, and a subset of them express both receptors (Figure 3B), a finding consistent with previous reports (Nagai et al., 2006). Next, we asked whether HSPCs have functional NF- $\kappa$ B activity upon TLR stimulation. To study this, we took advantage of a different transgenic NF- $\kappa$ B-GFP reporter mouse, one with GFP expression under the control of the NF- $\kappa$ B regulatory elements (Magness et al., 2004). LPS stimulation *in vivo* upregulated GFP expression from 8% basally to 35%–40% in both LSK cells and LT-HSCs (Figures 3C and 3D). More directly, when purified LSK cells and LT-HSCs were stimulated with LPS and Pam3CSK4 *in vitro*, NF- $\kappa$ B activation was again evident by both the percent of GFP<sup>+</sup> cells and mean fluorescence intensity (Figures 3E and 3F). These results demonstrate that both LSK cells and LT-HSCs contain functional TLR/NF- $\kappa$ B signaling that can be directly activated by TLR-4 and TLR-2 ligands.

To determine whether NF- $\kappa$ B regulates cytokine production in LSK cells, we took advantage of mice deleted for genes important to the NF- $\kappa$ B pathway. NFKB1 (also known as p50) is one of the main subunits of the NF- $\kappa$ B family of transcription factors, and p50 deficiency results in defective NF- $\kappa$ B activity in various immune cells (Sha et al., 1995). In contrast, miR-146a is an important negative regulator of NF- $\kappa$ B activation by targeting its upstream signaling transducers TRAF6 and IRAK1. Thus, deficiency in miR-146a leads to enhanced NF- $\kappa$ B activity (Boldin et al., 2011; Zhao et al., 2011, 2013). When we subjected LSK cells from *Nfkb1*<sup>-/-</sup> mice (p50 knockout [KO]) and *Mir146a*<sup>-/-</sup> mice (miR KO) to the single-cell chip analysis, we found that production of all 12 cytokines was significantly attenuated in p50 KO LSK cells and enhanced in miR KO LSK cells in comparison to wild-type (WT) LSK cells (Figures 4A–4C and S3). Although 37.9% of WT LSK cells produced cytokines under costimulation, the percentage rose to an impressive 69.7% in miR KO and dropped to a mere 2.4% in p50 KO LSK cells (Figure 4A). Furthermore, the percentage of cytokine-producing LSK cells in *Mir146a*<sup>-/-</sup> *Nfkb1*<sup>-/-</sup> (miR/p50 double knockout [DKO]) mice was less than in WT mice, indicating that the lack of p50 is dominant over miR-146a deficiency (Figures 4A–4C and S3). In addition to the percentage of responding cells, the amount of cytokines produced on a per-cell basis was also enhanced in miR KO LSK cells and reduced in p50 LSK cells for all the proinflammatory cytokines (Figure 4B). These data demonstrate that cytokine production in LSK cells is regulated by the level of



**Figure 2. Single-Cell Analysis of Cytokine Production by Various HSPC Subsets in Response to TLR Stimulation**

(A–E) Data from single-cell cytokine chip analysis. The six cell groups analyzed are BM (total BM cells), LSK cells, LMPPs (LSK CD34<sup>+</sup>Fit3<sup>hi</sup>), MPPs (LSK CD34<sup>+</sup>Fit3<sup>int</sup>), ST-HSCs (LSK CD34<sup>+</sup>Fit3<sup>-</sup>), and LT-HSCs (LSK CD34<sup>-</sup>Fit3<sup>-</sup>), all under LPS and Pam3CSK4 stimulation for 12 hr.

(A) Polyfunctionality and population-level statistics of the cell types analyzed. The percentage of cells secreting different number of cytokines is shown for each cell type studied. Different colors represent cells producing different number of cytokines from 1 to 12 (labeled with different colors); the number on top of each cell type represents the total percentage of cells secreting detectable amount of any of the 12 cytokines.

(B) Comparison of HSPC subsets by individual cytokines. Four out of 12 cytokines are shown here, and the rest are shown in Figure S4B. Each plot is composed of several thousand individual dots from several thousand single cells. The six cell groups arranged from left to right are total BM, LSK cells, LMPPs, MPPs, ST-HSCs, and LT-HSCs under LPS and Pam3CSK4 stimulation. The numbers on top represent the percentage of cytokine-producing cells identified by the gate (the dotted line), and the bars represent the mean intensity of only the cytokine-producing cells (average intensity of the cells above the dotted line).

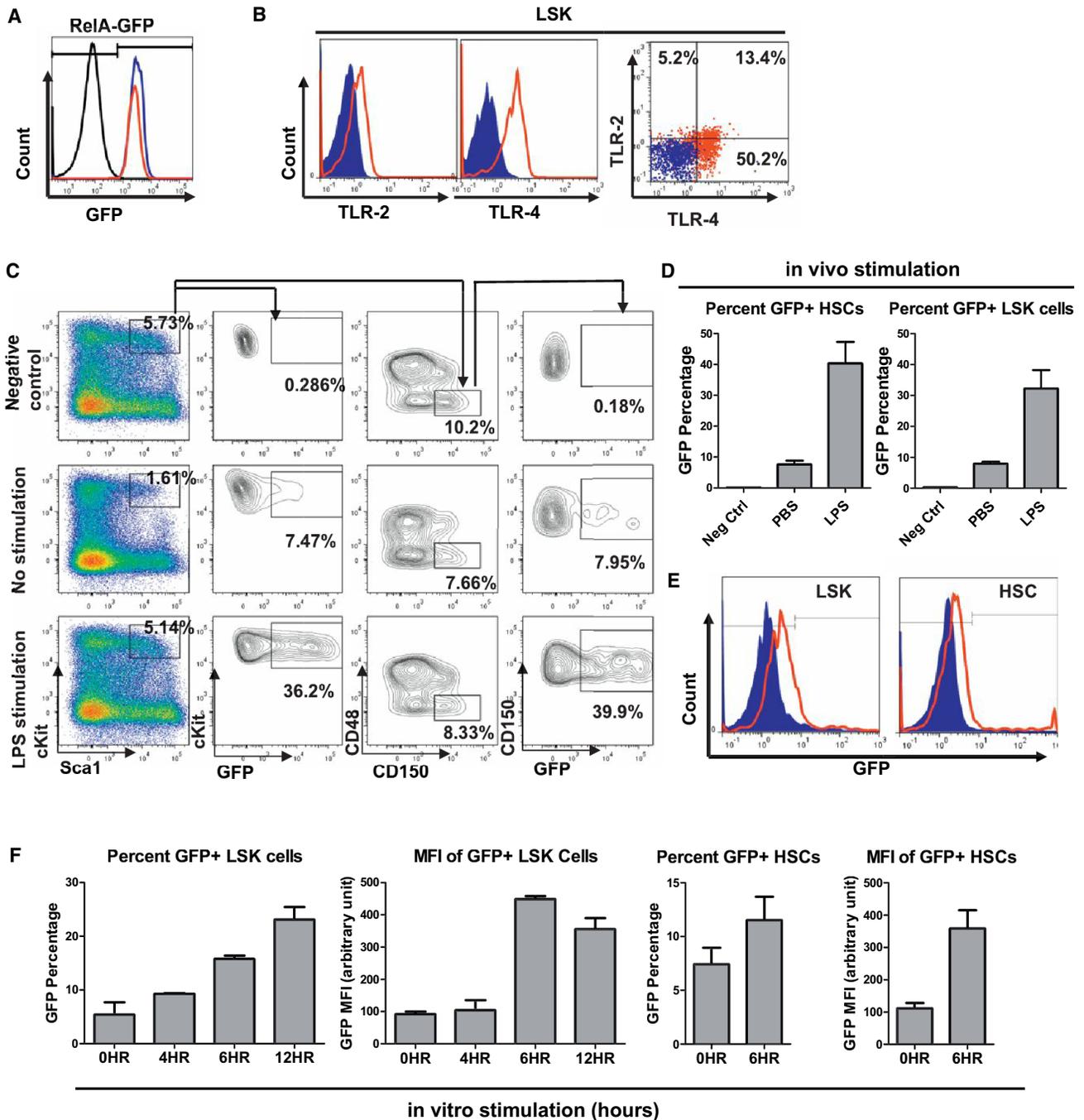
(C) A summary heat map showing the percentage of the six groups of cells that secrete any individual cytokine under different stimulations.

(D and E) Principal component analysis of all six cell groups that reduce a 12-dimensional cytokine data set from six different cell groups into two principal components (PC1 and PC2).

(D) PC1 represents the overall cytokine production capacity of a cell and is positively correlated to the overall intensity of all 12 cytokines (see Figure S4C). After all 12 cytokine intensities of each cell are converted into a single PC1 value, the relative frequency of cells with each PC1 value is calculated for all six cell subsets. Graphs show the relative frequency of cells (y axis) of the six cell groups against PC1 (x axis). Each cell type is represented by a different color.

(E) PC2 represents the level of biased cytokine production profile and is positively correlated to the lymphoid group of cytokines, including IL-2, IL-17a, IL-4, IL-12, IL-1 $\beta$ , and IFN- $\gamma$ , but is negatively correlated to the myeloid group of cytokines, including TNF- $\alpha$ , IL-6, and GM-CSF (see Figures S4D and S4E). Graphs show the relative frequency of cells (y axis) of the six cell groups against PC2 (x axis). Each cell type is represented by a different color.

See also Figure S4.



**Figure 3. NF-κB Activation through TLRs in HSPCs**

(A–F) Data from FACS analysis.

(A) Representative FACS histogram of LSK cells (blue) and HSCs (LSK CD150<sup>+</sup>CD48<sup>-</sup>; red) from RELA-GFP transgenic mice. LSK cells from a WT C57Bl/6 mice were used as a negative control (black).

(B) Representative FACS plots of TLR-2 and TLR-4 surface expression of WT LSK cells. Blue represents isotype antibody control, and red represents fluorescence-conjugated antibodies against TLR-2 (left), TLR-4 (middle), or both (right).

(C–F) Determination of TLR and NF-κB functionality in LSK cells and HSCs with mice in which GFP production is under NF-κB regulatory control.

(C and D) Experiments from in vivo stimulation.

(C) LPS-stimulated WT mouse is used as a negative control (top row) for GFP expression. No stimulation (middle row) corresponds to PBS-treated NF-κB-GFP transgenic reporter mice. LPS stimulation (bottom row) corresponds to LPS-stimulated NF-κB-GFP transgenic reporter mice (2 mg/kg body weight LPS for 6 hr).

(D) Quantification of GFP<sup>+</sup> percentages in LSK cells and HSCs from negative control, unstimulated, and LPS-stimulated NF-κB-GFP transgenic reporter mice (n = 3).

(E and F) Experiments from in vitro stimulation of purified LSK cells and HSCs.

(legend continued on next page)

NF- $\kappa$ B activity. Tuning the NF- $\kappa$ B activity up or down is sufficient to increase or decrease the cytokine production of LSK cells in response to TLR stimulation. We also measured the coproduction of any two cytokines at the single-cell level (Figures 4D and 4E). Correlation coefficients (R values) were calculated for all cytokine pairs for WT, p50 KO, and miR KO LSK cells, higher R values indicating tighter coregulation. In general, WT LSK cells showed higher correlation coefficients than either p50 KO or miR KO LSK cells, indicating a better coregulation of these cytokines (Figure 4E, green blocks). This may be functionally important, given that many of these cytokines are often produced together by a single cell; e.g., myeloid cells produce TNF- $\alpha$ , IL-6, and IL-12 during an inflammatory response, and CD4 T<sub>H1</sub> cells produce IL-2 and IFN- $\gamma$ . Imbalance in cytokine levels could lead to undesired effects and pathologies. When we performed a principal component analysis on LSK cells of all genotypes, a myeloid cytokine group including IL-6, TNF- $\alpha$ , IL-1 $\beta$ , and GM-CSF and a lymphoid cytokine group including IL-2, IL-4, and IL-17a were again identified (Figure 4F). Plotting the measurements onto the dominant 2D principal component space revealed diminished production of both groups of cytokines in p50 KO LSK cells and enhanced production in miR KO LSK cells. Interestingly, enhancement in the myeloid component was even more prominent in miR KO LSK cells, suggesting miR-146a-deficient LSK cells have especially enhanced myeloid cytokine production. Overall, these results demonstrate that TLR-stimulation-mediated cytokine production in HSPCs is exquisitely regulated by the level of NF- $\kappa$ B activity.

#### Functional Significance of HSPC-Produced Cytokines in Regulating Myelopoiesis In Vitro

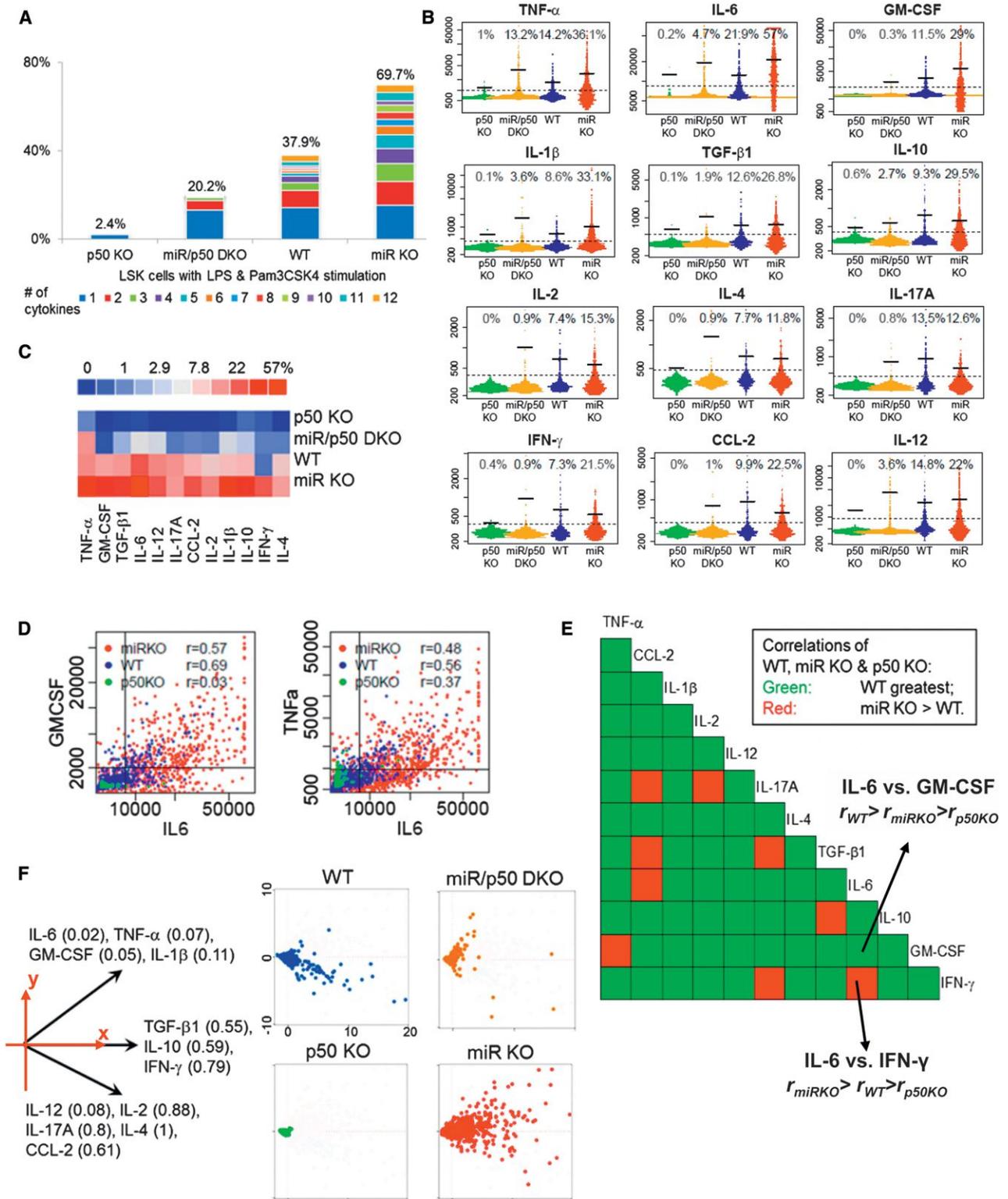
Cytokines produced by HSPCs are also produced by mature immune cells in the BM and periphery. However, we hypothesize that the location of HSPCs in the stem cell niche may represent an inherent advantage, allowing HSPC-produced cytokines to efficiently regulate their own fate in a more timely fashion during an infection or inflammatory challenge. To assess the functional significance of HSPC-produced cytokines, we first compared the cytokine production capacity of LSK cells with that of more differentiated progenitor cells and mature immune cells. To this end, purified LSK cells and lineage-committed progenitor cells (Lin<sup>-</sup>cKit<sup>+</sup>Sca1<sup>-</sup> and Lin<sup>-</sup>cKit<sup>-</sup>Sca1<sup>+</sup>) from BM as well as CD11b<sup>+</sup> myeloid, CD4<sup>+</sup> T, CD8<sup>+</sup> T, and B220<sup>+</sup> B cells from spleen were stimulated with LPS plus Pam3CSK4 for 24 hr. Then, culture medium was collected for multiplexed ELISA in order to measure a panel of 15 cytokines. Surprisingly, in the absence of survival and proliferative advantage after stimulation (Figure S5A), LSK cells produced far more cytokines in both quantity and breadth than mature myeloid and lymphoid cells in 24 hr with LPS and Pam3CSK4 stimulation. More impressively, even with much stronger stimuli, such as CpG, anti-CD3 and anti-CD28, or 100-fold more LPS, T, B, and myeloid cells were still significantly less potent cytokine producers than LSK cells (Figure 5A). In addition, LSK cells produced a wide range of myeloid and

lymphoid cytokines, whereas mature myeloid cells and lymphocytes showed narrower cytokine production profile.

Next, we determined whether the cytokines produced by HSPCs are able to influence hematopoiesis in vitro. First, we showed that a fraction of LT-HSCs, LSK cells, and myeloid progenitor cells expressed various cytokine receptors, including IL-6R $\alpha$ , IFN- $\gamma$ R, TNF-R1, and TNF-R2, on their surfaces (Figure S6A). This is consistent with previous studies that provided direct and indirect evidence of cytokine receptor expression in HSPCs (Baldrige et al., 2010, 2011; Maeda et al., 2009; Pronk et al., 2011). Next, we measured myeloid differentiation of LSK cells under LPS and Pam3CSK4 stimulation. Using an IL-6 neutralizing antibody, we were able to determine the effect of taking away IL-6 produced by LSK cells on myelopoiesis (Figures 5B and 5C). Interestingly, in comparison to the isotype antibody control, the neutralization of IL-6 had a significant effect on myeloid differentiation. Specifically, the percent of CD11b<sup>+</sup> cells produced from LSK cells decreased by about 50%, and the number of CD11b<sup>+</sup> cells showed a 2-fold reduction over a 4-day period (Figures 5B and 5C). To extend the study to include several other abundantly produced cytokines, we found that the neutralization of TNF- $\alpha$  or GM-CSF, but not IFN- $\gamma$ , also had an inhibitory effect on the generation of myeloid cells (Figures 5D and 5E). Interestingly, neutralization of IL-6, IFN- $\gamma$ , or GM-CSF all decreased the number of LSK cells, suggesting that these cytokines produced by LSK cells have a positive effect on their own proliferation and/or survival (Figure 5E). It is worth noting that we have previously shown that IL-6 can directly induce LSK cell proliferation by BrdU incorporation (Zhao et al., 2013). To further demonstrate the functional importance of HSPC-produced cytokines, we compared cytokines produced by BM cells that were depleted of HSPCs to that of total BM cells in inducing myelopoiesis. To this end, we stimulated equal number of total BM cells and Sca1-depleted BM cells with LPS and Pam3CSK4 in vitro for 24 hr and then used the conditioned media to stimulate LSK cells. We saw up to 40% reduction in the number CD11b<sup>+</sup> myeloid cells after only 2 days of incubation with conditioned medium from Sca1-depleted BM cells (Figure 5F). This suggests that HSPCs, accounting for less than 1% of total BM cells, make a significant contribution to the cytokine milieu and have a stimulatory effect on myelopoiesis. Instead of depleting cytokine, we next asked whether increased cytokine production by LSK cells could enhance myelopoiesis. To this end, we used miR-146a-deficient LSK cells that showed exaggerated cytokine production (Figure 4). Purified LSK cells from WT or miR KO mice were stimulated with LPS and Pam3CSK4 in vitro for 24 hr, and then the conditioned media were used to stimulate newly purified WT LSK cells. After 2 days, we saw a modest but consistent increase in numbers of both LSK cells and CD11b<sup>+</sup> cells when WT LSK cells were cultured with miR KO-LSK-cell-conditioned medium in comparison to WT-LSK-cell-conditioned medium (Figure 5G). This suggests that with just 24 hr of cytokine accumulation, the exaggerated cytokine production by miR KO LSK

(E) Representative FACS histograms of LSK cells and HSCs sorted from NF- $\kappa$ B-GFP transgenic reporter mice. Blue represents unstimulated LSK cells or HSCs and red represents cells stimulated in vitro with LPS (100 ng/ml) for 6 hr.

(F) Quantification of GFP<sup>+</sup> percentages and GFP mean fluorescence intensity of LSK cells at 0, 4, 6, and 12 hr and HSCs at 0 and 6 hr (n = 3). Data are presented as mean  $\pm$  SEM.



**Figure 4. Regulation of HSPC Cytokine Production by NF- $\kappa$ B**

(A–F) Data from single-cell cytokine chip analysis. LSK cells were sorted from WT, *Nfkb1*<sup>-/-</sup> (p50 KO), *Mir146a*<sup>-/-</sup> (miR KO), and *Nfkb1*<sup>-/-</sup>*Mir146a*<sup>-/-</sup> (miR/p50 DKO) mice and were stimulated with LPS and Pam3CSK4 for 12 hr.

(A) Polyfunctionality and population level statistics of LSK cells of different genetic models. The percentage of LSK cells secreting a different number of cytokines is shown for each cell type studied. Different colors represent cells producing different number of cytokines from 1 to 12 (labeled by different colors); the number on top of each cell type represents the total percentage of cells secreting detectable amount of any cytokine.

(legend continued on next page)

cells is sufficient to support enhanced myelopoiesis and LSK cell proliferation and/or survival. Consistent with this *in vitro* finding, miR KO mice exhibit a significant myeloproliferative disease after chronic inflammatory stimulation as a result of enhanced HSPC proliferation and myeloid differentiation, whereas deleting IL-6 in miR KO mice effectively ameliorates the myeloproliferative condition (Zhao et al., 2013). Complementing our previous reports that have shown the involvement of hyperactivated T cells and myeloid cells in miR-146a-deficiency-mediated pathologies (Zhao et al., 2011, 2013), this study shows that exaggerated cytokine production by miR-146a-deficient HSPCs is also a contributor to enhanced myelopoiesis in miR-146a KO mice.

### Functional Significance of HSPC-Produced Cytokines in Regulating Myelopoiesis *In Vivo*

In addition to *in vitro* functional significance, we next determined whether HSPC-produced cytokines are important in regulating hematopoiesis *in vivo*. Because current technology does not allow us to delete a cytokine specifically and selectively in stem and progenitor cells while allowing cytokine production in mature cells, to overcome this hurdle, we have created leucopenic conditions *in vivo* in which mature myeloid and lymphoid cells are severely depleted, whereas stem and progenitor cells are preserved. We created these conditions in mice by injecting a chemotherapeutic drug, 5-fluorouracil (5-FU) (Figures 6A–6C) or by transplanting stem and progenitor cells after lethal irradiation (Figures 6D–6F). LPS was injected into mice in order to stimulate myelopoiesis when neutropenia became the most severe, as measured by periodic sampling of peripheral blood. 5-FU is a chemotherapeutic drug that induces a significant reduction in mature cells in both BM and periphery while stimulating stem and progenitor cells to cycle (Harrison and Lerner, 1991). After 5-FU injection, myeloid cells started to decline, and, by day 5, 80% of myeloid cells were depleted in WT, miR-146a KO, IL-6 KO, and miR-146a/IL-6 DKO mice (Figure 6A). At this time, low-dose LPS was injected in order to stimulate myelopoiesis. Consistent with *in vitro* studies, IL-6 KO mice showed a 2-fold decrease in CD11b<sup>+</sup> and Gr1<sup>+</sup> myeloid cells, whereas miR KO mice showed a modest increase in comparison to WT mice (Figures 6B and 6C). In addition, DKO mice showed a similar level of reduction as IL-6 KO mice, indicating that loss of IL-6 is the dominant factor. To perform stem cell transplant, 500,000 purified Lin<sup>-</sup>cKit<sup>+</sup> cells from WT, miR KO, IL-6 KO, or miR/IL-6 DKO

mice were injected into lethally irradiated WT recipient mice. All four groups of mice showed severe leucopenia 6 days after irradiation and cell injection, and approximately 8,000 myeloid cells and 20,000 total white blood cells were left in the peripheral blood. This represented less than 1% of the normal level and 4% of the number of transplanted HSPCs (Figures 6D–6F). LPS stimulation promoted myelopoiesis that was mildly increased in mice receiving miR KO HSPCs but was significantly attenuated in mice receiving IL-6 KO or miR/IL-6 DKO HSPCs, indicating that IL-6 produced by HSPCs is an important factor in promoting myeloid cell recovery during neutropenia (Figures 6D–6F). Overall, these *in vivo* experiments show that, in mice with severe leucopenia, LPS-stimulated production of cytokines, especially IL-6, by endogenous or transplanted HSPCs has a significant positive impact on stress-induced myelopoiesis.

Within the BM niche, we speculate that the cytokines released by HSPCs upon TLR stimulation can act on themselves or neighboring HSPCs. To further delineate whether HSPC-produced cytokines mediate hematopoiesis predominantly through an autocrine or paracrine fashion, we injected NF- $\kappa$ B-GFP reporter mice with both LPS and BrdU in order to determine the relationship between cytokine-producing HSPCs and proliferating HSPCs. GFP expression, an indicator of NF- $\kappa$ B activation and cytokine production, and BrdU incorporation, a marker of proliferation, were costained in myeloid progenitor cells, LSK cells, and LT-HSCs at 4, 12, and 24 hr (Figure 6G). The result showed that LSK cells and LT-HSCs rapidly turned on NF- $\kappa$ B in response to LPS stimulation, whereas myeloid progenitor cells went into cycle quickly. As time elapsed, LT-HSCs and LSK cells started to proliferate, whereas NF- $\kappa$ B activity gradually dampened. Interestingly, throughout the stimulation, cells with NF- $\kappa$ B activity and cells that rapidly proliferated represented two largely nonoverlapping populations in all the stem and progenitor subsets. Although this result has multiple potential interpretations, the fact that we have been unable to capture a large fraction of cells that are simultaneously positive for BrdU and GFP throughout the 24 hr interval suggests to us the following mechanism: LPS-mediated NF- $\kappa$ B activation in HSPCs does not appear to directly turn on a proliferation program, and, instead, cytokines are induced that, in turn, act on cytokine receptors on neighboring HSPCs to stimulate proliferation and differentiation. The nature of this type of paracrine signaling and the distinction between the proliferative and the NF- $\kappa$ B-activated subsets within HSPCs require further investigation.

(B) Comparison of the cytokine secretion capacity of LSK cells from different genetic models. Each plot represents cytokine intensity scatter plots for a single protein with the four different mouse models arranged from left to right (p50 KO, green; miR/p50 DKO, yellow; WT, blue; miR KO, red). Each plot is composed of several thousand individual dots from several thousand single cells. The numbers on top represent the percentage of cytokine-producing cells above the gate (the dotted line), and the bars represent the mean intensity of only the positive cytokine-producing cells (average intensity of the cells above the dotted line).

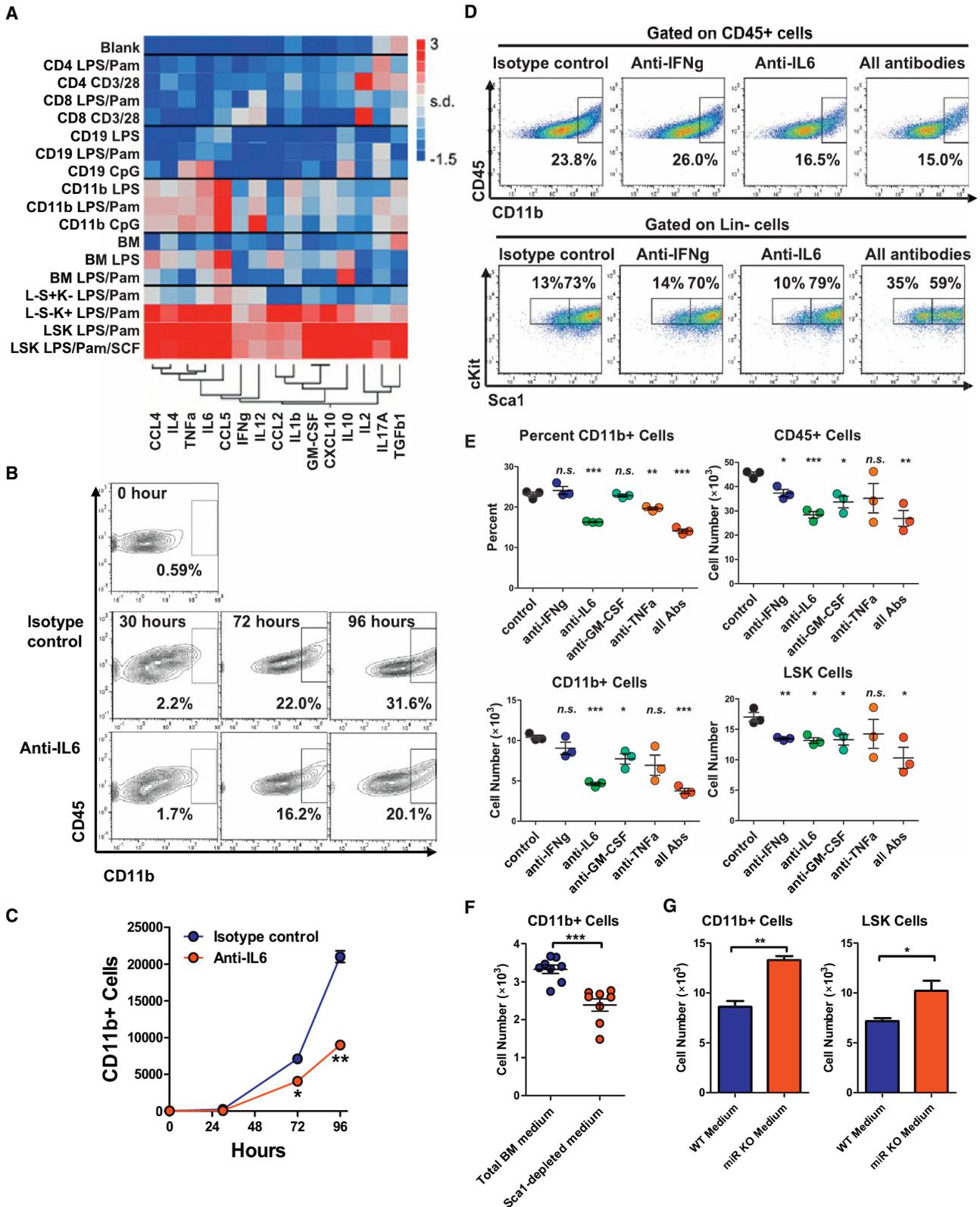
(C) A summary heat map showing the percentage of LSK cells that secrete any individual cytokine.

(D) 2D scatter plots showing protein pair correlations. Results from p50 KO (green), WT (blue), and miR KO (red) LSK cells are plotted. Correlation coefficients (R value) of the two cytokines for the different mouse models are shown on the plots.

(E) A half matrix summarizing the comparison of correlation coefficients (R values) of any given cytokine pairs between WT, p50 KO, and miR KO LSK cells. Each square block, intercepted by two cytokines with one on the top and one to the right, represents correlation coefficients between the two cytokines. Green represents the case when WT LSK has the highest correlation, such as IL-6 versus GM-CSF, and red represents when miR KO LSK has the highest correlation, such as IL-6 versus IFN- $\gamma$ . Correlations for all protein pairs from the p50 KO LSK cells are always lower than either WT or miR KO.

(F) Principal component analysis of LSK cells from the four genetic models. The data are plotted onto the 2D space by the top two principal components. This reduced space is the one that can explain the most (>60%) information of the data. The directions of the two protein groups are represented by the two vectors aligned to the x and y axes.

See also Figure S3.



**Figure 5. In Vitro Analysis of Functional Significance of HSPC-Produced Cytokines**

(A) Multiplexed ELISA quantification of cytokines in bulk cell-culture medium in order to compare cytokine production of different FACS-purified cell subsets stimulated for 24 hr with various stimulations. Cell types include LSK cells, L<sup>-</sup>S<sup>-</sup>K<sup>+</sup> (Lin<sup>-</sup>Sca1<sup>-</sup>cKit<sup>+</sup>), and L<sup>-</sup>S<sup>+</sup>K<sup>-</sup> (Lin<sup>-</sup>Sca1<sup>+</sup>cKit<sup>-</sup>) sorted from BMs and CD4<sup>+</sup>T, CD8<sup>+</sup>T, CD19<sup>+</sup>B, and CD11b<sup>+</sup> myeloid cells sorted from spleens. Stimulations include LPS/Pam (100 ng/ml LPS and 1  $\mu$ g/ml Pam3CSK4), LPS/Pam/stem cell (legend continued on next page)

## DISCUSSION

In this study, we have shown that ST-HSCs and MPPs can translate danger signals arising from an infection into cytokine signals that can directly regulate stress-induced hematopoiesis. Significantly, the cytokine production ability of HSPCs trumps both mature myeloid and lymphoid cells in terms of speed, magnitude, and breadth. In the BM stem cell niche, this property of HSPCs may play an important role in providing a rapid response time from the encounter of an infection to the output of myeloid cells. In addition to residing in the BM, HSPCs are known to egress from the marrow and traffic through the blood and lymphatic circulation to peripheral organs, including spleen, liver, lymph nodes, gut, and adipose tissues, where they may be exposed to a heavy burden of danger signals. These extramedullary sites may provide excellent opportunities for HSPCs to coordinate rapid stress-induced hematopoiesis upon TLR stimulation (Han et al., 2010; Jaiswal and Weissman, 2009; Massberg et al., 2007; Wesemann et al., 2013).

Given that all the cytokines produced by HSPCs are also produced by mature immune cells and some nonhematopoietic cells, isolating the effect of cytokines produced by HSPCs from mature cells *in vivo* is difficult, if not impossible, given current technical limitations in specifically deleting a gene in HSCs while turning it back on in mature cells. Therefore, we have relied on *in vitro* studies to demonstrate the functional importance of HSPC-produced cytokines in regulating myelopoiesis. In addition, there are situations when mature immune cells are significantly depleted, such as during sepsis, after chemotherapy, and during the initial recovering phase of stem cell transplant. During these situations, when mature immune cells are depleted and HSPCs are overrepresented, the ability of HSPCs to respond to stress directly in order to produce cytokines may become critical in mediating rapid hematopoietic cell recovery. We have created these situations in mice to mimic neutropenic conditions after chemotherapy or stem cell transplant and have shown that the ability of HSPCs to produce IL-6 is particularly important in mediating stress-induced myelopoiesis. In clinical practice, G-CSF and GM-CSF are used in certain situations after chemotherapy or stem cell transplant to stimulate neutrophil recovery in neutropenic patients who are susceptible to fatal infections (Bennett et al., 2013). A better understanding of the cytokine-mediated hematopoiesis in neutropenic conditions

will provide insight in the development of potentially more effective hematopoietic stimulating factors.

In addition to neutropenic conditions, we reason that this stem and progenitor cell property is also important under physiological conditions, despite of the rarity of HSPCs among differentiated hematopoietic cells. First, the speed, magnitude, and breadth of the cytokines produced by HSPCs in comparison to mature immune cells are impressive. Furthermore, this has not taken into the account the unique location of HSPCs in the stem cell niche. Location and organization of HSPCs within the BM niche have long been appreciated to play an important role in their self-renewal and proliferative properties (Shen and Nilsson, 2012). Perhaps groups of HSPCs residing in close proximity represent a unique advantage for rapid autocrine or paracrine-mediated hematopoiesis. Through BrdU incorporation in NF- $\kappa$ B-GFP reporter mice, we have shown that HSPCs with NF- $\kappa$ B activity and that are actively proliferating are largely two distinct populations. We speculate that a potential paracrine regulatory signaling may be at work within the BM that involves one subset of HSPCs responding to TLR stimulation by rapidly turning on NF- $\kappa$ B and producing copious amounts of cytokines and a neighboring cell population with cytokine receptors that can undergo rapid proliferation and differentiation in response to cytokine stimulation. In support of the notion, although some LSK cells express both TLR-4 and IL-6R $\alpha$ , it appears that a fraction of LSK cells express only the TLR or the cytokine receptor (Figure S6B), suggesting that HSPCs contain heterogeneous subsets with intrinsic differences in their ability to respond to TLR and cytokine stimuli.

The questions regarding the extent of involvement of HSPC-produced cytokines in hematopoiesis in a nonneutropenic host and the nature of HSPC heterogeneity on the basis of differential cytokine production or receptor expression remain open and will be an area of significant interest for future studies. In addition, we are also currently redesigning the microfluidic platform in order to recover individual cells from the chip after proteomic analysis for subsequent lineage and functional analysis. Furthermore, there are perhaps limitations in the *in vitro* on-chip stimulation, which may not provide the optimal conditions for culturing undifferentiated HSPCs. This may result in a reduction in the cells' functional robustness and an underestimate of the true percentage of cytokine-producing cells *in vivo*, despite our poststimulation analysis to ensure that HSPCs remain phenotypically

factor (SCF; 100 ng/ml LPS, 1  $\mu$ g/ml Pam3CSK4, and 50 ng/ml SCF), CpG (1  $\mu$ M), LPS (10  $\mu$ g/ml), CD3/28 (1  $\mu$ g/ml anti-CD3 and 1  $\mu$ g/ml anti-CD28). Blank, fresh medium; BM, culture medium from unstimulated total BM cells. Data are presented as a heat map with low cytokine level represented in blue, intermediate level represented in white, and high level represented in red. Cytokines are clustered into groups.

(B–E) Analysis of *in vitro* myelopoiesis from WT LSK cells by FACS. Sorted WT LSK cells were stimulated with LPS (100 ng/ml), Pam3CSK4 (1  $\mu$ g/ml), and SCF (50 ng/ml) in the presence of cytokine-neutralizing antibodies. Control, isotype antibody control; all Abs, a combination of anti-IL-6, anti-IFN- $\gamma$ , anti-GM-CSF, and anti-TNF- $\alpha$  neutralizing antibodies.

(B and C) Time course of myelopoiesis from WT LSK cells in the presence of anti-IL-6 neutralizing antibody.

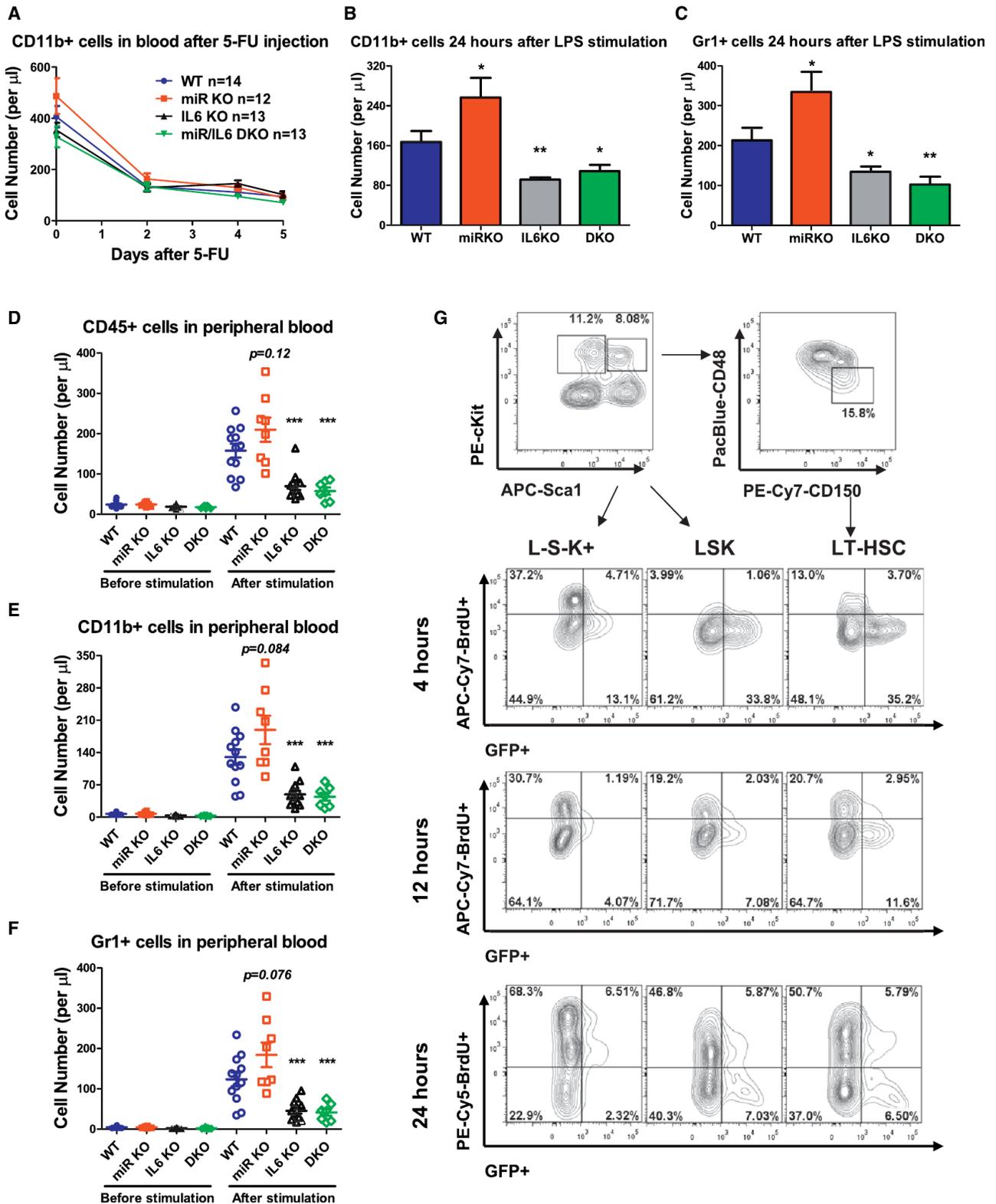
(D and E) Myelopoiesis from WT LSK cells in the presence of various neutralizing antibodies. Cells were analyzed on day 3 by FACS for percent and number of myeloid cells and LSK cells. The representative FACS plots (B and D); quantification ( $n = 3$ ; C and E).

(F and G) FACS analysis of *in vitro* myelopoiesis from WT LSK cells in the presence of conditioned medium.

(F) WT total BM cells or WT BM cells depleted of Sca1<sup>+</sup> cells were stimulated with LPS and Pam3CSK4 for 24 hr in order to induce cytokine production into the culturing medium. Then, the culturing medium was used to stimulate freshly purified WT LSK cells. Number of myeloid and LSK cells was analyzed on day 2 ( $n = 8$ ).

(G) WT or miR-146a KO LSK cells were stimulated with LPS and Pam3CSK4 for 24 hr in order to induce cytokine production into the culturing medium. Then, the culturing medium was used to stimulate freshly purified WT LSK cells. The number of myeloid and LSK cells was analyzed on day 2 ( $n = 3$ ). Data are presented as mean  $\pm$  SEM.

See also Figure S5.



**Figure 6. In Vivo Analysis of Functional Significance of HSPC-Produced Cytokines**

(A–C) For 5-FU-induced neutropenia, WT, miR-146a KO, IL-6 KO, and miR-146a/IL-6 DKO mice were injected with 5-FU (250 mg/kg of body weight, i.p.). (A) Peripheral blood was analyzed on days 0, 2, 4, and 5 by FACS in order to determine the number of CD11b<sup>+</sup> myeloid cells. On day 5, LPS (0.3 mg/kg of body weight, i.p.) was injected, and mice were bled 24 hr later in order to study LPS-induced myelopoiesis by analyzing the number of CD11b<sup>+</sup> (B) or Gr1<sup>+</sup> (C) cells.

(legend continued on next page)

undifferentiated by FACS and morphologically intact by microscopy. The ability to recover individual cells poststimulation will allow us to better assess cell viability.

Functions of NF- $\kappa$ B in immune cells have been extensively studied. However, knowledge of the functional role of NF- $\kappa$ B in HSPCs has been limited. In this study, we have connected a known function of NF- $\kappa$ B, the regulation of cytokine production, to two unexpected multipotent hematopoietic cell populations, ST-HSCs and MPPs. What is also intriguing is that some LT-HSCs appear able to respond to TLR stimulation by the activation of NF- $\kappa$ B, and all HSCs appear to have at least the p65 subunit of NF- $\kappa$ B. Thus, despite NF- $\kappa$ B by itself being insufficient to enable cytokine secretion in LT-HSCs, NF- $\kappa$ B is present from the inception of hematopoiesis and may regulate other aspects of HSC biology.

## EXPERIMENTAL PROCEDURES

### Animal Models

All mice (WT, *Nfkb1*<sup>-/-</sup>, *Mir146a*<sup>-/-</sup>, *Il6*<sup>-/-</sup>, *Nfkb1*<sup>-/-</sup>*Mir146a*<sup>-/-</sup>, *Il6*<sup>-/-</sup>*Mir146a*<sup>-/-</sup>, RELA-GFP knockin, and NF- $\kappa$ B-eGFP reporter mice) were on a C57BL/6 genetic background and housed under a specific pathogen-free condition at the California Institute of Technology. Mice used for all experiments were age- and sex-matched 6- to 8-week-old female mice. All experiments were approved by the institutional Animal Care and Use Committee of the California Institute of Technology.

### FACS Sort

In general, 15 to 20 mice of the same genotype were used for FACS sorting in order to obtain sufficient stem and progenitor cells. First, BM cells were subjected to magnetic bead selection (Miltenyi Biotec) according to manufacturer's protocol in order to deplete lineage-positive cells, and then they were sorted on a BD FACSAria sorter at the California Institute of Technology FACS Core. More details can be found in the Supplemental Experimental Procedures.

### Single-Cell Cytokine Chip Analysis

We integrated upstream FACS purification techniques with the single-cell barcode chip in order to study the functional proteomics from phenotypically defined single cells. The chips used in this study have > 5,000 microchambers of about 100 picoliter volume to enable sensitive detection of proteins. Within each microchamber, a panel of 12 cytokines (TNF- $\alpha$ , GM-CSF, IL-6, IL-12p40, IFN- $\gamma$ , IL-2, IL-4, IL-10, TGF- $\beta$ 1, CCL-2, IL-17A, and IL-1 $\beta$ ) can be simultaneously measured by sandwich ELISA-like assay. The manufacture procedure of the chip has been described in our previous study (Ma et al., 2011), and the detailed experimental steps can be found in the Supplemental Experimental Procedures. In brief, FACS-purified cells were cultured at 37°C in 5% CO<sub>2</sub> cell incubator with medium alone, medium plus LPS, or medium plus LPS and Pam3CSK4 for 12 hr. At the end of stimulation, the chip was imaged with a high-resolution bright-field microscope. Cell number in each chamber was counted, and cell viability was assessed to exclude fragmented or non-light-reflective cells by trained personnel in a blind manner. Then, cells were washed off, and the chips were developed via an immuno-sandwich assay. A GenePix 4400A microarray scanner was used to scan slides, and data were analyzed with GenePix Pro 7. Each single-cell chip analysis of a specific cell subset, genotype, and stimulating condition was performed at least two times.

### Multiplexed Cytokine Analysis of Bulk Cell-Culture Medium

For multiplexed cytokine analysis of bulk cell-culture medium, 100,000 cells of each FACS-purified subset were stimulated in 100  $\mu$ l medium with various stimulations for 24 hr in a 96-well plate. Then, the medium was collected and concentrated by 4-fold before being subjected to multiplexed ELISA quantification of 15 different cytokines and chemokines. For multiplexed ELISA quantification, the detection method was identical to the single-cell chip analysis described above. The difference is that the sample here consists of culturing medium only. In brief, the ELISA chip was first blocked with 3% BSA in PBS buffer and then hybridized with an antibody-single-stranded DNA conjugate cocktail and washed with 3% BSA in PBS followed by the application of a medium sample. The assay was completed by applying secondary biotinylated antibodies and streptavidin-cy3 in sequence. Multiplexed cytokine analysis of a specific cell subset, genotype, and stimulating condition was performed at least two times.

### In Vitro Analysis of Myeloid Differentiation

For analysis of myeloid differentiation under cytokine-neutralizing antibodies, sorted LSK cells (15,000 cells per 100  $\mu$ l medium) were stimulated with LPS (100 ng/ml), Pam3CSK4 (1  $\mu$ g/ml), and stem cell factor (50 ng/ml) in the presence of various cytokine-neutralizing antibodies. Antibody concentrations used were isotype control (1  $\mu$ g/ml), anti-IFN- $\gamma$  (1  $\mu$ g/ml), anti-IL-6 (1  $\mu$ g/ml), anti-GM-CSF (5  $\mu$ g/ml), and anti-TNF- $\alpha$  (5  $\mu$ g/ml; eBioscience). Cells were analyzed on day 3 by FACS for the percent and number of myeloid cells and LSK cells. For the analysis of myeloid differentiation under conditioned medium, WT or miR-146a KO LSK cells (30,000 cells per 100  $\mu$ l medium) were stimulated with LPS (100 ng/ml) and Pam3CSK4 (1  $\mu$ g/ml) for 24 hr in order to induce cytokine production into the culturing medium. Then, the culturing medium from either WT or miR KO LSK cells was collected and used to stimulate freshly purified WT LSK cells. The number of myeloid and LSK cells was analyzed after 2 days by FACS. All experiments were performed two times with three biological replicates, each of which was sorted from pooled BM cells of six to ten mice.

### In Vivo Analysis of Myeloid Differentiation

For 5-FU-induced leucopenia, WT, miR-146a KO, IL-6 KO, and miR-146a/IL-6 DKO mice were injected with 5-FU (250 mg/kg of body weight, intraperitoneally [i.p.] injected; Sigma-Aldrich). Peripheral blood was obtained on days 0, 2, 4, and 5 for FACS analysis in order to determine the severity of neutropenia. On day 5, LPS (0.3 mg/kg of body weight, i.p.) was injected, and mice were bled 24 hr later for FACS analysis. For stem cell transplant study after lethal irradiation, Lin<sup>-</sup>cKit<sup>+</sup> cells were purified from the BM of WT, miR-146a KO, IL-6 KO, and miR-146a/IL-6 DKO mice. 500,000 cells were injected intravenously into each lethally irradiated (1,000 rad in one dose) WT recipient mice. On day 6, mice were bled for FACS analysis and LPS injection (0.3 mg/kg of body weight, i.p.) was given; 48 hr later, mice were bled again for FACS analysis. Data represent cumulative results from two independent mouse experiments.

### Computational Algorithm and Statistical Analysis

In Figures 3 and 5, F tests were used to compare variances, and then the appropriate two-sided Student's t tests were applied. All figures with error bars were graphed as mean  $\pm$  SEM. For all heat maps, scale bars represent mean  $\pm$  SD. \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001.

For all single-cell cytokine chip analysis, custom-written software routines in R language were used to process, analyze, and visualize the single-cell functional assay results. In brief, the algorithm converts raw fluorescence images into numerical fluorescence intensity values for each assayed protein within a given microchamber matched with the number of cells. The number of cells within each microchamber was determined manually by microscopy.

(D–F) For stem cell transplant studies, Lin<sup>-</sup>cKit<sup>+</sup> cells were purified from the BM of WT, miR-146a KO, IL-6 KO, and miR-146a/IL-6 DKO mice. 500,000 cells per mouse were injected intravenously into lethally irradiated WT recipient mice. On day 6, mice were bled for FACS analysis in order to determine the degree of neutropenia. LPS was injected (0.3 mg/kg of body weight, i.p.), and, 48 hr later, mice were bled in order to study LPS-induced myelopoiesis by analyzing the number of total white blood cells (D), CD11b<sup>+</sup> cells (E), and Gr1<sup>+</sup> cells (F).

(G) LPS (2 mg/kg of body weight, i.p.) and BrdU (1 mg, i.p.) were injected into NF- $\kappa$ B-GFP transgenic reporter mice, which were harvested for FACS analysis at 4, 12, and 24 hr. Representative FACS plots of GFP expression and BrdU incorporation of L<sup>-</sup>S<sup>-</sup>K<sup>+</sup> cells, LSK cells, and HSCs. Data are presented as mean  $\pm$  SEM. See also Figure S6.

The average background signal levels from all zero-cell microchambers were used to set the gate in order to separate nonproducing cells from cytokine-producing cells. Detailed statistical analysis method for principal component analysis (Figures 1D, 2D, 2F, and 4F) can be found in the Supplemental Experimental Procedures and our previous publication (Ma et al., 2011). These types of statistical analysis and graphical representations are routinely used to analyze large-scale multidimensional data sets from numerous cell subsets (Bendall et al., 2011).

### SUPPLEMENTAL INFORMATION

Supplemental Information contains Supplemental Experimental Procedures and six figures and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2014.01.007>.

### AUTHOR CONTRIBUTIONS

J.L.Z., D.B., C.M., R.M.O., and J.R.H. conceived the study. J.L.Z., C.M., and A.M. designed and performed the experiments. R.D. helped with data collection. J.L.Z., D.B., C.M., and J.R.H. analyzed the data. J.L.Z. and D.B. wrote the manuscript.

### ACKNOWLEDGMENTS

The authors wish to thank Caltech animal facility and flow cytometry core facility and Drs. Alejandro Balazs, Devdoot Majumdar, and Michael Bethune of the D.B. lab for their help. RELA-GFP knockin and NF- $\kappa$ B-eGFP reporter mice were obtained from Dr. Manolis Pasparakis of the University of Cologne and Dr. Christian Jobin of the University of North Carolina, respectively. The work was supported by research grants R01AI079243 (D.B.), R01CA170689 (J.R.H.), National Research Service Award F30HL110691 (J.L.Z.), UCLA/Caltech Medical Scientist Training Program (J.L.Z. and A.M.), Rosen Fellowship (C.M.), NIH New Innovator Award DP2GM111099 (R.M.O.), the Pathway to Independence Award R00HL102228 (R.M.O.), and an American Cancer Society Research grant (R.M.O.) with core facilities support from 5U54CA119347 (J.R.H.).

Received: October 1, 2013  
Revised: November 28, 2013  
Accepted: January 14, 2014  
Published: February 20, 2014

### REFERENCES

- Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.J., Thoren, L.A., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: a revised road map for adult blood lineage commitment. *Cell* **121**, 295–306.
- Baldrige, M.T., King, K.Y., Boles, N.C., Weksberg, D.C., and Goodell, M.A. (2010). Quiescent haematopoietic stem cells are activated by IFN- $\gamma$  in response to chronic infection. *Nature* **465**, 793–797.
- Baldrige, M.T., King, K.Y., and Goodell, M.A. (2011). Inflammatory signals regulate hematopoietic stem cells. *Trends Immunol.* **32**, 57–65.
- Baltimore, D. (2011). NF- $\kappa$ B is 25. *Nat. Immunol.* **12**, 683–685.
- Beerman, I., Bhattacharya, D., Zandi, S., Sigvardsson, M., Weissman, I.L., Bryder, D., and Rossi, D.J. (2010). Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc. Natl. Acad. Sci. USA* **107**, 5465–5470.
- Bendall, S.C., Simonds, E.F., Qiu, P., Amir, A.D., Krutzik, P.O., Finck, R., Bruggner, R.V., Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696.
- Bennett, C.L., Djulbegovic, B., Norris, L.B., and Armitage, J.O. (2013). Colony-stimulating factors for febrile neutropenia during cancer therapy. *N. Engl. J. Med.* **368**, 1131–1139.
- Boldin, M.P., Taganov, K.D., Rao, D.S., Yang, L., Zhao, J.L., Kalwani, M., Garcia-Flores, Y., Luong, M., Devrekanli, A., Xu, J., et al. (2011). miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *J. Exp. Med.* **208**, 1189–1201.
- Challen, G.A., Boles, N.C., Chambers, S.M., and Goodell, M.A. (2010). Distinct hematopoietic stem cell subtypes are differentially regulated by TGF- $\beta$ 1. *Cell Stem Cell* **6**, 265–278.
- Copley, M.R., Beer, P.A., and Eaves, C.J. (2012). Hematopoietic stem cell heterogeneity takes center stage. *Cell Stem Cell* **10**, 690–697.
- De Lorenzi, R., Gareus, R., Fengler, S., and Pasparakis, M. (2009). GFP-p65 knock-in mice as a tool to study NF- $\kappa$ B dynamics in vivo. *Genesis* **47**, 323–329.
- Essers, M.A., Offner, S., Blanco-Bose, W.E., Waibler, Z., Kalinke, U., Duchosal, M.A., and Trumpp, A. (2009). IFN $\alpha$  activates dormant haematopoietic stem cells in vivo. *Nature* **458**, 904–908.
- Forsberg, E.C., Serwold, T., Kogan, S., Weissman, I.L., and Passegué, E. (2006). New evidence supporting megakaryocyte-erythrocyte potential of flk2/flt3+ multipotent hematopoietic progenitors. *Cell* **126**, 415–426.
- Han, J., Koh, Y.J., Moon, H.R., Ryoo, H.G., Cho, C.H., Kim, I., and Koh, G.Y. (2010). Adipose tissue is an extramedullary reservoir for functional hematopoietic stem and progenitor cells. *Blood* **115**, 957–964.
- Harrison, D.E., and Lerner, C.P. (1991). Most primitive hematopoietic stem cells are stimulated to cycle rapidly after treatment with 5-fluorouracil. *Blood* **78**, 1237–1240.
- Jaiswal, S., and Weissman, I.L. (2009). Hematopoietic stem and progenitor cells and the inflammatory response. *Ann. N Y Acad. Sci.* **1174**, 118–121.
- Janeway, C.A., Jr., Travers, P., Walport, M., and Shlomchik, M.J. (2001). *Immunobiology: The Immune System in Health and Disease*. Appendix III. Cytokines and Their Receptors, Fifth Edition. (New York: Garland Science).
- King, K.Y., and Goodell, M.A. (2011). Inflammatory modulation of HSCs: viewing the HSC as a foundation for the immune response. *Nat. Rev. Immunol.* **11**, 685–692.
- Ma, C., Fan, R., Ahmad, H., Shi, Q., Comin-Anduix, B., Chodon, T., Koya, R.C., Liu, C.C., Kwong, G.A., Radu, C.G., et al. (2011). A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells. *Nat. Med.* **17**, 738–743.
- Maeda, K., Malykhin, A., Teague-Weber, B.N., Sun, X.H., Farris, A.D., and Coggeshall, K.M. (2009). Interleukin-6 aborts lymphopoiesis and elevates production of myeloid cells in systemic lupus erythematosus-prone B6.Sle1.Yaa animals. *Blood* **113**, 4534–4540.
- Magness, S.T., Jijon, H., Van Houten Fisher, N., Sharpless, N.E., Brenner, D.A., and Jobin, C. (2004). In vivo pattern of lipopolysaccharide and anti-CD3-induced NF- $\kappa$ B activation using a novel gene-targeted enhanced GFP reporter gene mouse. *J. Immunol.* **173**, 1561–1570.
- Massberg, S., Schaerli, P., Knezevic-Maramica, I., Köllnberger, M., Tubo, N., Moseman, E.A., Huff, I.V., Junt, T., Wagers, A.J., Mazo, I.B., and von Andrian, U.H. (2007). Immunosurveillance by hematopoietic progenitor cells trafficking through blood, lymph, and peripheral tissues. *Cell* **131**, 994–1008.
- Megias, J., Yáñez, A., Moriano, S., O'Connor, J.E., Gozalbo, D., and Gil, M.L. (2012). Direct Toll-like receptor-mediated stimulation of hematopoietic stem and progenitor cells occurs in vivo and promotes differentiation toward macrophages. *Stem Cells* **30**, 1486–1495.
- Mossadegh-Keller, N., Sarrazin, S., Kandalla, P.K., Espinosa, L., Stanley, E.R., Nutt, S.L., Moore, J., and Sieweke, M.H. (2013). M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature* **497**, 239–243.
- Nagai, Y., Garrett, K.P., Ohta, S., Bahrun, U., Kouro, T., Akira, S., Takatsu, K., and Kincade, P.W. (2006). Toll-like receptors on hematopoietic progenitor cells stimulate innate immune system replenishment. *Immunity* **24**, 801–812.
- Naik, S.H., Perié, L., Swart, E., Gerlach, C., van Rooij, N., de Boer, R.J., and Schumacher, T.N. (2013). Diverse and heritable lineage imprinting of early haematopoietic progenitors. *Nature* **496**, 229–232.

Pronk, C.J., Veiby, O.P., Bryder, D., and Jacobsen, S.E. (2011). Tumor necrosis factor restricts hematopoietic stem cell activity in mice: involvement of two distinct receptors. *J. Exp. Med.* *208*, 1563–1570.

Sha, W.C., Liou, H.C., Tuomanen, E.I., and Baltimore, D. (1995). Targeted disruption of the p50 subunit of NF-kappa B leads to multifocal defects in immune responses. *Cell* *80*, 321–330.

Shen, Y., and Nilsson, S.K. (2012). Bone, microenvironment and hematopoiesis. *Curr. Opin. Hematol.* *19*, 250–255.

Wesemann, D.R., Portuguese, A.J., Meyers, R.M., Gallagher, M.P., Cluff-Jones, K., Magee, J.M., Panchakshari, R.A., Rodig, S.J., Kepler, T.B., and

Alt, F.W. (2013). Microbial colonization influences early B-lineage development in the gut lamina propria. *Nature* *501*, 112–115.

Zhao, J.L., Rao, D.S., Boldin, M.P., Taganov, K.D., O'Connell, R.M., and Baltimore, D. (2011). NF-kappaB dysregulation in microRNA-146a-deficient mice drives the development of myeloid malignancies. *Proc. Natl. Acad. Sci. USA* *108*, 9184–9189.

Zhao, J.L., Rao, D.S., O'Connell, R.M., Garcia-Flores, Y., and Baltimore, D. (2013). MicroRNA-146a acts as a guardian of the quality and longevity of hematopoietic stem cells in mice. *Elife* *2*, e00537.

# Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming

Xiao Hu,<sup>1,5</sup> Lei Zhang,<sup>1,5</sup> Shi-Qing Mao,<sup>1,5</sup> Zheng Li,<sup>1</sup> Jiekai Chen,<sup>2</sup> Run-Rui Zhang,<sup>1</sup> Hai-Ping Wu,<sup>1</sup> Juan Gao,<sup>1</sup> Fan Guo,<sup>1</sup> Wei Liu,<sup>1</sup> Gui-Fang Xu,<sup>1</sup> Hai-Qiang Dai,<sup>1</sup> Yujiang Geno Shi,<sup>3</sup> Xianlong Li,<sup>4</sup> Boqiang Hu,<sup>4</sup> Fuchou Tang,<sup>4</sup> Duanqing Pei,<sup>2</sup> and Guo-Liang Xu<sup>1,\*</sup>

<sup>1</sup>Shanghai Key Laboratory of Molecular Andrology, The State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

<sup>2</sup>CAS Key Laboratory of Regenerative Biology, South China Institute for Stem Cell Biology and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China

<sup>3</sup>Division of Endocrinology, Diabetes, and Hypertension, Department of Medicine and BCMP, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

<sup>4</sup>Biodynamic Optical Imaging Center, College of Life Sciences, Peking University, Beijing 100871, China

<sup>5</sup>These authors contributed equally to this work

\*Correspondence: glxu@sibs.ac.cn

<http://dx.doi.org/10.1016/j.stem.2014.01.001>

## SUMMARY

Tet-mediated DNA oxidation is a recently identified mammalian epigenetic modification, and its functional role in cell-fate transitions remains poorly understood. Here, we derive mouse embryonic fibroblasts (MEFs) deleted in all three *Tet* genes and examine their capacity for reprogramming into induced pluripotent stem cells (iPSCs). We show that Tet-deficient MEFs cannot be reprogrammed because of a block in the mesenchymal-to-epithelial transition (MET) step. Reprogramming of MEFs deficient in TDG is similarly impaired. The block in reprogramming is caused at least in part by defective activation of key miRNAs, which depends on oxidative demethylation promoted by Tet and TDG. Reintroduction of either the affected miRNAs or catalytically active Tet and TDG restores reprogramming in the knockout MEFs. Thus, oxidative demethylation to promote gene activation appears to be functionally required for reprogramming of fibroblasts to pluripotency. These findings provide mechanistic insight into the role of epigenetic barriers in cell-lineage conversion.

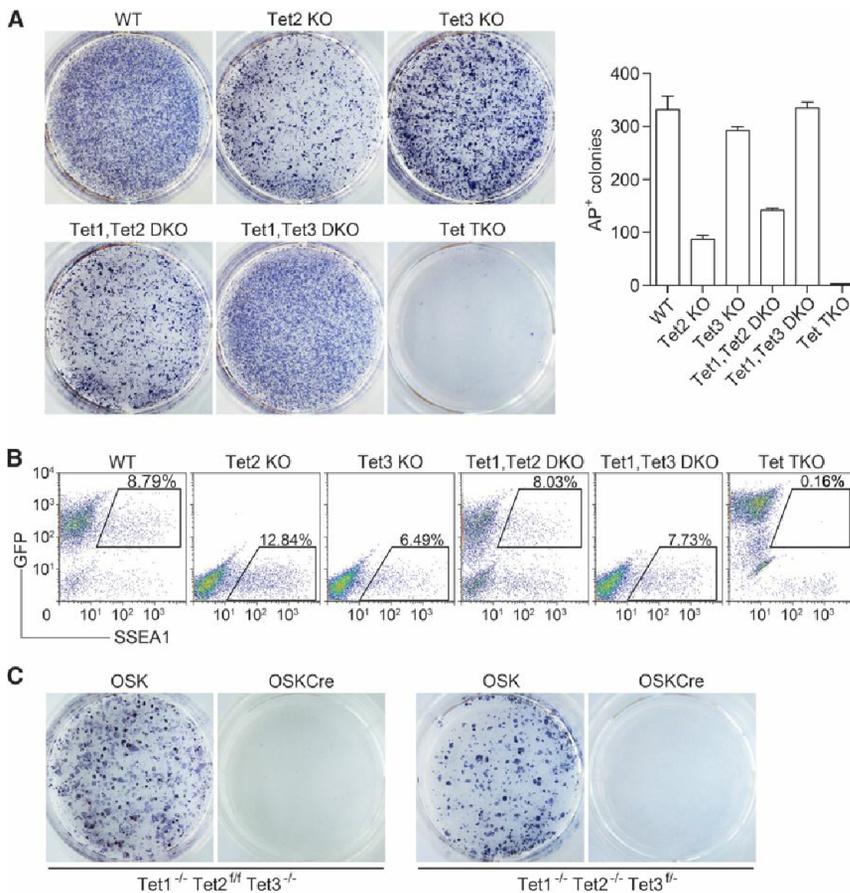
## INTRODUCTION

Pluripotent cells, such as those in early embryos, proliferate and differentiate into distinctive cell lineages. Lineage commitment is established and maintained in development by epigenetic programming of gene-expression profiles, in which DNA methylation plays a prominent role (Goll and Bestor, 2005; Jaenisch and Bird, 2003). DNA methylation patterns are faithfully propagated in cells undergoing mitosis. Early embryonic genes, such as the pluripotency genes *Oct4* and *Nanog*, undergo silencing

and de novo DNA methylation in their promoter and enhancer regions during cell differentiation and maintain their hypermethylated states in differentiated somatic cells (Epsztejn-Litman et al., 2008; Li et al., 2007). The developmental and cell-differentiation processes therefore entail dynamic regulation of genomic methylation accompanied by gene-expression changes.

Lineage commitment can be reversed in vivo and in vitro through natural and experimental reprogramming such as nuclear transfer of a somatic nucleus into an enucleated oocyte and factor-induced conversion of somatic cells to induced pluripotent stem cells (iPSCs). Because each cell identity is supposedly defined by a unique methylation profile that underpins its lineage commitment and serves as a barrier between different cell types, methylation reprogramming is a mechanistically vital process underlying cell-type switch. For example, demethylation of pluripotency genes is a hallmark of somatic-cell reprogramming into a pluripotent state (Gurdon and Melton, 2008; Takahashi and Yamanaka, 2006). Transcriptional activation of epigenetically silenced genes thus necessitates the demethylation of critical regulatory elements in DNA during experimental reprogramming.

The Tet family of DNA dioxygenases catalyzes oxidation of methylcytosines to hydroxymethylcytosines (5hmC), formylcytosines (5fC), and carboxylcytosines (5caC) (He et al., 2011; Ito et al., 2011; Tahiliani et al., 2009). An active mode of DNA demethylation has been proposed to encompass Tet-mediated oxidation of methylcytosines and excision of the higher oxidation products by DNA glycosylase TDG (He et al., 2011; Pastor et al., 2013; Seisenberger et al., 2013). However, the possibility of replication-associated passive dilution of 5-methylcytosine (5mC) and its oxidation products (Inoue and Zhang, 2011) has complicated the dissection of the role of Tet-initiated active demethylation in development and cell reprogramming. Whereas overexpression of Tet1 and Tet2 has been shown to promote iPSC formation through the reactivation of silenced pluripotency genes (Bagci and Fisher, 2013; Costa et al., 2013; Doege et al., 2012), the importance of and mechanism behind the Tet-mediated oxidation in cell reprogramming have remained largely undefined.



**Figure 1. Tet Triple Knockout Prevents Reprogramming of MEFs**

(A) Alkaline phosphatase (AP) staining of cells 10 days after retroviral transduction of Oct4, Sox2, and Klf4 (OSK). Wild-type (WT), Tet2 KO, Tet3 KO, Tet1,Tet2 double knockout (DKO), Tet1,Tet3 DKO, and Tet1,Tet2,Tet3 triple knockout (TKO) MEFs were assayed for iPSC formation. AP-positive colony numbers are shown on the right. Data represent means  $\pm$  SD of three independent experiments.

(B) FACS analysis of SSEA1-positive cells at reprogramming day 10. WT, Tet1,Tet2 DKO, and TKO MEFs were sorted from chimera embryos injected with GFP-labeled ESCs. Other MEF types were isolated from homozygous embryos generated by mating heterozygotes. SSEA1-positive cells induced from each MEF type are those within polygon gates.

(C) No AP-positive colonies were generated from [Tet2-floxed, Tet1,Tet3 DKO] and [Tet3-floxed, Tet1,Tet2 DKO] MEFs upon acute deletion of Tet2 or Tet3 by Cre expression. See also Figures S1 and S2.

In this study, we generated mouse embryonic fibroblasts (MEFs) deleted of components of the putative active DNA demethylation pathway and tested these MEFs for iPSC induction and the reactivation of genes critical for cellular reprogramming. We demonstrate here that demethylation promoted by Tet and TDG is directly involved and essential in the reactivation of miRNAs that enables a mesenchymal-to-epithelial transition to initiate the reprogramming process.

## RESULTS

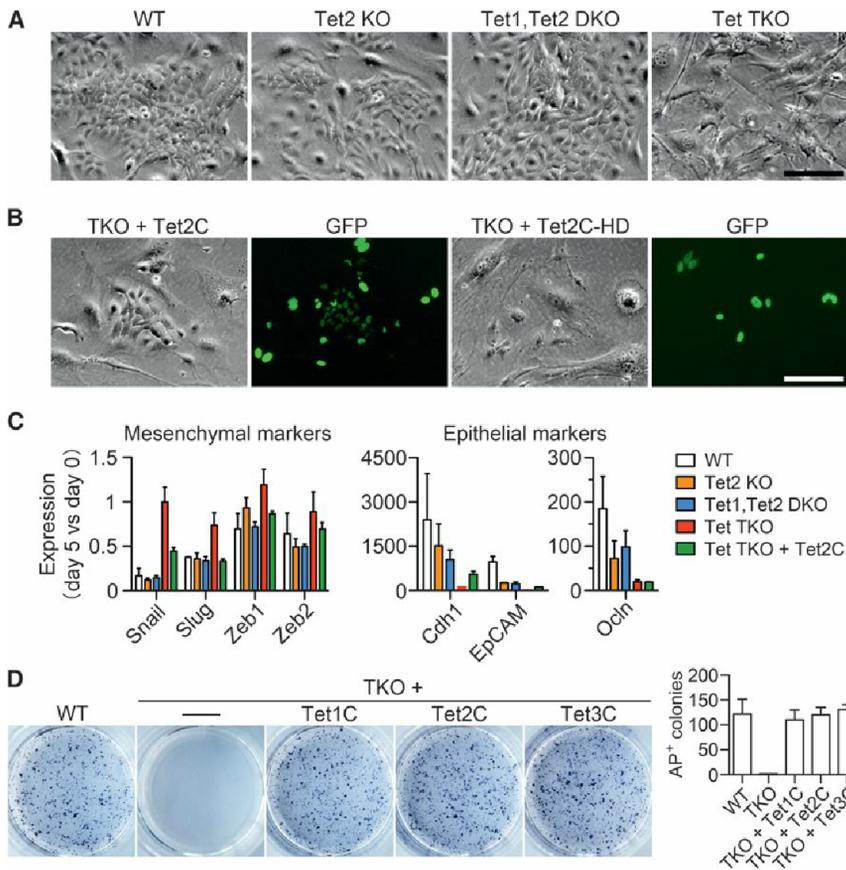
### Tet Dioxygenases Are Essential for Fibroblasts to Undergo Reprogramming

Although Tet1 and Tet2 have been shown to facilitate cell reprogramming (Costa et al., 2013; Doege et al., 2012), it remains unclear whether Tet-mediated DNA oxidation is essential for this process. To assess the functional importance of Tet enzymes, we generated mouse embryonic fibroblasts from embryonic stem cells (ESCs) deficient in Tet genes for the generation of iPSCs (see Supplemental Experimental Procedures available online). ESCs lacking all three Tet genes seemed normal in self-renewal and pluripotency (Figure S1). MEFs generated from chimeric embryos with blastocyst injection of these ESCs proliferated at a similar rate as wild-type control MEFs and those isolated from homozygous embryos resulting from heterozygous mating (Figure S2A). Considering the dynamic expression of three Tet genes during iPSC generation (Figure S2B), we deter-

mined and compared the reprogramming efficiency among MEFs deficient in a single Tet gene and in combinations, using the three-factor (Oct4, Sox2, and Klf4) induction system (Figure S2C). MEFs with Tet1 deletion had slightly increased reprogramming efficiency as described (Chen et al., 2013), and inactivation of Tet3 had little effect, based on the comparison of alkaline phosphatase (AP)-positive colony numbers (Figure 1A). Although inactivation of Tet2 reduced the reprogramming by  $\sim$ 70%, AP-positive colonies could still appear. In addition, MEFs from Tet1,Tet2 and Tet1,Tet3 double knockouts also generated numerous colonies and SSEA1-positive cells. Strikingly, inactivation of all three Tet genes completely abolished the reprogramming potential of MEFs as evidenced by the failure to obtain AP- (Figure 1A) and SSEA1-positive cells (Figure 1B). The resistance of triple knockout (TKO) MEFs to reprogramming was further validated using high-performance engineered factors (Wang et al., 2011) in the presence of c-Myc and improved medium (Chen et al., 2011; Figure S2D). The reprogramming deficiency of TKO MEFs could not be ascribed to inherent genomic or epigenomic alterations potentially arisen from the constitutive Tet deletion, because AP-positive colonies also did not appear upon the acute deletion of Tet2 from the Tet1,Tet3 double knockout (DKO) MEFs or the acute deletion of Tet3 from Tet1,Tet2 DKO MEFs by the Cre recombinase expression (Figure 1C). Moreover, Tet TKO MEFs could be rescued for reprogramming in latter experiments. These observations indicate that the Tet enzymes are indispensable for factor-driven reprogramming of somatic cells.

### Tet TKO MEFs Fail to Undergo Mesenchymal-to-Epithelial Transition during Reprogramming

Factor-driven reprogramming is a multistep process initiated by mesenchymal-to-epithelial transition (MET) as an essential event



**Figure 2. Tet-Deficient MEFs Are Unable to Initiate Mesenchymal-to-Epithelial Transition during Reprogramming**

(A) Phase-contrast photographs of cells under reprogramming (day 5) showing lack of the morphologic shift characteristic of MET in Tet TKO MEFs. The scale bar represents 100  $\mu$ m.

(B) Rescue of MET in TKO MEFs by ectopic expression of the Tet2 C-terminal catalytic domain (Tet2C). The MEFs used were derived from TKO ESCs labeled with lentiviral H2B-enhanced GFP (EGFP). Tet2C-HD carries two point mutations in the active site of the enzyme. The scale bar represents 100  $\mu$ m.

(C) Failed downregulation of mesenchymal and upregulation of epithelial markers in Tet TKO MEFs as demonstrated by qRT-PCR. The expression levels were normalized to *Gapdh*. Data represent means  $\pm$  SD of three independent experiments.

(D) AP staining of colonies from WT MEFs, Tet TKO MEFs, and TKO MEFs transduced with the indicated Tet at reprogramming day 10. AP-positive colony numbers are shown on the right. Data represent means  $\pm$  SD of three independent experiments.

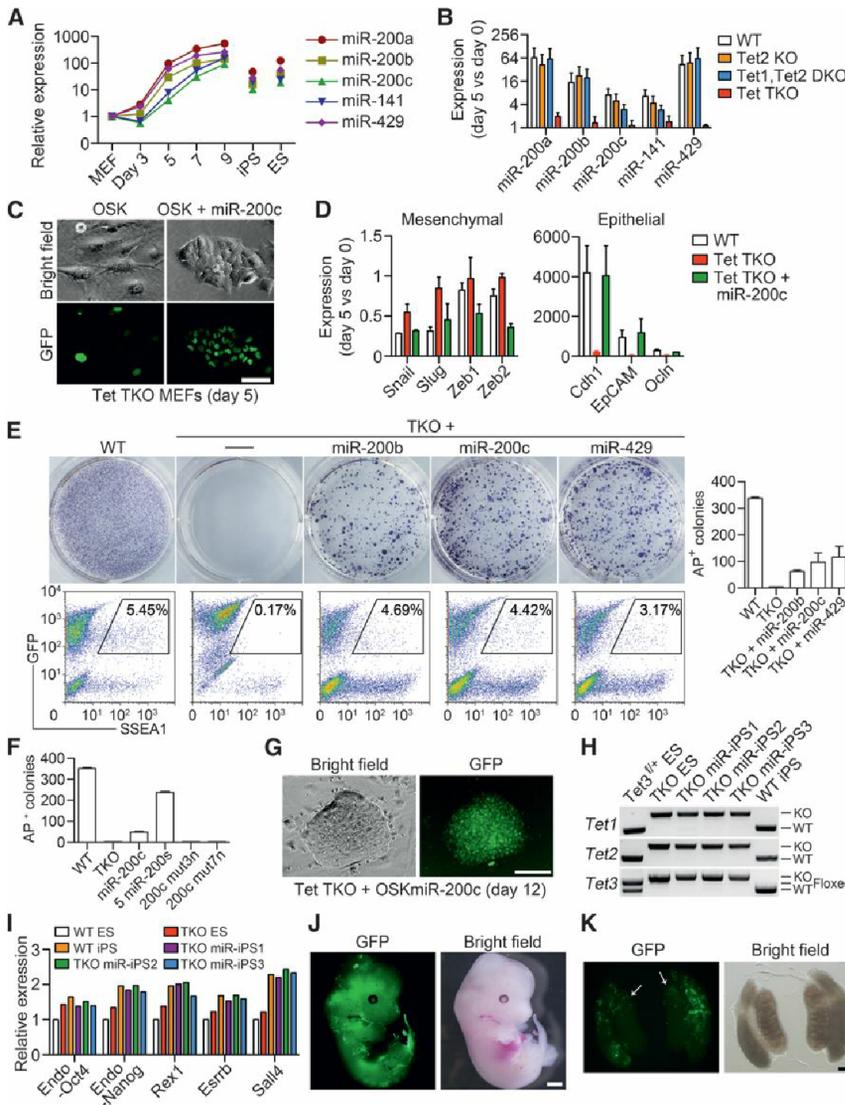
for MEFs to be successfully reprogrammed into iPSCs (Li et al., 2010; Samavarchi-Tehrani et al., 2010). The fact that AP and SSEA1, markers of early reprogramming, were not induced in TKO MEFs suggested an early reprogramming defect. Indeed, Tet TKO MEFs showed no sign of epithelium-like morphological shift, whereas wild-type, single, or double knockout MEFs exhibited an obvious MET, starting from day 4 after transduction of the Oct4, Sox2, and Klf4 (OSK) reprogramming factors (Figures 2A and S2C). The MET process in TKO MEFs could be rescued by ectopic expression of the wild-type catalytic domain, but not the inactive form of Tet2 (Figure 2B).

E-cadherin (Cdh1), EpCAM, and occludin (Ocln) are components of epithelial cell junctions, and without their expression, cells are unable to form colonies. Snail, Slug, Zeb1, and Zeb2 are transcription factors that promote the epithelial-mesenchymal transition (EMT) and maintain the mesenchymal phenotype by directly repressing epithelial gene expression (Thiery et al., 2009). Quantitative RT-PCR (qRT-PCR) analysis confirmed persistent expression of the mesenchymal markers *Snail*, *Slug*, *Zeb1*, and *Zeb2* and lack of activation of the epithelial markers *Cdh1*, *EpCAM*, and *Ocln* in Tet-deficient MEFs (Figure 2C), indicating a resistance to MET at the molecular level. The reprogramming capacity of the TKO MEFs appeared to be fully restored by ectopic expression of the catalytic domain from any of the Tet proteins (Figure 2D). Our data suggest that Tet-deficient MEFs fail to initiate the MET process due to their inability to downregulate the key mesenchymal regulators and that MET depends on the catalytic function of Tet enzymes.

**Tet-Deficient MEFs Fail to Reactivate MicroRNAs Critical for MET**

The miR-200 s, miR-200a, miR-200b, miR-200c, miR-141, and miR-429 are causatively involved in both cancer metastasis and experimental cell reprogramming by modulating the expression of transcription factors such as Zeb1 and Zeb2 that repress epithelial markers including E-cadherin (Gregory et al., 2008; Samavarchi-Tehrani et al., 2010; Wang et al., 2013). We confirmed that all five members of the miR-200 family were substantially upregulated in MEFs undergoing reprogramming (Figure 3A). Interestingly, the expression of miR-200 family was diminishing in Tet-deficient MEFs, most severely in TKO cells (Figure 3B). Remarkably, ectopic expression of miR-200c in the TKO MEFs restored the MET process as evident in the cell morphology, formation of colonies, and epithelial marker expression (Figures 3C and 3D). In fact, each of the three miR-200 members tested, but not the two independent miR-200c mutants containing base changes in the miRNA seed region, enabled a marked rescue in terms of AP-positive colony and SSEA-positive cell numbers (Figures 3E and 3F). Combination of the five members led to a rescue of up to 80% of the reprogramming efficiency of the wild-type MEFs (Figure 3F). Importantly, iPSC lines could be established from the TKO MEFs upon miR-200 rescue. These iPSCs displayed the typical ESC morphology (Figures 3G and 3H) and expressed endogenous pluripotency genes at similar levels as in ESCs (Figure 3I). When injected into mouse blastocysts, they contributed to the generation of chimeric embryos (Figure 3J) and to the germline in embryonic gonads (Figure 3K).

To investigate whether the failure in MET is indeed the major reason for reprogramming deficiency of TKO MEFs, we derived two types of epithelium-like cells, keratinocytes and neural



**Figure 3. Impediment to Activation of miRNAs Crucial for MET in Tet-Deficient MEFs**

(A) Activation of the miR-200 family miRNAs during reprogramming. Expression levels normalized to *Rnu6* are relative to the value in untransduced MEFs which is set to 1.

(B) Failed upregulation of the miR-200 family miRNAs in Tet-deficient MEFs transduced with OSK (day 5). Data represent means  $\pm$  SD of two independent experiments.

(C) Restoration of MET in Tet-deficient MEFs by ectopic expression of miR-200c. Representative photographs of cells under induction with OSK alone and together with the ectopic expression of miR-200c (day 5) are shown. Nuclear H2B-GFP identifies Tet-deficient MEFs originally from the labeled TKO ESCs. The scale bar represents 50  $\mu$ m.

(D) Downregulation of mesenchymal genes and upregulation of epithelial genes in TKO MEFs undergoing OSK-driven reprogramming in the presence of ectopic miR-200c. Data represent means  $\pm$  SD of two independent experiments.

(E) AP (upper) and SSEA1 (lower) staining of cells from indicated MEFs rescued with miR-200b, miR-200c, or miR-429 at reprogramming day 10. Wild-type MEFs and TKO MEFs without ectopic miRNA expression were used for comparison. AP-positive colony numbers are shown on the right. Data represent means  $\pm$  SD of two independent experiments.

(F) Rescue assay with five miRNAs in combination and miR-200c mutants. AP-positive colonies emerged from WT and TKO MEFs without and with the rescue by indicated miRNAs were scored at reprogramming day 10. *mut3n* and *mut7n* are two mutants of miR-200c, harboring 3- and 7-base change, respectively, in the seed region.

(G) Formation of iPSC colonies from TKO MEFs rescued for MET with ectopic miR-200c. GFP expression confirms the cell origin from TKO ESCs used to generate MEFs. The scale bar represents 80  $\mu$ m.

(H) Genotype confirmation of three independent TKO iPSC lines 1–3 (denoted as miR-iPSC). Genomic PCR of *Tet3f*/+ iPSCs, TKO ESCs, and wild-type iPSCs provides controls for identifying the bands representing the WT and knockout (KO) alleles of the three *Tet* genes.

(I) Expression of pluripotency markers in Tet TKO miR-iPSCs.

(J) Chimera formation assay. Representative images of whole-mount embryos of E12.5 are shown. Note that green fluorescent cells were from GFP-labeled Tet TKO miR-iPSCs. The scale bar represents 1 mm.

(K) Representative images of chimeric genital ridges isolated from an E12.5 embryo. GFP-positive cells (arrows) identify those originally from Tet TKO miR-iPSCs. The scale bar represents 200  $\mu$ m.

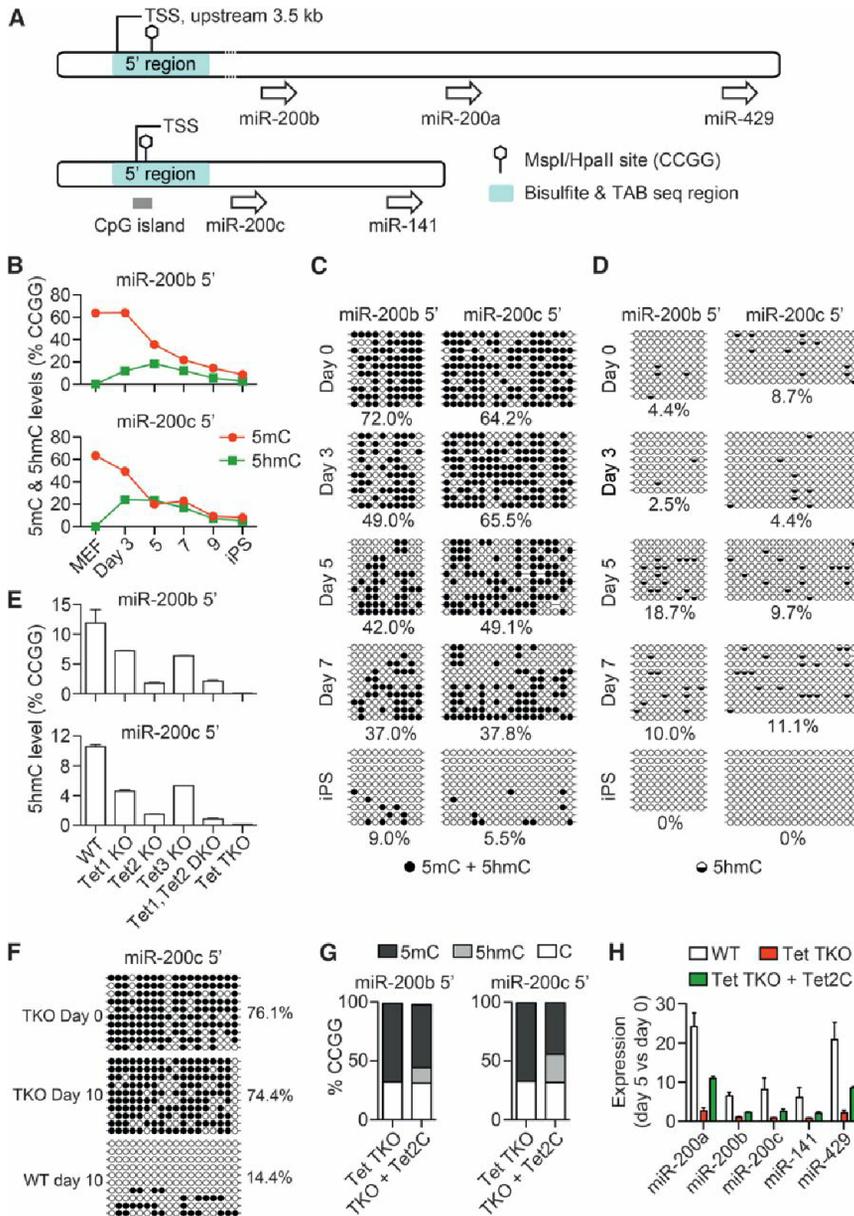
See also Figure S3.

progenitor cells (NPCs), and examined their reprogramming in the absence of Tets. As indicated in the AP colony-formation assay, neonatal keratinocytes [*Tet1*<sup>-/-</sup> *Tet2*<sup>-/-</sup> *Tet3*<sup>f/+</sup>] could be reprogrammed efficiently upon conditional deletion of the only functional *Tet3* allele (Figures S3A–S3D). NPCs derived from TKO ESCs also underwent efficient programming (Figures S3E–S3H). Moreover, inactivation of floxed *Tet2*, in [*Tet1*<sup>-/-</sup> *Tet2*<sup>f/f</sup> *Tet3*<sup>-/-</sup>] MEFs at a stage (days 4–6) subsequent to the MET did not seem to have any effect whereas earlier inactivation led to no AP-positive colonies (Figures S3I–S3K). These results demonstrate that the expression of the miR-200 family that is critical for the MET relies on Tet enzymes and that the

inability of the Tet-deficient MEFs to be reprogrammed can be primarily attributed to a failure to activate their expression and initiate MET.

**Impaired Oxidative Demethylation of miRNA Genes in Tet-Deficient MEFs**

Having revealed the lack of miRNA reactivation as a major cause for blocking MET in Tet-deficient MEFs, we reasoned that persistent epigenetic silencing might have rendered these somatic cells unresponsive to reprogramming. The miR-200 family has two clusters in the mouse genome, with miR-200b, miR-200a, and miR-429 sharing one transcript and miR-200c



**Figure 4. Tet-Dependent Oxidative Demethylation and Activation of MET-Related miRNA Genes**

(A) Schematic illustration of the miR-200 family gene clusters. Hexagons denote the C(C)GG sites analyzed in the glucosylated hydroxymethyl-sensitive qPCR (GlucMS-qPCR) assay in (B), (E), and (G). TSS denotes putative transcription start sites. (B) Dynamic regulation of methylation and hydroxylation at miR-200b and miR-200c 5' region (putative promoter region) during reprogramming as reflected by the changes of the 5mC and 5hmC levels determined by GlucMS-qPCR assay. The y axis shows % of modification (5mC or 5hmC) in the C(C)GG site analyzed. The cells of days 3–9 were FACS-sorted SSEA1-positive cells.

(C) Profiles of 5mC + 5hmC in the putative promoter regions of miR-200b and miR-200c in cells before (MEF) and during reprogramming (OSK days 3, 5, and 7) revealed by bisulfite sequencing. The cells of days 3–7 were FACS-sorted SSEA1-positive cells.

(D) Profiles of 5hmC revealed by Tet-assisted bisulfite (TAB) sequencing. The cells of days 3–7 were FACS-sorted SSEA1-positive cells.

(E) Impaired 5mC hydroxylation in Tet-deficient MEFs. The y axis shows % of 5hmC in the C(C)GG site determined by GlucMS-qPCR assay. Data represent means ± SD of two independent experiments.

(F) Impaired demethylation in the 5' region of miR-200c in TKO MEFs during reprogramming. MEFs or total cells under reprogramming at day 10 were analyzed by bisulfite sequencing.

(G) Restoration of 5mC hydroxylation in Tet-deficient MEFs by ectopic Tet2C. The 5hmC levels at the indicated miR loci in MEFs under reprogramming at day 5 were determined by GlucMS-qPCR assay.

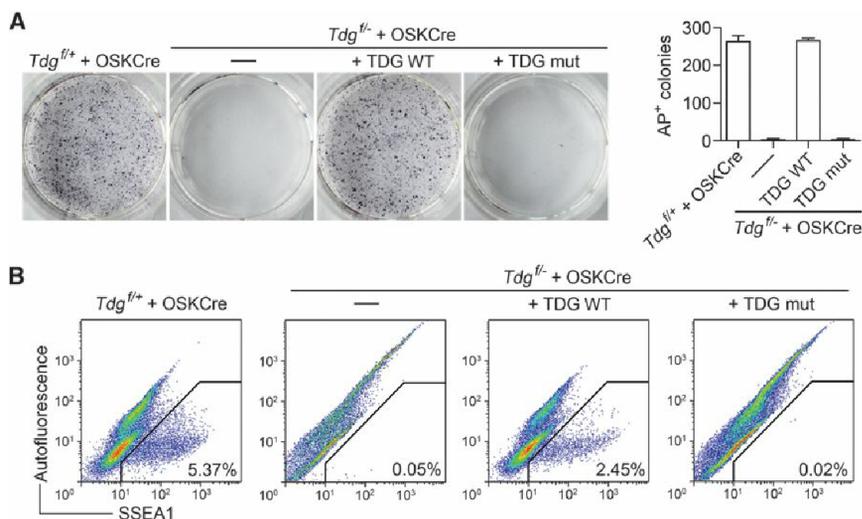
(H) Partial rescue of miRNA activation in Tet-deficient MEFs under reprogramming by ectopic expression of the Tet2C. Data represent means ± SD of two independent experiments.

See also Figure S4.

and miR-141 sharing another (Figure 4A). In order to analyze the dynamics of DNA methylation during reprogramming, we sorted SSEA1-positive cells (1%–10% of the total population) at reprogramming days 3–9, which are considered a population poised to become iPSCs (Polo et al., 2012). In wild-type MEFs under reprogramming, the putative promoter loci (or 5' region) of miR-200b and miR-200c and other examined loci within the gene body registered a dynamic reprogramming of DNA modifications as reflected in the results of the GlucMS-qPCR assay that quantifies 5mC and 5hmC at selected CCGG sites (Figure 4B; data not shown). The hypermethylation state decreases gradually, in parallel with the appearance of 5hmC, which peaked around day 5. Detailed bisulfite sequencing analysis in the 5' regions of both clusters revealed hypermethylation in MEFs at the start (day 0), hypomethylation in iPSCs, and intermediate levels at days 3–7

(Figure 4C). Because conventional bisulfite sequencing cannot distinguish 5hmC from 5mC, we applied Tet-assisted bisulfite sequencing (Yu et al., 2012b) to analyze the occurrence of 5hmC. Strikingly, the 5hmC levels in day 5 MEFs undergoing reprogramming reached 18.7% and 9.7%, respectively, in the 5' regions (Figure 4D). Most importantly, Tet triple knockout abolished 5hmC formation with Tet2 knockout (KO) having the greatest effect among the three individual knockouts (Figure 4E). As a consequence of failed hydroxylation, hypermethylation as exemplified by the 5' region of miR-200c was still retained in TKO cells by day 10 (Figure 4F).

In order to gain a global view of the dynamics of methylcytosine and hydroxymethylcytosine during reprogramming, we performed reduced representation bisulfite sequencing (RRBS) and Tet-assisted-RRBS (TA-RRBS) of cellular DNA at reprogramming days 0 and 5. A marked increase in 5hmC was noticed in various genomic regions in the initial days, accompanying the



**Figure 5. TDG Knockout Prevents Reprogramming of MEFs (see also Figure S5)**

(A) AP staining of cells 10 days after transduction. Quantification of AP-positive colonies is shown at the right. MEFs with two *Tdg* genotypes (*f/+* and *f/-*) were transduced with retroviral Oct4, Sox2, and Klf4 (OSK) reprogramming factors and the Cre-EGFP fusion. TDG-depleted MEFs (*Tdg f/-* + Cre) were also cotransduced with the wild-type (WT) or the mutant form (mut; N151A) of TDG to test for rescue. Data represent means  $\pm$  SD of two independent experiments.

(B) Representative FACS plots of MEFs stained for SSEA1.

See also Figure S5.

decrease in 5mC (Figures S4A and S4B). An inverse correlation between 5hmC and 5mC was also seen among promoters (Figure S4C). Gene Ontology analysis indicated that the most-enriched promoters were associated with cell adhesion (Figure S4D), consistent with the idea that MET-related genes are a group of targets for Tet-catalyzed hydroxylation during the early phase of reprogramming. Among the enriched targets subject to methylation reprogramming were the miR-200 clusters (Figure S4E). Despite the fact that the *Oct4* locus also gained 5hmC (Figure S4E), the combined level of 5mC and 5hmC did not seem to decline and its expression was lacking in the early phase of reprogramming as shown in latter experiments.

A further functional validation of Tet enzymes in the epigenetic regulation of miRNAs was provided by the observation that ectopic expression of the Tet2 catalytic domain restored the 5hmC level in the 5' regions of the two clusters in the TKO MEFs (Figure 4G) and resulted in upregulation of the miRNAs, albeit to a lesser degree than in the wild-type MEFs (Figure 4H). Taken together, these data demonstrate that Tet-dependent 5mC hydroxylation demethylates and reactivates MET-promoting miRNAs in reprogramming MEFs.

### TDG Is Essential for MET and Reprogramming MEFs to iPSCs

DNA hydroxylation in reprogramming MEFs can have several ramifications for epigenetic regulation. Apart from the possibility of 5hmC serving as a stable mark, which is unlikely the major function of 5hmC predicated on its transitory occurrence, DNA demethylation may happen in part by hydroxylation, followed by passive dilution or further conversion to 5fC and 5caC (He et al., 2011; Ito et al., 2011) and subsequent removal by an active mechanism. TDG is implicated in DNA demethylation initiated by Tet enzymes due to its ability to recognize and excise the higher oxidation products 5fC and 5caC (He et al., 2011; Maiti and Drohat, 2011), although the biological significance of the Tet-TDG functional connection has not been fully evaluated. We further extended our findings of the coupling between Tet and TDG by demonstrating their interdependency in the activation of a methylated reporter gene (Figure S5A). In order to explore the potential involvement of Tet-catalyzed higher oxidation in cell reprogram-

ming, we established inducible *Tdg* knockout MEFs from embryos carrying a floxed and a null allele and tested them

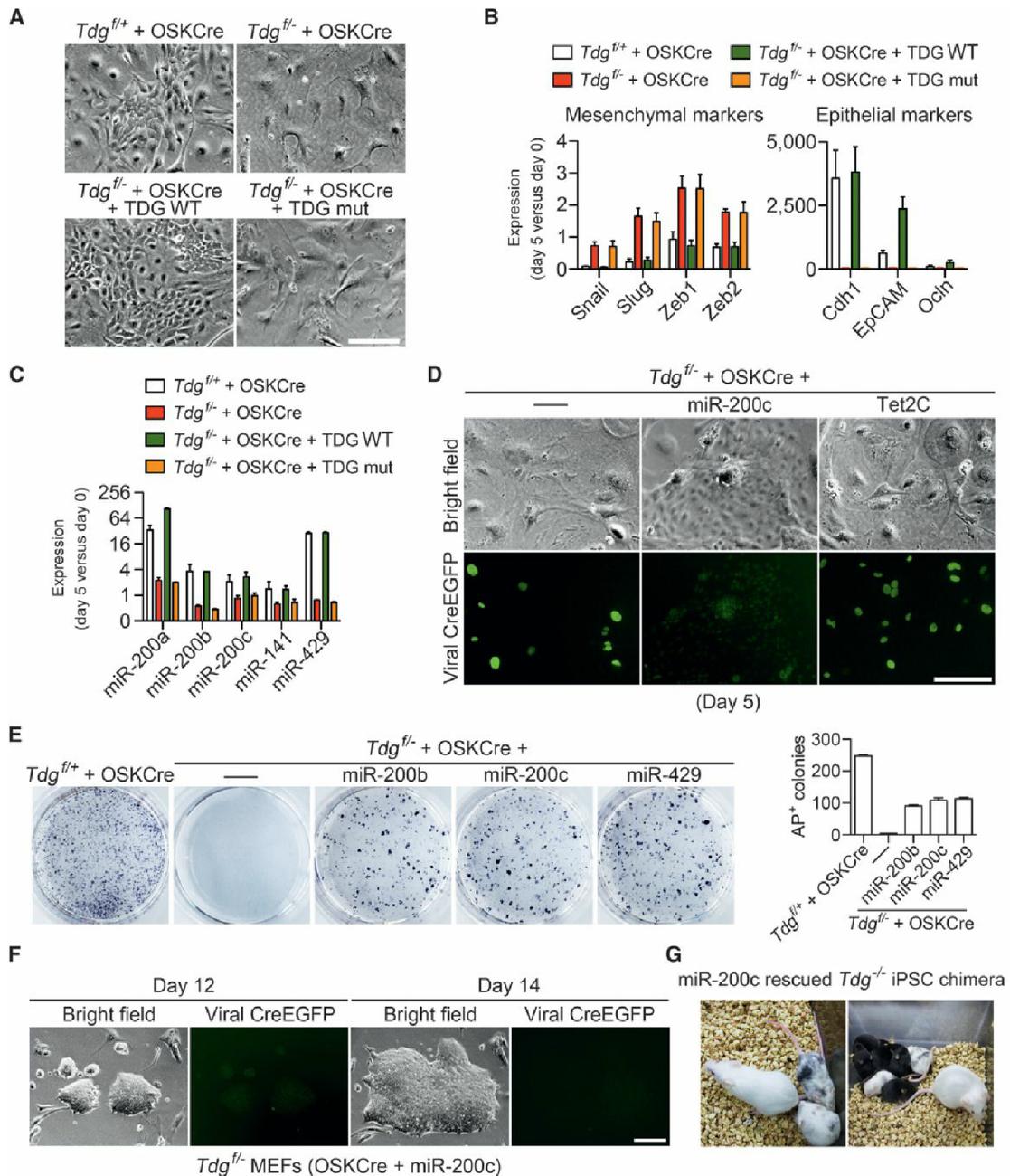
for iPSC induction. The Cre-mediated deletion of *Tdg* did not affect the proliferation of MEFs (Figure S5B). Upon transduction with Oct4, Sox2, Klf4, and Cre recombinase, the control MEFs with a wild-type *Tdg* allele in addition to a floxed allele showed a normal reprogramming capacity. However, MEFs carrying a null allele together with a floxed allele did not result in any iPSCs positive for AP and SSEA1 (Figures 5A and 5B). Re-expression of wild-type TDG, but not the catalytic mutant (N151A; Figure S5C), restored full reprogramming capacity (Figure 5).

Similar to the Tet TKO MEFs, TDG-null MEFs did not initiate the MET as judged by morphology and mesenchymal/epithelial marker expression (Figures 6A and 6B). Likewise, the blockage of MET initiation in *Tdg*-deficient MEFs could be ascribed to the failed activation of regulatory miRNAs (Figure 6C), and the MET resumed upon forced expression of miR-200c, but not Tet2C (Figure 6D). As evidenced by the formation of AP-positive colonies, robust reprogramming proceeded in the absence of TDG upon rescue with miR-200c or other two miRNAs, miR-200b and miR-429, with an efficiency of above 30% of the *Tdg*-proficient wild-type cells (Figure 6E). iPSC lines could be established from rescued *Tdg* KO MEFs (Figure 6F), and they were pluripotent as demonstrated by the generation of high-grade chimeric mice (Figure 6G). Contrary to the drastic effect on MEFs, *Tdg* deletion did not prevent keratinocyte and NPCs from reprogramming (Figures S6A–S6D). Cre-mediated late deletion from *Tdg*-floxed MEFs did not seem to impair the generation of AP-positive colonies whereas early deletion prior to MET led to no single colony (Figures S6E and S6F).

The data above led us to conclude that the TDG is required for MET by contributing to the reactivation of the critical miRNA genes in the early phase of reprogramming. Because the epigenetic regulation, MET, and iPSC formation in TDG-null MEFs all depend on the catalytic activity of TDG, we infer that the formation and excision of 5fC and 5caC are likely a critical step that at least partially contributes to the demethylation process.

### Coordinated Action of Tet and TDG in Oxidative DNA Demethylation

In order to dissect the demethylation pathway underlying miRNA reactivation, we attempted to detect the higher oxidation



**Figure 6. TDG Deficiency Blocks MET**

(A) Lack of MET in MEFs depleted of TDG and rescue of MET upon re-expression of a functional TDG. Representative phase-contrast photographs of transduced MEFs (day 5) are shown. The scale bar represents 100  $\mu$ m.

(B) Effect of TDG on the expression of mesenchymal and epithelial markers in MEFs under reprogramming at day 5. Expression levels were normalized to *Gapdh*. Data represent means  $\pm$  SD of three independent experiments.

(C) Effect of TDG on the expression of the miR-200 family miRNAs in MEFs under reprogramming at day 5. The expression levels were normalized to *Rnu6*. Data represent means  $\pm$  SD of three independent experiments.

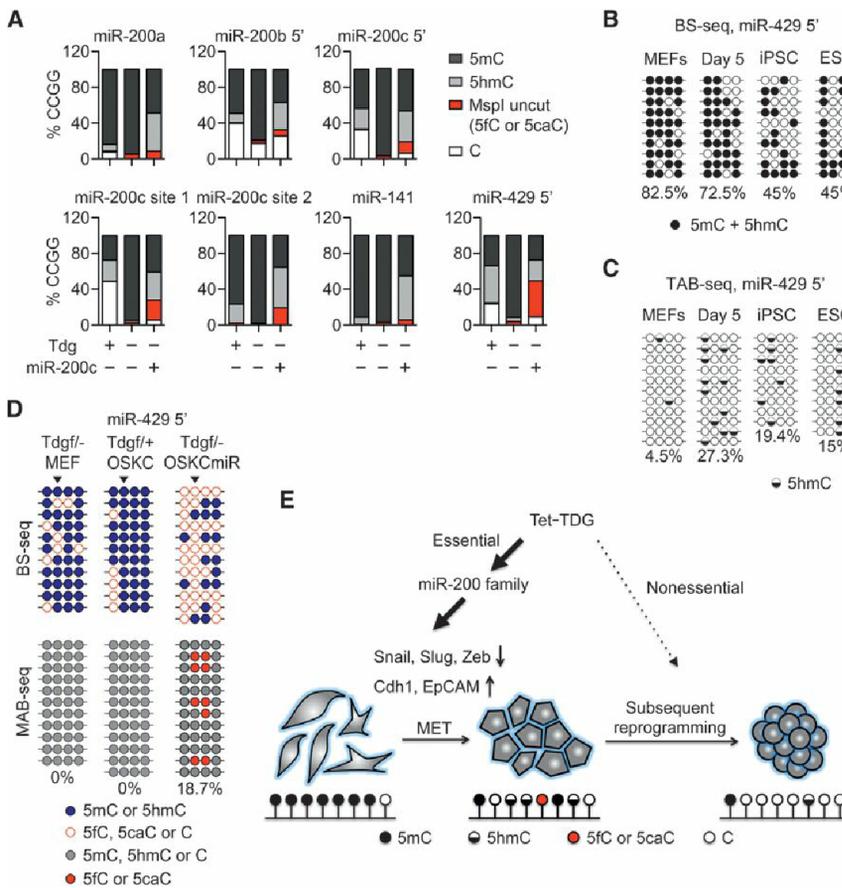
(D) Restoration of MET in TDG-depleted MEFs upon ectopic expression of miR-200c. Note that the retroviral Cre-EGFP fusion gene is still weakly expressed at day 5 in intermediate MEFs rescued for MET by ectopic miR-200c expression. The scale bar represents 100  $\mu$ m.

(E) AP staining of day 10 colonies arising from *Tdg<sup>fl/fl</sup>*, *Tdg<sup>fl/-</sup>*, and *Tdg<sup>-/-</sup>* MEFs transduced with OSKCre with or without rescue by the indicated miRNA. AP-positive colony numbers are shown on the right. Data represent means  $\pm$  SD of two independent experiments.

(F) Representative photographs of colonies induced from TDG-depleted MEFs rescued with the ectopic expression of miR-200c (days 12 and 14). Note the silencing of the retroviral CreEGFP fusion gene. The scale bar represents 100  $\mu$ m.

(G) Photographs of chimeras generated from miR-200c-rescued *Tdg<sup>-/-</sup>* iPSCs. Rescued *Tdg<sup>-/-</sup>* iPSCs on a B6 genetic background were injected into ICR blastocysts.

See also Figure S6.



**Figure 7. Dynamic Methylation Reprogramming through Coupled Action of Tet and TDG during Reprogramming**

(A) GlucMS-qPCR analysis of cytosine modifications at the miR-200 family loci in intermediate cells from wild-type (+) and *Tdg*-deficient (-) MEFs rescued with (+) and without (-) ectopic miR-200c at reprogramming day 5. The level of combined 5fC and 5caC at a selected C(C)GG site was derived from the amount of DNA resistant to *MspI* cleavage (*MspI* uncut). (B) Profiles of 5mC and 5hmC in the 5' region of miR-429 during iPSC induction. Note that a significant proportion of the black circles ought to be 5hmC (compare to C), as 5hmC and 5mC are indistinguishable in this bisulfite-sequencing (BS-seq) analysis. Day 5 cells were unsorted population undergoing reprogramming. (C) Profiles of 5hmC revealed by TAB-seq. Day 5 cells were unsorted population. (D) MAB sequencing detection of 5fC and 5caC (red) in a 5' region of miR-429 at reprogramming day 5 in indicated MEFs transduced with Cre (C) and with Cre-miR-200c (CmiR), respectively. BS-seq profiles are shown on the top for comparison. Untransduced *Tdg*<sup>-/-</sup> MEFs served as a control. In MAB-seq, 5fC and 5caC, which are resistant to *M.SssI* methylation in vitro but sensitive to the subsequent bisulfite conversion, are read as "C" whereas all other forms (5mC, 5hmC, and C) are read as "5mC". Arrowheads indicate the selected C(C)GG site analyzed in (A). (E) Proposed role of oxidative demethylation during somatic cell reprogramming. Tet- and TDG-mediated demethylation is required for the reactivation of miRNA genes critical for MET at the initiation phase of reprogramming. See also Figure S7.

derivatives of 5mC in reprogramming MEFs. We took advantage of the observation that 5fC/5caC modifications, but not 5mC and 5hmC, block *MspI* (CCGG) restriction digestion (He et al., 2011; Ito et al., 2011) to estimate the frequency of 5fC/5caC at the specific *MspI* sites located in the miR-200 loci (Figure 4A). Whereas TDG-proficient intermediate cells and unrescued TDG-deficient cells lacked detectable 5fC/5caC, these modified species were readily detected in TDG-deficient cells rescued for reprogramming with ectopic miR-200c (Figure 7A). The abundance of 5fC/5caC varied among different loci, with the highest level detected in the 5' region of miR-429. Like those of miR-200b and miR-200c (Figures 7B–7D), the 5' region of miR-429 also exhibited hypermethylation in starting MEFs (Figure 7B) and methylation reduction while gaining 5hmC in reprogramming intermediates and iPSCs (Figure 7C).

In order to provide independent validation for 5fC/5caC deposition, we used a method for base-resolution profiling, termed methylation-assisted bisulfite sequencing (MAB-seq). The miR-429 region was selected due to the relative abundance of 5fC/5caC found at a CpG site in *Tdg*-null reprogramming cells (Figure 7A). MAB sequencing revealed a considerable amount (~19%) of 5fC and 5caC in TDG-null cells undergoing reprogramming upon miRNA rescue (Figure 7D). No 5fC/5caC was detected in corresponding TDG-proficient intermediate cells or control MEFs harboring an intact wild-type or a floxed *Tdg* allele.

As the total level of 5fC, 5caC, and unmodified cytosines (C) in this region determined by conventional bisulfite profiling was 64.6%, C could be inferred to account for 45.9%. Compared to the high levels of combined 5mC and 5hmC in MEFs (around 80%; Figures 7B–7D), Tet-promoted 5mC oxidation in the rescued reprogramming *Tdg*-deficient cells had not only generated 5fC and 5caC accumulation but had also led to more unmodified cytosines, presumably by replication-dependent passive dilution of the resultant oxidation products.

In summary, the higher oxidation derivatives appeared dynamic in intermediate reprogramming stages primarily due to constant removal by TDG as well as passive dilution. They accumulated to a certain degree in TDG-knockout cells rescued for reprogramming by ectopic miR-200c. Tet-mediated oxidation of 5mC coupled with TDG-initiated base excision occurs in intermediate cells undergoing reprogramming, leading to the demethylation and reactivation of miRNA genes critical for MET (Figure 7E).

## DISCUSSION

The physiological significance of DNA oxidation in epigenetic regulation has remained poorly understood, especially with regard to its importance in lineage commitment and cell reprogramming. Whereas previous reports have suggested the

involvement of Tet-mediated hydroxylation in somatic cell reprogramming (Costa et al., 2013; Doege et al., 2012; Gao et al., 2013), our work indicates that none of the individual Tet knockouts abolish the capacity of MEFs to be reprogrammed. The requirement for DNA oxidation can only be demonstrated by eliminating all three Tet genes simultaneously. Deletion of the glycosylase TDG also renders MEFs incapable of reprogramming due to an impediment to epigenetic activation of miRNAs crucial to mesenchymal-to-epithelial transition, as found in Tet-deficient MEFs.

DNA demethylation can proceed either by active removal or the passive dilution of methylcytosines and its derivatives, as a result of DNA replication. Recent studies document a prominent role of replication-coupled passive demethylation in cell-fusion-based reprogramming (Tsubouchi et al., 2013), early embryonic development (Inoue and Zhang, 2011), as well as primordial germ cell development (Hackett et al., 2012; Kagiwada et al., 2013; Seisenberger et al., 2012; Vincent et al., 2013). Contrary to these observations, we demonstrate that both Tet and TDG are essential for miRNA gene demethylation and factor-induced cell reprogramming, in support of an active mode of DNA demethylation. TDG, initially identified as a thymine DNA glycosylase, is not required for DNA repair but rather plays an important role in epigenetic regulation (Cortázar et al., 2011; Cortellino et al., 2011). A role in active demethylation has been proposed for TDG, based on its capability to specifically recognize and excise the Tet-generated higher oxidation products 5fC and 5caC (He et al., 2011; Maiti and Drohat, 2011). While still lacking evidence in a physiological reprogramming setting, we suppose that Tet and TDG-mediated demethylation is likely to play a locus-specific role whereas replication-mediated passive demethylation is more prevalent in settings where genome-wide erasure has to occur efficiently.

In both experimental and natural reprogramming models, DNA demethylation has been intimately linked to the activation of pluripotency loci. In particular, cell fusion and iPSC induction experiments have implicated Tet-mediated hydroxylation in the epigenetic reactivation of silent pluripotency genes, a perceived bottleneck in the path toward the establishment of pluripotency (Doege et al., 2012; Piccolo et al., 2013). Oocyte Tet3 provides a reprogramming activity for pluripotency gene reactivation during the early embryonic development after nuclear transfer and natural fertilization (Gu et al., 2011). In this work, we also detect Tet-mediated 5hmC deposition at pluripotency loci in addition to other genes, including those related to cell adhesion, during iPSC induction of mouse fibroblasts (Figures S4 and S7). Interestingly, our genetic ablation studies indicate that Tet function is not essential for the demethylation and reactivation of pluripotency genes, as iPSCs can still be generated from Tet-deficient MEFs upon the rescue of MET by a single miRNA. Rather, Tet enzymes have an indispensable role in promoting demethylation and reactivation of miRNAs critical for the fibroblasts to be converted into epithelia at the onset of reprogramming, thus breaking down an epigenetic barrier imposed on the regulatory miRNA genes (Figure 7E). As Tet- or TDG-deficient cells expressing ectopic miRNAs are amenable to reprogramming, the activation of pluripotency genes presumably encompasses passive demethylation involving DNA replication. It is likely that both active (Tet-mediated) and passive (replication-dependent) de-

methylation mechanisms contribute to pluripotency gene demethylation for the reprogramming of wild-type cells.

Deletion of individual Tet genes, or Tet1 and Tet2 in combination, in mouse is compatible with embryonic development (Dawlaty et al., 2011, 2013; Gu et al., 2011). In light of the functional redundancy of the Tet genes as revealed in this experimental reprogramming system, the assessment of the biological significance of DNA oxidation in development awaits the generation and examination of triple knockout animals. Interestingly, a report by Song et al. (2013) published during the revision of this manuscript implicates the downregulation of Tet activity in the epigenetic inactivation of miR-200, which supposedly contributes to EMT in the process of mammary tumorigenesis and metastasis. It remains to be addressed whether and to what extent Tet- and TDG-mediated DNA demethylation control MET-related miRNAs and other important developmental regulators in physiological and disease processes.

## EXPERIMENTAL PROCEDURES

### Animal Use and Care

Animal procedures were carried out according to the ethical guidelines of the Institute of Biochemistry and Cell Biology.

### Derivation of Tet-Deficient MEFs

To prepare Tet1, Tet2, Tet3 triple knockout (TKO) MEFs, Tet TKO ESCs were labeled with constitutively expressing lentiviral GFP and then injected into mouse blastocysts to obtain chimeric E12.5 embryos. MEFs were isolated, and the GFP-positive cells were sorted by fluorescence-activated cell sorting (FACS). More information is available in Supplemental Experimental Procedures.

### Retroviral Production and iPSC Induction

Retroviral production and infection followed the previously published protocol (Takahashi and Yamanaka, 2006). For retroviral production, Plat-E cells were seeded at  $7 \times 10^6$  cells per 100 mm dish 1 day before transfection. Nine micrograms of pMXs-based retroviral constructs were transfected into Plat-E cells using Lipofectamine 2000 reagents (Invitrogen) according to the manufacturer's recommendations. Eight to ten hours later, the medium was replaced. Another 48 hr later, virus-containing supernatants were collected and filtered through a 0.45  $\mu$ m polyvinylidene fluoride filter (Millipore) and supplemented with 4  $\mu$ g/ml polybrene (Sigma). MEFs (seeded at  $5 \times 10^4$  cells per each well in a 6-well plate 1 day before infection) were incubated with virus-containing supernatants for 12 hr. After two rounds of infection, cells were replated onto mitomycin-C-treated MEF feeder layers and the medium was changed into optimized medium (Chen et al., 2011). GFP-positive or alkaline-phosphatase-positive colonies were scored. Alkaline phosphatase staining was performed with NBT/BCIP (Roche).

In the Tet TKO MEFs rescue experiments using Tet1C, Tet2C, or Tet3C, cells were cultured with optimized medium without vitamin C.

For the establishment of miR-200c-rescued Tet TKO and Tdg KO iPSC lines, Tet TKO and Tdg fl/fl MEFs were transduced with OSK/miR-200c and OSKCre/miR-200c retroviruses, respectively. Cells under induction were cultured in optimized medium. iPSC colonies were picked, transferred onto feeder layers, and cultured in ESC medium supplemented with 2i and leukemia inhibitory factor.

### Bisulfite Sequencing and TAB-Seq

For bisulfite sequencing, 100 ng of genomic DNA were treated with the EZ DNA Methylation-Direct Kit (Zymo Research). For Tet-assisted bisulfite sequencing (TAB-seq), genomic DNA was treated according to the protocol published recently (Yu et al., 2012a). Briefly, 1  $\mu$ g of genomic DNA was sonicated into fragments of 300 bp to 1 kb in size. The fragments were then glucosylated, oxidized with recombinant mouse Tet1C, and treated with bisulfite sequentially. Specific genomic regions were PCR amplified and cloned into pMD-19T (Takara) for sequencing.

**GlucMS-qPCR**

Genomic DNA was extracted with the DNeasy Blood and Tissue kit (QIAGEN). Glucosylated hydroxymethyl-sensitive quantitative PCR (GlucMS-qPCR) was performed using the EpiMark 5-hmC and 5-mC Analysis Kit (New England Biolabs).

**MAB-Seq**

One microgram of genomic DNA was methylated by M.SssI (NEB) following the NEB's instruction. Complete methylation of the DNA was confirmed by restriction analysis. The methylated DNA was then purified by phenol-chloroform extraction and bisulfite sequencing performed as described above. The methodology will be described elsewhere.

**Flow Cytometry**

Single-cell suspensions were obtained by repetitive pipetting and filtered through a 40  $\mu$ m cell strainer. Cells were incubated with anti-mouse SSEA-1 PE (eBioscience) and analyzed on a FACS Aria (BD Biosciences). Data were analyzed with FlowJo software (Tree Star).

**RNA Extraction, Reverse Transcription, and qPCR**

Total RNA was extracted from cells with Trizol reagent (Life Technologies) and reverse-transcribed using the PrimeScript RT Reagent Kit with a gDNA Eraser according to manufacturer's instructions (Takara). Quantitative real-time PCR was performed using SYBR Premix EX Taq (Takara) on Bio-Rad CFX96. For the quantification of miRNA expression, total RNA was reverse-transcribed with the PrimeScript miRNA qPCR Starter Kit (Takara), and qPCR was performed following the vendor's instructions.

**ACCESSION NUMBERS**

The Gene Expression Omnibus accession number for the RRBS and TA-RRBS data reported in this paper is GSE52741.

**SUPPLEMENTAL INFORMATION**

Supplemental Information for this article includes Supplemental Experimental Procedures and seven figures and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2014.01.001>.

**ACKNOWLEDGMENTS**

We thank Jiemin Wong, Pentao Liu, and Colum Walsh for critical reading of the manuscript. This work was supported by grants from the Ministry of Sciences and Technology of China (2012CB966903 and 2014CB964802), the National Science Foundation of China (31230039 and 31221001), the "Strategic Priority Research Program" of the Chinese Academy of Sciences (XDA01010301), and the National Science and Technology Major Project "Key New Drug Creation and Manufacturing Program" of China (grant number 2014ZX09507-002) to G.-L.X.

Received: April 22, 2013

Revised: November 25, 2013

Accepted: December 26, 2013

Published: February 13, 2014

**REFERENCES**

Bagci, H., and Fisher, A.G. (2013). DNA demethylation in pluripotency and reprogramming: the role of tet proteins and cell division. *Cell Stem Cell* 13, 265–269.

Chen, J., Liu, J., Chen, Y., Yang, J., Chen, J., Liu, H., Zhao, X., Mo, K., Song, H., Guo, L., et al. (2011). Rational optimization of reprogramming culture conditions for the generation of induced pluripotent stem cells with ultra-high efficiency and fast kinetics. *Cell Res.* 21, 884–894.

Chen, J., Guo, L., Zhang, L., Wu, H., Yang, J., Liu, H., Wang, X., Hu, X., Gu, T., Zhou, Z., et al. (2013). Vitamin C modulates TET1 function during somatic cell reprogramming. *Nat. Genet.* 45, 1504–1509.

Cortázar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F., et al. (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* 470, 419–423.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., et al. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146, 67–79.

Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shliha, P.V., Fidalgo, M., Saunders, A., Lawrence, M., Dietmann, S., et al. (2013). NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* 495, 370–374.

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.W., Page, D.C., and Jaenisch, R. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* 9, 166–175.

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M., et al. (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev. Cell* 24, 310–323.

Doege, C.A., Inoue, K., Yamashita, T., Rhee, D.B., Travis, S., Fujita, R., Guarnieri, P., Bhagat, G., Vanti, W.B., Shih, A., et al. (2012). Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* 488, 652–655.

Epsztejn-Litman, S., Feldman, N., Abu-Remaileh, M., Shufaro, Y., Gerson, A., Ueda, J., Deplus, R., Fuks, F., Shinkai, Y., Cedar, H., and Bergman, Y. (2008). De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* 15, 1176–1183.

Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y., et al. (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* 12, 453–469.

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* 74, 481–514.

Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y., and Goodall, G.J. (2008). The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* 10, 593–601.

Gu, T.P., Guo, F., Yang, H., Wu, H.P., Xu, G.F., Liu, W., Xie, Z.G., Shi, L., He, X., Jin, S.G., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* 477, 606–610.

Gurdon, J.B., and Melton, D.A. (2008). Nuclear reprogramming in cells. *Science* 322, 1811–1815.

Hackett, J.A., Sengupta, R., Zylicz, J.J., Murakami, K., Lee, C., Down, T.A., and Surani, M.A. (2012). Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* 339, 448–452.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333, 1303–1307.

Inoue, A., and Zhang, Y. (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* 334, 194.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333, 1300–1303.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet. Suppl.* 33, 245–254.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., and Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J.* 32, 340–353.

Li, J.Y., Pu, M.T., Hirasawa, R., Li, B.Z., Huang, Y.N., Zeng, R., Jing, N.H., Chen, T., Li, E., Sasaki, H., and Xu, G.L. (2007). Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. *Mol. Cell. Biol.* 27, 8748–8759.

- Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q., et al. (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 7, 51–63.
- Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.* 286, 35334–35338.
- Pastor, W.A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell Biol.* 14, 341–356.
- Piccolo, F.M., Bagci, H., Brown, K.E., Landeira, D., Soza-Ried, J., Feytout, A., Mooijman, D., Hajkova, P., Leitch, H.G., Tada, T., et al. (2013). Different roles for Tet1 and Tet2 proteins in reprogramming-mediated erasure of imprints induced by EGC fusion. *Mol. Cell* 49, 1023–1033.
- Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151, 1617–1632.
- Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 7, 64–77.
- Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., and Reik, W. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* 48, 849–862.
- Seisenberger, S., Peat, J.R., Hore, T.A., Santos, F., Dean, W., and Reik, W. (2013). Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368, 20110330.
- Song, S.J., Polisenio, L., Song, M.S., Ala, U., Webster, K., Ng, C., Beringer, G., Brikbak, N.J., Yuan, X., Cantley, L.C., et al. (2013). MicroRNA-antagonism regulates breast cancer stemness and metastasis via TET-family-dependent chromatin remodeling. *Cell* 154, 311–324.
- Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., and Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935.
- 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.
- Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. *Cell* 139, 871–890.
- Tsubouchi, T., Soza-Ried, J., Brown, K., Piccolo, F.M., Cantone, I., Landeira, D., Bagci, H., Hochegger, H., Merkschlager, M., and Fisher, A.G. (2013). DNA synthesis is required for reprogramming mediated by stem cell fusion. *Cell* 152, 873–883.
- Vincent, J.J., Huang, Y., Chen, P.Y., Feng, S., Calvopiña, J.H., Nee, K., Lee, S.A., Le, T., Yoon, A.J., Faull, K., et al. (2013). Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. *Cell Stem Cell* 12, 470–478.
- Wang, Y., Chen, J., Hu, J.L., Wei, X.X., Qin, D., Gao, J., Zhang, L., Jiang, J., Li, J.S., Liu, J., et al. (2011). Reprogramming of mouse and human somatic cells by high-performance engineered factors. *EMBO Rep.* 12, 373–378.
- Wang, G., Guo, X., Hong, W., Liu, Q., Wei, T., Lu, C., Gao, L., Ye, D., Zhou, Y., Chen, J., et al. (2013). Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. *Proc. Natl. Acad. Sci. USA* 110, 2858–2863.
- Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Jin, P., Ren, B., and He, C. (2012a). Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nat. Protoc.* 7, 2159–2170.
- Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., et al. (2012b). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* 149, 1368–1380.

# Targeting Self-Renewal in High-Grade Brain Tumors Leads to Loss of Brain Tumor Stem Cells and Prolonged Survival

Zhe Zhu,<sup>1</sup> Muhammad Amir Khan,<sup>1</sup> Markus Weiler,<sup>2,4</sup> Jonas Blaes,<sup>2</sup> Leonie Jestaedt,<sup>5</sup> Madeleine Geibert,<sup>1</sup> Peng Zou,<sup>1</sup> Jan Gronych,<sup>3</sup> Olga Bernhardt,<sup>1</sup> Andrey Korshunov,<sup>6</sup> Verena Bugner,<sup>3</sup> Peter Lichter,<sup>3</sup> Bernhard Radlwimmer,<sup>3</sup> Sabine Heiland,<sup>5</sup> Martin Bendszus,<sup>5</sup> Wolfgang Wick,<sup>2,4</sup> and Hai-Kun Liu<sup>1,\*</sup>

<sup>1</sup>Helmholtz Young Investigator Group, Normal and Neoplastic CNS Stem Cells, DKFZ-ZMBH Alliance, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

<sup>2</sup>Clinical Cooperation Unit Neurooncology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

<sup>3</sup>Division of Molecular Genetics, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

<sup>4</sup>Department of Neurooncology, University Clinic Heidelberg and German Cancer Consortium (DKTK), Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

<sup>5</sup>Department of Neuroradiology, Heidelberg University Hospital, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

<sup>6</sup>Department of Neuropathology, Heidelberg University Hospital, Im Neuenheimer Feld 220, 69120 Heidelberg, Germany

\*Correspondence: l.haikun@dkfz-heidelberg.de

<http://dx.doi.org/10.1016/j.stem.2014.04.007>

## SUMMARY

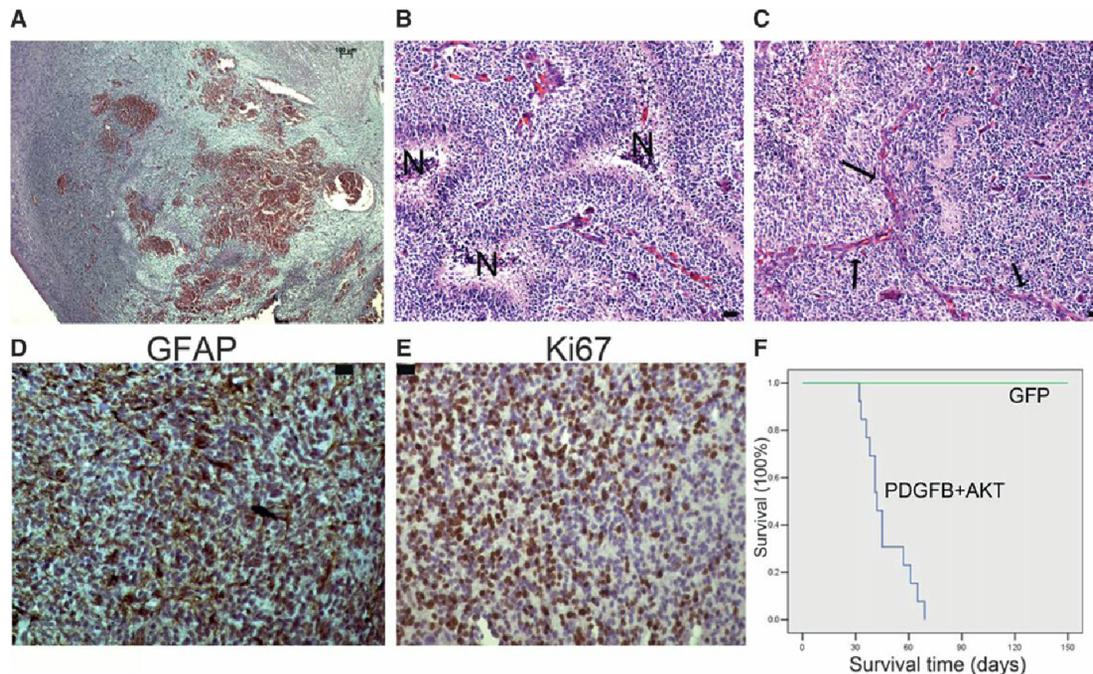
Cancer stem cells (CSCs) have been suggested as potential therapeutic targets for treating malignant tumors, but the *in vivo* supporting evidence is still missing. Using a GFP reporter driven by the promoter of the nuclear receptor *tailless* (*Tlx*), we demonstrate that *Tlx*<sup>+</sup> cells in primary brain tumors are mostly quiescent. Lineage tracing demonstrates that single *Tlx*<sup>+</sup> cells can self-renew and generate *Tlx*<sup>-</sup> tumor cells in primary tumors, suggesting that they are brain tumor stem cells (BTSCs). After introducing a BTSC-specific knock-out of the *Tlx* gene in primary mouse tumors, we observed a loss of self-renewal of BTSCs and prolongation of animal survival, accompanied by induction of essential signaling pathways mediating cell-cycle arrest, cell death, and neural differentiation. Our study demonstrates the feasibility of targeting glioblastomas and indicates the suitability of BTSCs as therapeutic targets, thereby supporting the CSC hypothesis.

## INTRODUCTION

The cancer stem cell (CSC) hypothesis provides an alternative model to explain the tumor cell heterogeneity (Reya et al., 2001). Like normal somatic adult stem cells, which are distributed in different tissues, CSCs are thought to be the less differentiated populations in malignant tissues and are considered to be the cells that are responsible for the maintenance of tumor tissues, as well as for the relapse of tumors after conventional treatment. In most tumor entities, the existence of CSCs has not been demonstrated *in vivo* with the genetic cell fate mapping approach, which is the most stringent criterion for the identification of normal tissue stem cells (Clevers, 2011). Using lineage

tracing, a very recent study performed in a mouse colon adenoma model showed that *Lgr5*-expressing cells are CSCs (Schipers et al., 2012). This study provided the first solid evidence for the presence of CSCs in unperturbed primary tumors. The CSC hypothesis rises the anticipation that targeting of CSCs in tumors will lead to an improved clinical outcome because they are thought to be the “root” of growing tumors. However, the *in vivo* experimental evidence supporting that is still missing in most tumor entities.

Brain tumor stem cells (BTSC) are one of the first CSCs identified in solid tumors (Singh et al., 2003, 2004). CD133 has been suggested to be a useful marker for isolation of BTSCs but the results obtained based on this method are controversial (Clevers, 2011). BTSCs share many similarities with normal neural stem cells (NSCs), i.e., expression of markers like Nestin and CD133, or sphere-forming ability when placed into medium containing growth factors (Vescovi et al., 2006). It would not be surprising if BTSCs could hijack the self-renewal pathway of normal NSCs, and similar results have been found in many other type of tumors (He et al., 2009). Several studies tried to target factors that are known to be important of NSC maintenance during brain tumorigenesis. Inhibitor of DNA binding 1 (*Id1*) is selectively expressed in NSCs in adult neurogenic niches, namely the subventricular zone (SVZ) and the subgranular zone (SGZ) and *Id* proteins are shown to be involved in regulation of stem cell self-renewal (Nam and Benezra, 2009). However, deletion of *Id1* in mouse brain tumors has modest effects on animal survival, although BTSCs of *Id1* mutants have reduced self-renewal capacity *in vitro* (Barrett et al., 2012). This leads to the assumption that the self-renewal of BTSCs does not predict brain tumor growth potential, however, it is noteworthy that *Id1* mutant animals have normal NSCs populations and normal neurogenesis (Barrett et al., 2012). A recent study using a cell ablation approach demonstrated that chemo-resistant Nestin expressing brain tumor cells are responsible for brain tumor propagation, and ablation of these cells led to prolonged survival of tumor-bearing mice (Chen et al., 2012), which support that self-renewing tumor



**Figure 1. Induction of High-Grade Brain Tumors in *Ntv-a* Mice**

(A) Hematoxylin and eosin (H&E) staining of GB induced with PDGFB and AKT in the Nestin-Tv-a model. Scale bar represents 100  $\mu$ m.

(B and C) Brain tumors induced with PDGFB and AKT have typical pathological features like pseudopalisading and necrosis (B, N indicates the necrotic areas) and vascular proliferation (C, arrows). Scale bar represents 50  $\mu$ m.

(D and E) The mouse tumors highly express the astrocyte marker GFAP (D) and the proliferation marker Ki67 (E). Scale bar represents 20  $\mu$ m.

(F) Kaplan-Meier survival curve of the *Ntv-a* mice induced with PDGFB and AKT ( $n = 15$ ). RCAS-GFP was used as control vector ( $n = 12$ ).

See also Figure S1.

cells are important therapeutic targets of brain tumors in vivo. The nuclear receptor tailless (Tlx, Nr2e1) is specifically expressed in adult NSCs and loss of Tlx in mice leads to the loss of self-renewal of NSCs, both in the SVZ and the SGZ (Liu et al., 2008, 2010; Shi et al., 2004; Zhang et al., 2008; Zou et al., 2012). Interestingly, Tlx is also overexpressed in human brain tumors and it is important for tumor initiation from the SVZ (Liu et al., 2010; Zou et al., 2012). These results suggest that Tlx is a crucial regulator for self-renewal of adult NSCs and brain tumorigenesis. However, the expression pattern and function of Tlx in established tumors is unknown.

To investigate whether BTSCs are important for tumor maintenance and survival, we generated a mouse brain tumor model, which allows cell-type-specific genetic manipulation in fully developed primary brain tumors. We found that the Tlx<sup>+</sup> cells are the slow-dividing and undifferentiated tumor cells in vivo, xenotransplantation, and lineage tracing directly demonstrates that they are the BTSCs. An inducible inactivation of Tlx in Nestin expressing brain tumor cells leads to a significant prolongation of survival of tumor-bearing animals. Tumor cells from Tlx-deficient tumors lost self-renewing capacity and resulted in less proliferating cells. In addition, we identified several important signaling pathways involved in cell-cycle arrest, cell death, and neural differentiation. Overall, this study demonstrates the direct genetic evidence of the importance of CSC in brain tumors and provides an attractive therapeutic target of high-grade gliomas.

## RESULTS

### A Mouse Model for Gene Targeting in Fully Developed High-Grade Brain Tumors

To establish a mouse model that allows gene targeting in established brain tumors, we made use of the RCAS/tv-a system to initiate brain tumors. The RCAS/tv-a system uses an avian retroviral vector, called RCAS, derived from avian leukosis virus (ALV, subgroup A), and a Nestin-tv-a (*Ntv-a*) transgenic mouse line that expresses Tv-a (the receptor for ALV-A) under the control of the *Nestin* gene promoter. This system makes it possible to transfer and express exogenous genes in Nestin-expressing neural progenitor cells and their progeny (Holland, 2001). With coinjection of RCAS viruses expressing AKT and platelet-derived growth factor B (PDGFB), we were able to induce high-grade brain tumors with full penetrance in *Ntv-a* mice (Figure 1A). Approximately 60% of the tumors were multifocal glioblastoma (GB) (Table S1 available online), which have classic histopathologic features of human GB, like pseudopalisading-associated necrosis and vascular proliferation (Figures 1B and 1C). These tumors express the astrocyte marker glial filament acid protein (GFAP) and the cell proliferation marker Ki67 (Figures 1D and 1E). All animals developed brain tumors within 3 months after injection of PDGFB and AKT (Figure 1F). To determine the cell of origin of brain tumors in this model, we injected the RCAS-GFP producing DF-1 cells into the *Ntv-a* mice, which allow visualization of infected cells with GFP signal. Three days after

injection, we found that most of GFP cells are located in the SVZ area and all express GFAP and Nestin (Figures S1A and S1B), suggesting the radial glia-like NSCs are the cells initially infected. It is known that the SVZ NSCs mainly give rise to newborn neurons in the olfactory bulb. One month after injection, we found many GFP<sup>+</sup> cells in the olfactory bulb (Figure S1C), which further demonstrate that the NSCs are targeted by RCAS approach. This is consistent with the working principle of this model that only the Nestin-tv-a-expressing cells are infected by the RCAS vectors (Holland et al., 1998). We also analyzed the tumor location in many animals (Table S1). Most of the tumors are multifocal, but there is no strong correlation between tumor grade and location. Nevertheless, SVZ is the most frequently associated region of tumor locations (Table S1), again suggesting the stem cell origin of these tumors. This mouse model allows induction of tumors without involving Cre/LoxP technology, thus the Cre/LoxP system can be used for additional genetic manipulations in mouse primary brain tumors. We then generated mouse models with the following genotypes, Ntv-a;Tlx-GFP and Ntv-a; Nestin-CreERT2;Tlx<sup>flox/flox</sup>. As we have shown in Tlx-GFP transgenic mice, in which the GFP expression is driven by a Tlx bacterial artificial chromosome (BAC)-based promoter, that NSCs are GFP<sup>+</sup> (Feng et al., 2013), we are able to visualize the Tlx expressing brain tumor cells via the GFP signal. Furthermore, the Ntv-a;Nestin-CreERT2;Tlx<sup>flox/flox</sup> mice can be used to induce Tlx inactivation in PDGFB/AKT-induced brain tumors, in a tamoxifen (TMX)-dependent manner.

### Tlx<sup>+</sup> Cells in Mouse Brain Tumors Are Slow-Dividing Cells

After induction of tumors in Tlx-GFP;Ntv-a mice, we observed heterogeneous expression of GFP in mouse brain tumors (Figure 2A). To further characterize the Tlx-GFP tumor cells *in vivo*, we performed additional immunohistochemistry (IHC) analysis, and we found that the GFP<sup>+</sup> cells were negative for Olig2 (Figure 2B), a common glioma marker and an oligodendrocyte marker in normal brain tissues. We found that some of the GFP<sup>+</sup> cells were also positive for GFAP, which is a marker for mature astrocyte and radial glia like NSCs in adult brain (Figure 2C). GFP<sup>+</sup> cells were also negative for doublecortin (DCX), a marker for immature neurons (Figure 2D). We found that a commonly used marker for NSCs, NESTIN, is widely expressed in the mouse tumors and that the Tlx-GFP<sup>+</sup> cells are just a subpopulation of NESTIN<sup>+</sup> cells (Figure 2E). It has been suggested that CSCs are slow-dividing cells that is one of the major features of many tissue-specific stem cells (Li and Clevers, 2010). By costaining of GFP with the cell proliferation marker Ki67 and another widely used stem cell marker, Sox2, we found that most of the GFP cells were negative for Ki67 (Figure 2F). Interestingly, the majority of the Sox2 cells were negative for Tlx-GFP but positive for Ki67 (Figure 2F). We further quantified the relative distribution of Tlx-GFP cells comparing with three markers (Sox2, GFAP, NESTIN), the result indicates that Tlx-GFP marks a unique population in primary brain tumors (Figure 2G). Using additional cell proliferation markers like proliferating cell nuclear antigen (PCNA) (Figure S2A) or MCM2 (Figure S2B) further confirmed that Tlx<sup>+</sup> cells are largely quiescent (Figure 2H). This result suggests that the Tlx-GFP<sup>+</sup> tumor cells are not fast-dividing cells *in vivo*, and it is known that most of

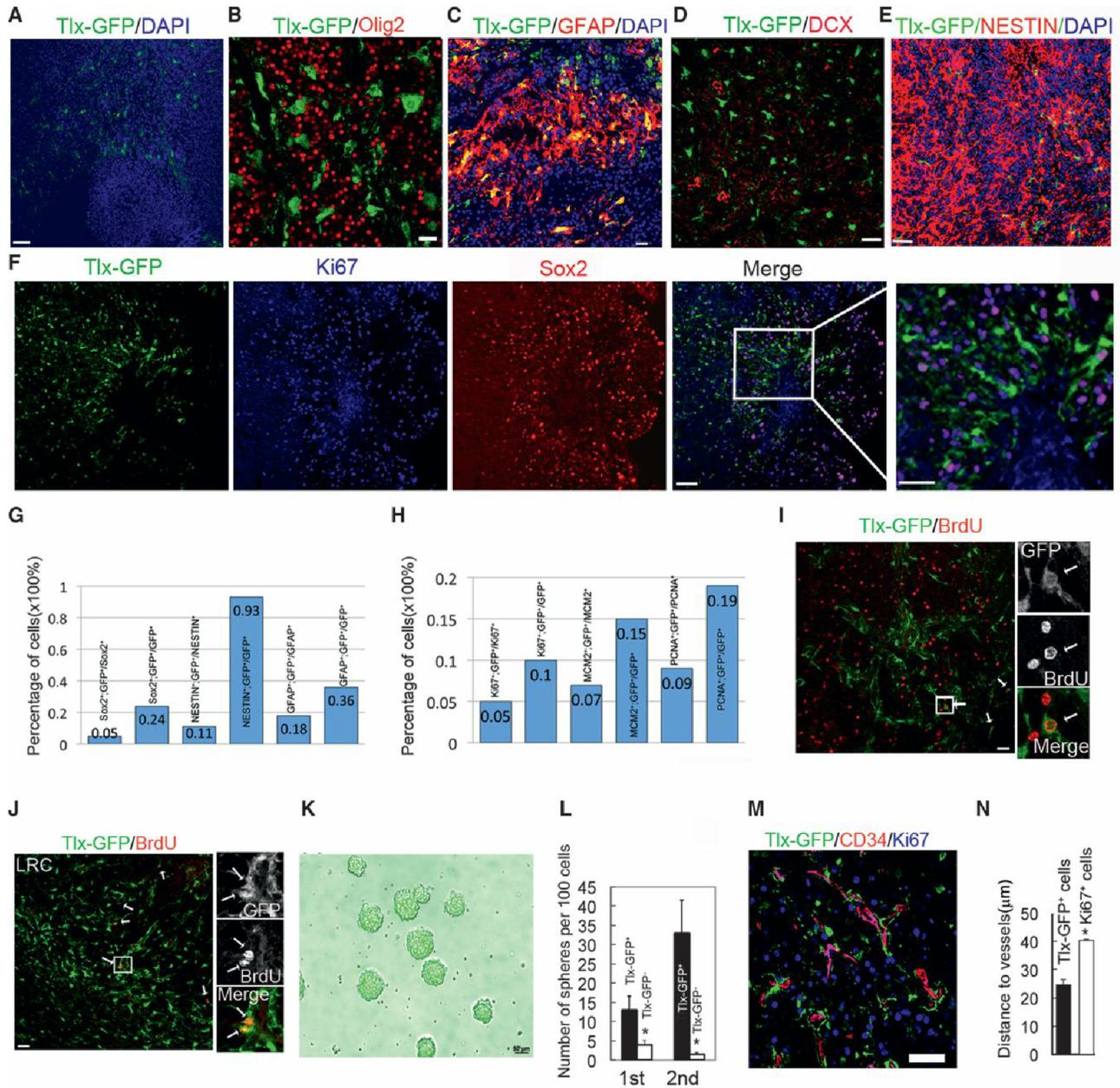
the Tlx-expressing cells in the SVZ are slowly dividing *in vivo* (Liu et al., 2008). This could be further confirmed by bromodeoxyuridine (BrdU) pulse labeling (2 hr) experiments, which showed that a small population (6.3%) of the BrdU-incorporating cells were positive for GFP and around 9.4% of the GFP<sup>+</sup> cells incorporated BrdU (Figure 2I). This is consistent with the results obtained by containing of Tlx-GFP with proliferation markers (Figure 2H). BrdU label retention experiments have been used to demonstrate the slow-dividing features of normal stem cells (Bickenbach, 1981). We performed BrdU injections for 3 consecutive days and mice were analyzed 2 weeks after the last BrdU injection, which allowed us to visualize the BrdU label-retaining cells (LRC) in tumor tissues. We observed that most of the BrdU label retaining cells (74.5%) are GFP<sup>+</sup> (Figure 2J, arrows), which strongly suggests that Tlx-GFP<sup>+</sup> cells are the slow-dividing cells in brain tumors. We also found that the Tlx-GFP<sup>+</sup> BrdU LRC cells are positive for GFAP or NESTIN, but are negative for Sox2 (Figures S2C–S2E). The Tlx-GFP<sup>-</sup> LRC cells are Olig2<sup>+</sup> and Sox2<sup>-</sup> (Figures S2F and S2G), suggesting these are the fully differentiated tumor cells.

The neurosphere assay has been widely used for the determination of the existence of self-renewing cells in brain tissues (Singec et al., 2006). Here, we found that when tumor cells are placed into tumorsphere culture medium containing FGF2 and EGF, the vast majority of tumorspheres are GFP<sup>+</sup> (Figure 2K). We isolated the Tlx-GFP<sup>+</sup> and GFP<sup>-</sup> cells via fluorescence-activated cell sorting (FACS) from primary tumors, these two populations were cultured in parallel using tumorsphere culture conditions. We found that only the Tlx-GFP<sup>+</sup> cells form tumorspheres and they can be passaged multiple times. The Tlx-GFP<sup>-</sup> cells cannot form tumorspheres efficiently and no spheres can be further passaged (Figure 2L).

It is intriguing that many GFP<sup>+</sup> cells are located in close proximity to the vasculature as shown by staining of an endothelial marker CD34 (Figure 2M), suggesting that the GFP<sup>+</sup> cells are located in a vascular niche. This is further supported by measuring the average distance between GFP cells and CD34<sup>+</sup> endothelial cells. The distance is significantly shorter than the distance between Ki67<sup>+</sup> and CD34<sup>+</sup> cells (Figures 2M and 2N). Although this is in line with a well-recognized feature of BTSCs, the slow-dividing feature of the perivascular cells does not support the notion that the vascular niche is supporting the proliferative capacity of stem cells. These results suggest that Tlx-GFP cells are the bona fide slow-dividing BTSCs in primary brain tumors.

### Xenotransplantation and Lineage Tracing Demonstrate that Tlx-GFP Cells Are BTSCs

BTSCs were initially defined as cells that can initiate new tumors when transplanted into immunocompromised mice (Singh et al., 2004). To test whether there is a difference between Tlx-GFP<sup>+</sup> and Tlx-GFP<sup>-</sup> cells regarding their potential to initiate tumors after transplantation, we performed the following serial transplantation experiment. Tlx-GFP<sup>+</sup> and Tlx-GFP<sup>-</sup> cells were isolated via FACS. A different number of cells (10<sup>5</sup>, 10<sup>4</sup>) were injected intracranially into the brains of nude mice. Animals were monitored with a 9.4 Tesla MRI device to determine tumor frequency (Figure 3A). We found that the Tlx-GFP cells form tumors very efficiently upon transplantation, and the tumors



**Figure 2. Characterization of Tlx-GFP<sup>+</sup> Cells in Primary Mouse Brain Tumors**

(A) GFP staining of mouse tumors developed in Ntv-a;Tlx-GFP mouse brain. Note that there is only a subset of tumor cells that are positive for GFP. Scale bar represents 20  $\mu$ m.

(B–E) Brain tumor sections derived from Ntv-a;Tlx-GFP mouse brain were stained for: GFP (green), Olig2 (red) (B); GFP (green), GFAP (red) (C); GFP (green) DCX (red)(D), and GFP (green), NESTIN (red)(E). Note that there is no colocalization between GFP, Olig2, and DCX. Only a subset of GFAP or NESTIN cells express Tlx-GFP. Scale bars represent 20  $\mu$ m.

(F) Staining of Ntv-a;Tlx-GFP brain tumor sections with GFP(green)/Ki67(blue)/Sox2(red) antibodies demonstrates that most of the Tlx-GFP cells are negative for Ki67 and Sox2 staining. This suggests they are quiescent in vivo. Note that most of the Ki67<sup>+</sup> cells express Sox2. Scale bars represent 50  $\mu$ m.

(G and H) Quantification of relative distribution of Tlx-GFP<sup>+</sup> cells comparing with different markers (G: Sox2, NESTIN, GFAP; H: Ki67, MCM2, and PCNA). Note that none of these markers can exclusively label Tlx-GFP cells.

(I) BrdU labeling was performed to see the proliferation in tumor mice. BrdU injection to Ntv-a;Tlx-GFP tumor-bearing mice was performed 2 hr before sacrificing the mice. BrdU (red) and GFP (green) costaining of tumor sections were performed. Note that only a very small populations of GFP cells are BrdU<sup>+</sup> (inset showing a higher magnification of a BrdU and GFP double-positive cell). Scale bar represents 20  $\mu$ m.

(J) Ntv-a;Tlx-GFP tumor mice received 3 days consecutive BrdU injections and the mice were sacrificed 3 weeks after the last injection. BrdU (red) and GFP (green) costaining of tumor sections were performed. Note that the BrdU label-retaining cells are GFP<sup>+</sup> (arrows). Inset: a higher magnification of a BrdU and GFP double-positive cell. Scale bar represents 20  $\mu$ m.

(legend continued on next page)

recapitulate the histological features of primary tumors, namely pseudopalisading-associated necrosis and vascular proliferation (Figures 3A and 3D). The Tlx-GFP<sup>-</sup> cells rarely form tumors upon transplantation of 10<sup>5</sup> cells (Figure 3A). We also analyzed the tumors initiated by Tlx-GFP<sup>+</sup> cells. Although we only injected Tlx-GFP-pure population, the heterogeneous expression pattern of Tlx-GFP is reestablished in transplanted tumors and they are still restricted in the slow-dividing subpopulation in primary tumors shown by costaining with PCNA or NESTIN (Figures 3B and 3C). This suggests the GFP<sup>+</sup> cells are able to generate GFP<sup>+</sup> and GFP<sup>-</sup> cells after transplantation. All animals developed tumors after injection of 10<sup>4</sup> Tlx-GFP<sup>+</sup> cells, but no tumors were observed with injection of the same amount of Tlx-GFP<sup>-</sup> cells (Figure 3D). Most importantly, the tumors initiated by Tlx-GFP<sup>+</sup> cells could be serially transplanted up to at least three times during our analysis time (Figures 3E and 3F). The one tumor initiated by the GFP<sup>-</sup> population (1/7) can be further transplanted, but the histology analysis indicates that it is a rather benign tumor without the histological features of GB (Figure 3G and 3H). Most importantly, we did not observe Tlx-GFP expression in transplanted tumors from Tlx-GFP<sup>-</sup> cells, indicating that reacquiring Tlx-GFP expression is not a frequent event in vivo. This result demonstrated that Tlx-GFP tumor cells can reestablish a heterogeneous tumor upon transplantation. We also performed a similar transplantation experiment using FACS-isolated primary Tlx-GFP<sup>+</sup> and Tlx-GFP<sup>-</sup> cells. The results are consistent with the experiment described above (Figure S3) suggesting that Tlx-GFP<sup>+</sup> cells are the cells that can serially transplant tumors.

The limitation of the xenotransplantation assay has been raised recently (Magee et al., 2012). This prompts us to perform the lineage tracing experiment to visualize the potential of Tlx<sup>+</sup> tumor cells. As shown in Figure 3I, Tlx-GFP;Tlx-CreERT2;Ntva;Confetti tumor-bearing mice were treated with TMX to induce Cre-dependent recombination only in the Tlx-GFP cells because Cre is controlled by the Tlx promoter. The confetti mice allow random labeling of cells with one of four random colors upon Cre activation (Snippert et al., 2010). In our model, the tracing was initiated in established tumors, thus no retracing experiment is needed, which is different from the lineage tracing experiment in a colon adenoma model (Schepers et al., 2012). We also showed above that Tlx-GFP cells in the tumor are the tumor cells because they initiate tumors upon transplantation. No recombination was observed without TMX treatment. After low-dosage TMX treatment, we can specifically induce spark-labeling only in the single Tlx-GFP<sup>+</sup> cells in side tumor (Figure 3J). In this experiment, the confocal setting can detect all colors during the image acquisition, also excluding clones with mixed colors. Seven days after TMX injection (as shown in Figure 3K, represented by YFP<sup>+</sup> clone), we found that most of the clones are two to three cells and they are grouped together. Thirteen days

after TMX injection, more cells were found but the cells started to migrate away from each other. It is important to note the glioblastoma is one of the most invasive solid tumors, and our data also demonstrated that invasive feature of these cells. Two months after TMX treatment, we started to see big clones that had infiltrated into different places inside tumors (Figure 3K). This experiment clearly demonstrates that a single Tlx<sup>+</sup> tumor cell can generate big clones of tumor cells over the long term. These results, together with the transplantation assay, strongly suggest that Tlx-GFP cells are BTSCs in vivo.

Recently, it was demonstrated that BTSCs are resistant to temozolomide (TMZ) treatment (Chen et al., 2012), most likely because of their slow-dividing feature as we described. To test whether Tlx-GFP cells become activated after TMZ treatment, we applied TMZ as described in Figure 3L and followed with CidU long-term tracing to label the slow-dividing Tlx-GFP cells. IdU was injected into tumor mice 2 hr before analysis. We found that without TMZ treatment, Tlx-GFP/CidU cells did not incorporate IdU during the analysis period, but the TMZ-treated Tlx-GFP/CidU cells became IdU<sup>+</sup> indicating that they reentered the cell-cycle after TMZ treatment (Figure 3K). This suggests that these cells are the cell-of-origin of relapsed tumors after chemotherapy.

### Inducible Inactivation of Tlx in Brain Tumors Leads to Prolonged Survival

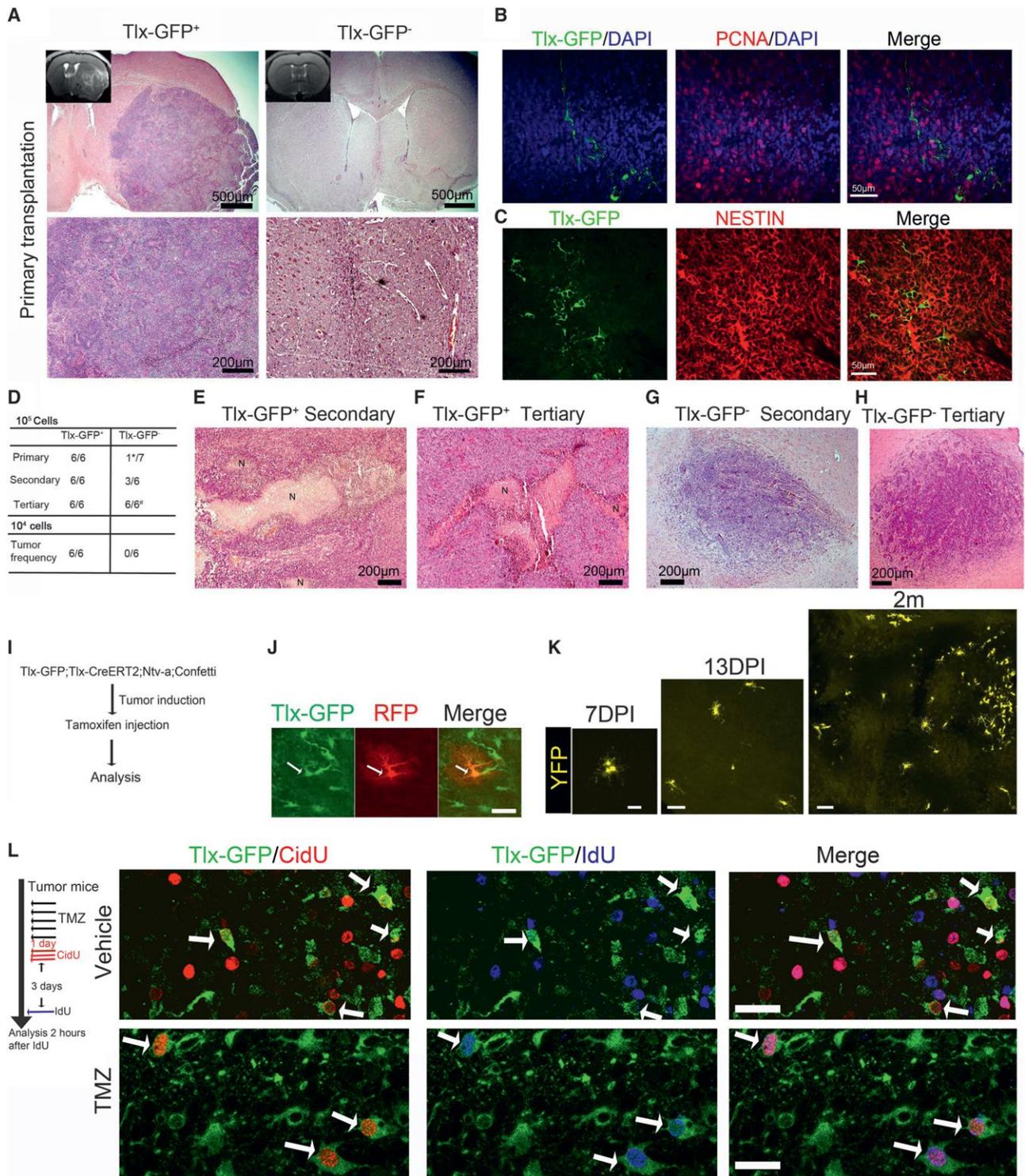
The fact that Tlx is an essential transcription factor for self-renewal of NSCs and for brain tumor initiation from NSCs, led us to investigate whether inactivation of Tlx can serve as a therapeutic strategy for treating GB. We generated the Ntv-a; Nestin-CreERT2;Tlx<sup>flox/flox</sup> mice for such an experiment. The Tlx-GFP population is also positive for NESTIN (Figure 2I), which suggests that we are able to induce the knockout of Tlx in brain tumors using Nestin-CreERT2 mice, in which a TMX-inducible Cre recombinase is expressed under the *Nestin* promoter. Based on the observation of the survival of PDGFB/AKT-injected animals, we decided to induce Tlx mutation with TMX 10 days before the tumor-bearing animals were expected to develop tumor-related symptoms (i.e., 2.5 weeks after RCAS-PDGFB/AKT injection). We confirmed that high-grade tumors were already developed in those animals at the stage they underwent TMX treatment (Figure 4A; N indicates necrosis, n = 5). Control animals were Ntv-a;Tlx<sup>flox/flox</sup> mice, which received identical treatment. We observed that the Tlx mutant tumor mice survived significantly longer than the control mice (Figure 4B), suggesting that targeting Tlx in brain tumors is beneficial for the survival of tumor-bearing animals. To confirm whether Cre recombination is efficient in mutant tumors, both RNA and protein were isolated for Tlx expression analysis, and we could confirm that Tlx RNA and protein levels were both significantly downregulated in the

(K) All tumorspheres derived from the Ntv-a;Tlx-GFP tumors are positive for Tlx-GFP, suggesting that only the Tlx-GFP cells can form tumorspheres in vitro. Note that around 94.7% of cells are GFP<sup>+</sup> in this culture.

(L) Ntv-a;Tlx-GFP brain tumors were separated into GFP<sup>+</sup> and GFP<sup>-</sup> populations via FACS before being subjected to tumorsphere culture experiment. Number of spheres formed by 100 cells were quantified. Note that spheres derived from GFP<sup>-</sup> can not be further passaged. p < 0.05.

(M) Ntv-a;Tlx-GFP brain tumors costained with CD34 (red), Ki67(blue), and GFP (green). Note that most GFP cells are negative for Ki67 and are associated with CD34<sup>+</sup> endothelial cells. Scale bar represents 50 μm.

(N) The average distance between Tlx-GFP cells and CD34<sup>+</sup> cells is shorter than that between Ki67 cells and CD34<sup>+</sup> cells. p < 0.05. See also Figure S2.



**Figure 3. Xenotransplantation and Lineage Tracing Demonstrate that Tlx-GFP Cells Are BTSCs**

(A) Tlx-GFP<sup>+</sup> or Tlx-GFP<sup>-</sup> (10<sup>5</sup>) cells were injected into the brain of nude mice. H&E staining indicated that Tlx-GFP<sup>+</sup> initiated tumors with GB features, but most of the mice injected with Tlx-GFP<sup>-</sup> cells did not develop tumors. Inset: MRI image of transplant-initiated tumors.

(B and C) Tumors initiated by Tlx-GFP<sup>+</sup> are heterogeneous. Costaining of Tlx-GFP with PCNA demonstrates that Tlx-GFP<sup>+</sup> cells remain to be slow-dividing (B) and it is restricted to a subpopulation of NESTIN expressing cells as in primary tumors (C).

(D) Summary of the transplantation experiment. When injected with 10<sup>5</sup> cells, all Tlx-GFP<sup>+</sup>-injected animals develop tumors, and these tumors can be serially transplanted at least for three rounds. The one tumor initiated by the Tlx-GFP<sup>-</sup> cells can be further transplanted. \*Only the animal developed tumor was used for

(legend continued on next page)

mutant tumors at the end point of the survival assay (Figures 4C and 4D). We found that the *Tlx* mutant tumors were smaller (Figure 4E). It is worth noting that most *Tlx* mutant animals still showed neurological symptoms and brain tumors could still be observed, although the *Tlx* mutant brain tumors were much smaller and less infiltrative (Figure 4F). Furthermore, we performed histological analyses of all tumors, and strikingly the majority of the *Tlx* mutant tumors did not have features of GB (e.g., pseudopalisading and vascular proliferation) (Figure 4G). This could be further confirmed through neuropathological assessment and grading according to the World Health Organization (WHO) classification of CNS neoplasias of all tumors by a trained neuropathologist. *Tlx* mutant tumors had a lower grade compared to the normal control tumors (Figure 4H). This result also provides a possible link between BTSCs and GB histopathology.

There are still tumor cells in the *Tlx* mutant mice. To assess the potential of *Tlx* mutant tumor cells, we performed the transplantation experiment after *Tlx* ablation. Primary tumors were disassociated and  $10^5$  primary tumor cells were injected in all the experiments. MRI was used to monitor tumors. The *Tlx* wild-type (WT) tumor cells transplant tumors very efficiently as shown in Figures 4I–4L. Histology analysis demonstrates that *Tlx* mutant tumors are very small and benign (Figure 4M), and we did not observe any histological features of GB. Importantly, the WT tumors can be serially transplanted but the *Tlx* mutant tumor failed to initiate any tumors for the secondary transplantation (Figures 4J, 4K, 4N, and 4O). These results strongly suggest that *Tlx* is required for tumor propagation in the transplantation assay.

These results suggest that death of animals in the *Tlx* mutant group was driven by cells derived from BTSCs before TMZ treatment, which was sufficient to cause damage of normal brain tissues and animal death as a consequence. To test this hypothesis experimentally, we decided to inactivate *Tlx* by injecting RCAS-Cre together with PDGFB and AKT to achieve an early deletion of *Tlx* gene during tumor development. We found that many animals did not get tumors in the Cre group compared to the group injected with RCAS-GFP together with PDGFB and AKT (Figure S4). This experiment suggests that early treatment of tumor stem cells will greatly improve treatment outcome.

### **Tlx Inactivation Leads to Loss of Self-Renewing BTSCs**

To analyze the cellular consequences in the primary tumor after *Tlx* deletion, we analyzed the cell proliferation in the tumors after the animals were sacrificed because of brain tumor-related symptoms and found that there were significantly less Ki67-ex-

pressing cells in the *Tlx* mutant tumors. This suggests there was a reduced proliferation upon *Tlx* deletion (Figure 5A). As *Tlx* is only expressed by the slow-dividing cells in brain tumors, we hypothesized that the decrease in Ki67<sup>+</sup> cells in *Tlx* mutant tumors was the consequence of *Tlx* inactivation in BTSCs. Interestingly, when we analyzed tumors 5 days after TMZ injection, we found that the number of Ki67<sup>+</sup> cells was not changed in the *Tlx* mutant tumors. This suggests that loss of *Tlx* results in a decreased generation of Ki67<sup>+</sup> proliferating cells (Figure 5B). To further investigate whether loss of *Tlx* leads to an impaired function of BTSCs, we performed a tumorsphere assay, which showed that *Tlx* mutant tumors generate significantly less tumorspheres (Figures 5C and 5D), and the mutant tumorspheres could not be further passaged and expanded (Figure 5D). To determine whether this is due to the loss of *Tlx* in BTSCs, we performed the tumorsphere culture assay using cells derived from the *Ntv-a;Nestin-CreERT2;Tlx<sup>flox/flox</sup>* tumors. Seventy-two hours after adding 4-hydroxytamoxifen (4-OHT) to the tumorsphere culture medium to induce *Tlx* deletion *in vitro*, we confirmed that *Tlx* was successfully removed from the cells (Figure 5E). We found that the *Tlx* mutant tumorspheres could not be passaged as efficiently as *Tlx*<sup>+</sup> tumor spheres (Figures 5E and 5F), suggesting that loss of *Tlx* in BTSCs leads to loss of self-renewal. To investigate whether loss of *Tlx* leads to an altered differentiation capacity of BTSCs, we established adherent cell cultures from untreated *Ntv-a;Nestin-CreERT2;Tlx<sup>flox/flox</sup>* tumors. After treatment with 4-OHT, the growth factors were removed from the medium to induce the differentiation process. We found a reduction of glial differentiation in *Tlx* mutant cells shown by GFAP or Olig2 staining (Figures 5G and 5H), whereas neuronal lineage differentiation was increased in *Tlx* mutant BTSCs shown by DCX staining (Figure 5I). These results demonstrated that loss of *Tlx* in primary brain tumors leads to loss of self-renewal and induction of neuronal differentiation, which may explain the survival effect of targeting *Tlx* in mouse gliomas.

### **Loss of Tlx in Brain Tumors Leads to the Induction of Pathways Regulating Apoptosis, Senescence, and Differentiation**

To investigate the molecular pathways that are altered after *Tlx* mutation in brain tumors, RNA was isolated from control and mutant tumors and subjected to microarray analysis providing insights into genome-wide expression changes after *Tlx* inactivation. Interestingly, many pathways involved in cell-cycle regulation, apoptosis, neural differentiation, and senescence were significantly altered (Table S2). Quantitative RT-PCR (qRT-PCR) analysis for some of the candidate genes confirmed that

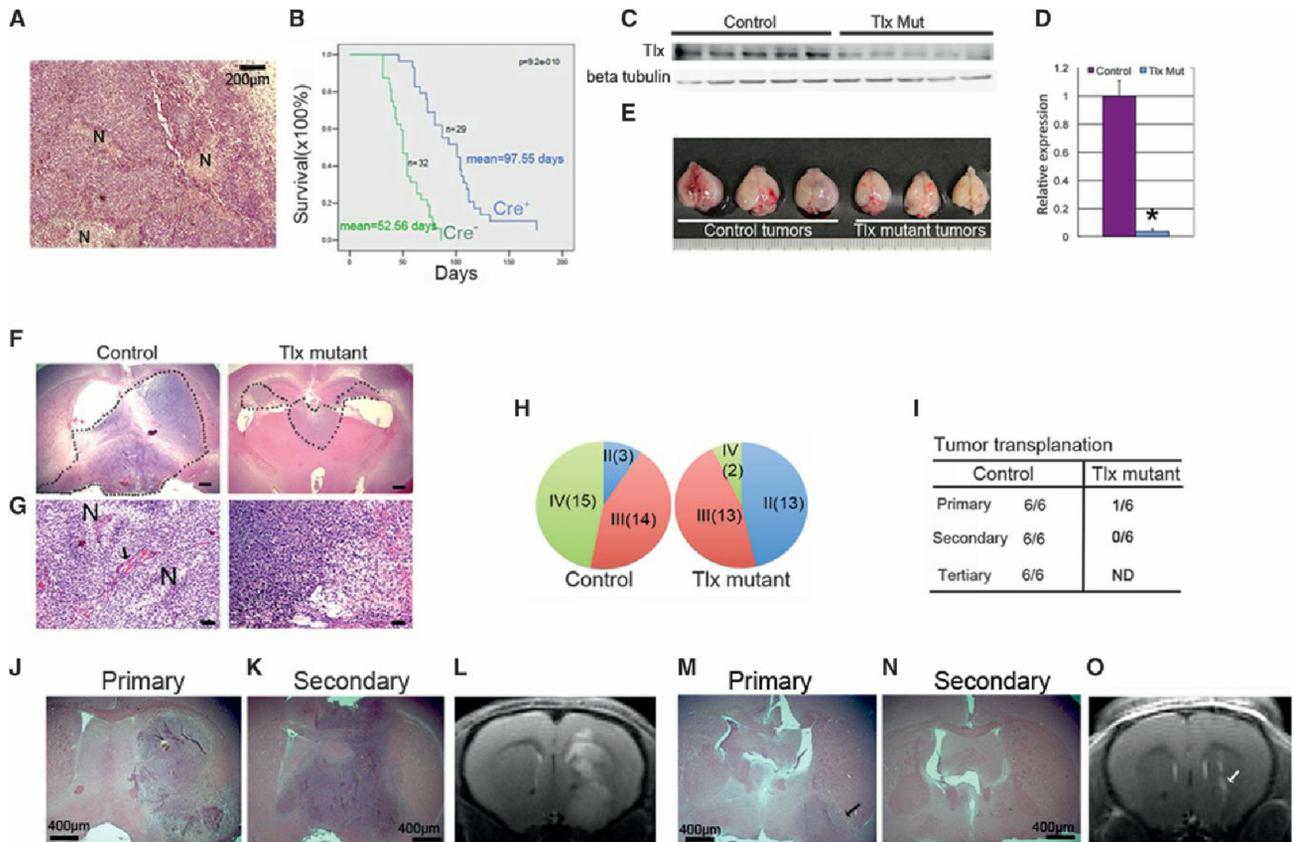
secondary transplantation, six animals received cells from *Tlx*-GFP<sup>-</sup> tumor and three developed tumors. #All tumors are low grade tumors. When injected with  $10^4$  cells, only the *Tlx*-GFP<sup>+</sup> cells initiated tumors.

(E–H) H&E staining of tumors from the serial transplantation experiment. Note that the *Tlx*-GFP<sup>+</sup> secondary and tertiary tumors have histological features of a GB (E and F), like necrosis (N), which was never observed from the *Tlx*-GFP<sup>-</sup>-derived tumors (G and H).

(I–K) Lineage tracing of *Tlx*<sup>+</sup> cells in primary tumors. (I) Experimental scheme. (J) TMZ induction of single cell recombination in *Tlx*-GFP tumor cells. GFP indicating *Tlx*<sup>+</sup> cells, RFP indicating cells being labeled by Cre recombination. (K) Representative picture of clone growth over time initiated by a *Tlx*<sup>+</sup> stem cell labeled with YFP after TMZ treatment. Scale bars represent 20  $\mu$ m if not indicated.

(L) *Tlx*-GFP tumor-bearing mice received TMZ treatment as described. CidU were injected into these mice 1 day after TMZ treatment, and three injections were performed in 1 day with 2 hr intervals. Three days later, animals were injected with IdU 2 hr before sacrificing. Brain sections were stained using antibodies against CidU (red), IdU (blue) and GFP (green). Note that in the vehicle-treated brains, GFP/CidU<sup>+</sup> cells do not incorporate BrdU indicating they are slow-dividing (arrows), but in the TMZ-treated brains, GFP/CidU<sup>+</sup> cells incorporate IdU indicate that they reenter cell-cycle (arrows).

See also Figure S3.



**Figure 4. Genetic Inactivation of Tlx in BTSCs Leads to the Loss of Self-Renewing BTSCs and Prolonged Survival**

(A) H&E staining demonstrates that malignant brain tumors were already developed at the time point of receiving TMX injection. N, necrosis.

(B) The Kaplan-Meier survival curve of mice receiving TMX. Note that inactivation of Tlx via TMX leads to a significant prolongation of animal survival. ( $p = 9.2E-10E$ ).

(C and D) RNA and protein were extracted from tumors of animals that were sacrificed at the onset of neurological symptoms. qRT-PCR analysis (C) demonstrating that Tlx is efficiently deleted upon TMX treatment. ( $n = 6, p < 0.0001$ ). This can be confirmed by the western blot results indicating that Tlx protein is absent in the mutant tumors. Data are represented as mean  $\pm$  SD if not specified.

(E) Brains obtained from Tlx wild-type tumors and Tlx mutant tumors.

(F and G) H&E staining of brain tumors after Tlx inactivation, tumor region was marked with a dash line. Note that the Tlx mutant tumors are much smaller and lack glioblastomas histological features like necrosis (N) associated with pseudopalisading and vascular proliferation (arrow). Scale bars represent 200  $\mu$ m (F) and 50  $\mu$ m (G).

(H) WHO classification of Tlx mutant and control brain tumors. Note that there are much less grade IV GBs in the Tlx mutant group.

(I) Tumor cells were isolated before TMX treatment and subjected to stem cell culture condition, 4-OTH was used for induction of Tlx mutation, and  $10^5$  cells of control and Tlx mutant were collected and injected to the nude mice brains. MRI were used to determine tumor frequency. Note that WT tumors can be serially transplanted, and Tlx mutant tumor cells can not initiate any tumors for the secondary transplantation experiment. \*All Tlx mutant cell initiated tumors are local and do not shown any invasive pattern under MRI.

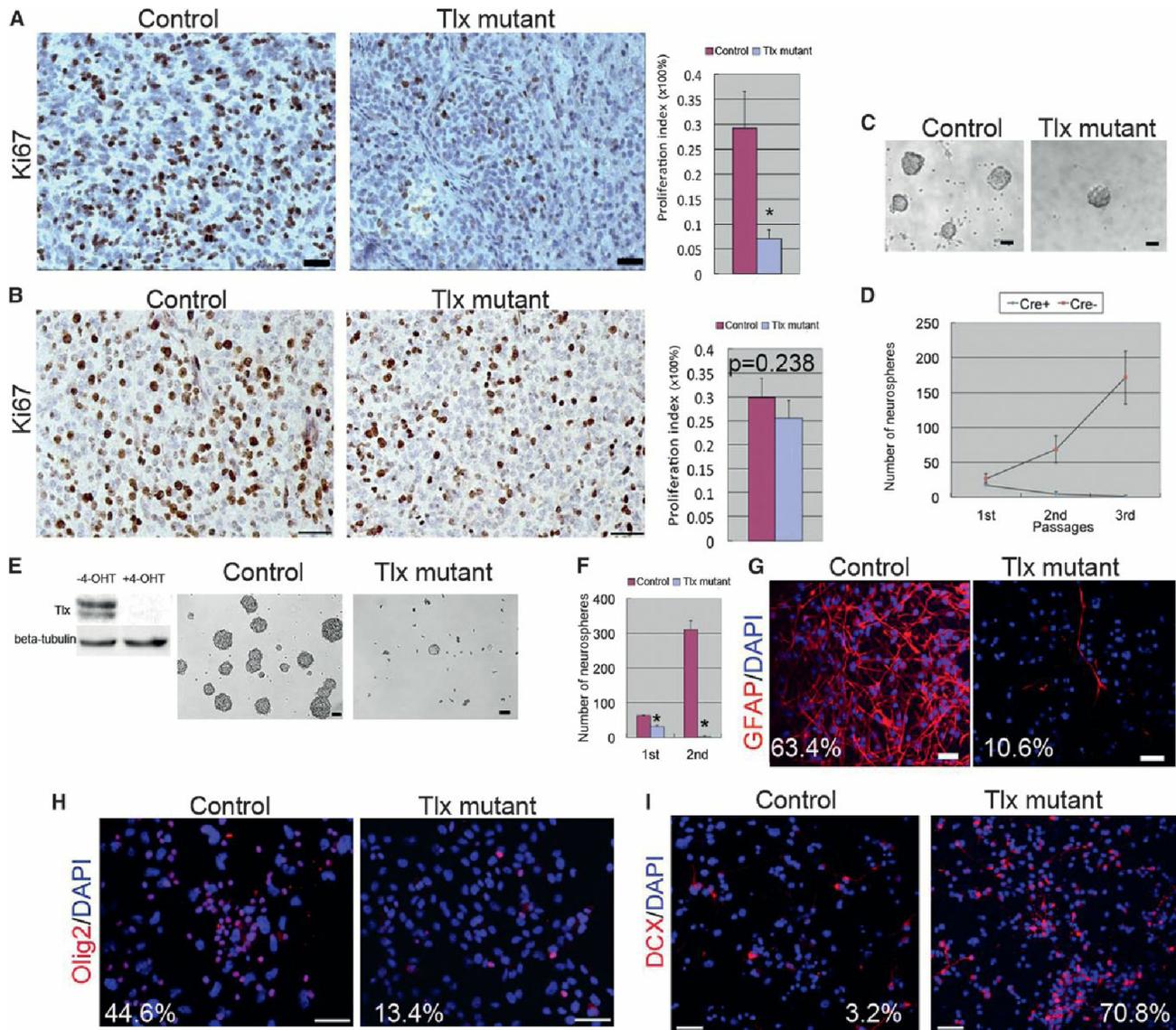
(J-O) Histology of transplanted tumors from the experiment described in (I). Note that high grade tumors can be initiated with WT cells (J and K), which can be clearly visualized with 9.4 tesla MRI, whereas the tumor initiated from the Tlx mutant population is very small (M, arrow) and can not be further transplanted (N). No tumor can be seen with the MRI for the secondary transplantation (O, arrow indicates the injection site).

See also Figure S4.

several important tumor suppressor genes were significantly upregulated. For instance, CDKN2A, CDKN2B, and PML are all upregulated in the Tlx mutant tumors (Figures 6A–6C), suggesting a general inhibition of tumor growth and induction of cell-cycle arrest and cell death after Tlx inactivation. Also, we confirmed that factors that are important for neuronal differentiation were induced in the Tlx mutant tumors as well, for example SMACC1, Dlx2, and TGF $\beta$ R1 (Figures 6D–6F). These results suggest that loss of Tlx leads to the induction of cell death and differentiation pathways, which is consistent with the phenotype

we observed in the mice brain tumors. Dlx2 and TGF $\beta$ R1 have been suggested as essential regulators for neuronal differentiation (Brill et al., 2008; Shah et al., 1996), which explains in part the increase of neuronal differentiation of Tlx mutant BTSCs (Figures 5G–5I).

The microarray data reflects general changes of genes in a mixed population. To further determine which molecular pathways were altered in BTSCs after Tlx deletion, we isolated the BTSCs from Nestin-CreERT2;Tlx<sup>flox/flox</sup>;Ntva tumor mice untreated with TMX. The cells are cultured under monolayer

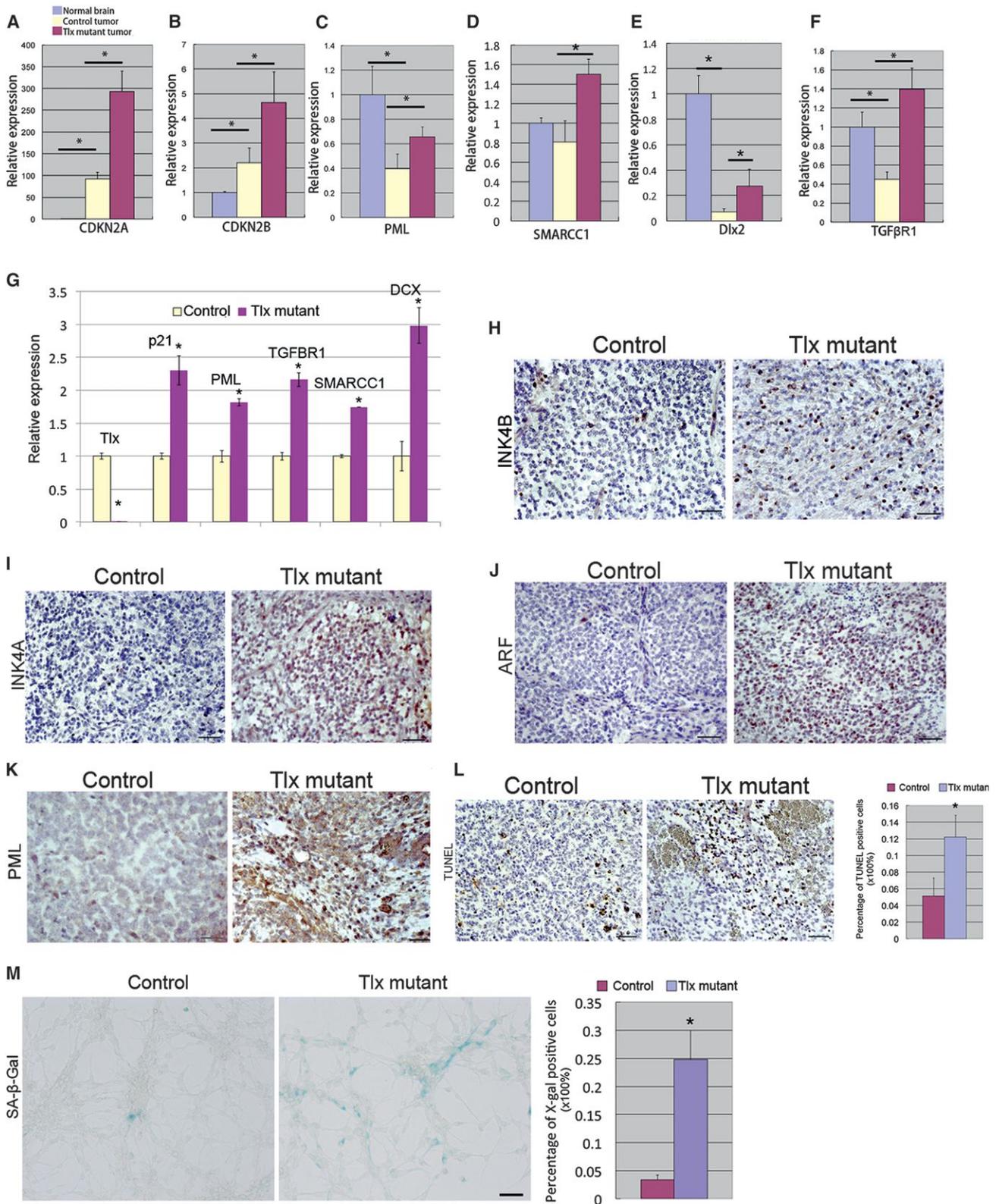


**Figure 5. Loss of Tlx in BTSCs Leads to the Loss of Self-Renewal and Induction of Neuronal Differentiation**

(A) Ki67 staining of Tlx mutant and control brain tumors. Proliferation index means percentage of Ki67<sup>+</sup> cells in tumors ( $n = 4$ ,  $p < 0.01$ ). Scale bar represents 50  $\mu\text{m}$ . (B) Five days after TMX induction of Tlx mutation, brain tumors were analyzed for Ki67 staining. Note that no significant differences were observed between normal tumors and Tlx mutant tumors. ( $n = 4$ ,  $p = 0.238$ ). Scale bar represents 50  $\mu\text{m}$ . (C and D) Fewer tumorspheres were derived from Tlx mutant tumors (C). Tumorspheres obtained from Tlx mutant tumors cannot be passaged in vitro whereas control tumorspheres can be passaged and expanded (D). Scale bars represent 50  $\mu\text{m}$ . (E and F) Tumorsphere culture from the *Ntv-a;Nestin-CreERT2;Tlx<sup>fllox/fllox</sup>* tumors untreated with TMX. 4-OHT was added to the tumorspheres culture medium to induce Tlx mutation in vitro. Western blot shows that Tlx can be efficiently removed after 72 hr of 4-OHT treatment (E). Much less tumorspheres were obtained from Tlx mutant cells after the first passage (E), the Tlx mutant tumorspheres cannot be further passaged and expanded (F,  $p < 0.001$  both in the first and second passages). Scale bars represent 50  $\mu\text{m}$ . (G–I) 4-OHT was added to adherent culture of the *Ntv-a;Nestin-CreERT2;Tlx<sup>fllox/fllox</sup>* tumors. Growth factors were then removed to induce differentiation. Antibodies against GFAP (G), Olig2 (H), and DCX (I) were used for staining of differentiated cells. Note that there are less GFAP and Olig2<sup>+</sup> cells in the Tlx mutant culture, but an increase of DCX<sup>+</sup> cells was observed in the Tlx mutant culture. Scale bars represent 50  $\mu\text{m}$ .

condition and 4-OHT treatment was performed in vitro that allowed inactivation of Tlx in BTSCs (Figure 6G). We also included two known Tlx target genes (PTEN and p21) that were reported to be repressed by Tlx (Sun et al., 2007; Zhang et al., 2006). RNA expression analysis shown in Figure 6G suggests that as known Tlx target genes, only p21 (CDKN1A), but not

PTEN, is upregulated upon Tlx inactivation in BTSCs. PML, TGF $\beta$ R1, SMARCC1, and DCX are upregulated suggesting that these genes are initially upregulated in Tlx mutant BTSCs, although it is not known yet whether the expression of these genes are repressed by Tlx protein directly, whereas CDKN2A and CDKN2B are not altered in Tlx mutant BTSCs (data not



**Figure 6. Loss of Tlx in Brain Tumors Leads to Changes in Multiple Essential Pathways Regulating Apoptosis, Senescence, and Differentiation**

(A–F) qRT-PCR assay using RNA isolated from normal brain, control, and Tlx mutant tumors. Results show the induction of the expression of CDKN2A, CDKN2B, PML, SMARCC1, Dlx2, and TGFβR1 in Tlx mutant tumors (n = 6 for each group, p < 0.05).

(legend continued on next page)

shown), suggesting that the upregulation of these two genes are secondary consequences upon *Tlx* deletion in primary tumors. It is important to note that we have shown that *Tlx* overexpression in NSCs does not lead to changes of *CDKN2A/CDKN2B* expression (Liu et al., 2010), which is consistent with the current results.

It is intriguing that we found factors like p21, *CDKN2A*, *CDKN2B*, and *PML* are all upregulated in the *Tlx* mutant, although not necessarily only in BTSCs. Additionally, we confirmed the upregulation of *CDKN2A* (*ARF/INK4A*), *CDKN2B* (*INK4B*), and *PML* in *Tlx* mutant tumors by IHC (Figures 6H–6K), suggesting upregulation of these genes in many tumor cells. These factors are known to be involved in the regulation of cell-cycle arrest, apoptosis, and senescence (Gil and Peters, 2006; Sharpless and DePinho, 2007). To investigate whether any of these phenotypes is induced in the *Tlx* mutant tumors, we first analyzed apoptosis in the *Tlx* mutant brain tumors. TUNEL assay was performed, and we found that there was a significant increase of TUNEL<sup>+</sup> cells in the *Tlx* mutant tumors (Figure 6L), suggesting that loss of *Tlx* also leads to an increase in apoptosis of tumor cells. Furthermore, in order to detect whether there is a senescence phenotype after *Tlx* inactivation, we performed senescence-associated beta-gal (*SA-β-Gal*) staining on tumor sections. However, we could not detect any cellular senescence in situ (data not shown). One explanation could be that the cells that underwent senescence had been already removed by macrophages from the primary tumors (Kay, 1975). To further investigate whether there was an induction of senescence, we isolated BTSCs from *Nestin-CreERT2;Tlx<sup>fllox/fllox</sup>;Ntv-a* tumor animals that have not been treated with TMX. We treated the cells with 4-OHT for 72 hr as described above, 4 days after removal of 4-OHT from the medium, the *SA-β-Gal* assay was performed. We observed an increase of *β-Gal*<sup>+</sup> cells in the *Tlx* mutant culture (Figure 6M), indicating an induction of senescence after inactivation of *Tlx* in BTSCs. The senescence phenotype is likely caused by activation of p21 in BTSCs. These results suggest that the p21, *INK4A/ARF*, *INK4B*, and *PML* pathways are induced upon removal of *Tlx* in BTSCs, which leads to prolonged survival of tumor-bearing mice.

### ***Tlx* Is Important for Self-Renewal of Human BTSCs**

We have shown previously that *Tlx* is upregulated in human brain tumors (Liu et al., 2010). Most importantly, by analyzing the Cancer Genome Atlas (TCGA) data set (<http://hgserver1.amc.nl/cgi-bin/r2/main.cgi>), we found that *Tlx* high expression correlates with poor survival in human GB patients, indicating that *Tlx* is a prognosis marker and a therapeutic target for human GB (Figure 7A).

To investigate whether loss of *Tlx* in human BTSCs will lead to similar phenotype as we observed in mice, we obtained three

different patient-derived human glioblastoma stem cells. Two small hairpin RNAs (shRNAs) against the human *Tlx* gene were used for knock down (KD) of *Tlx* in human BTSCs, as shown in Figure 7B. *Tlx* expression is silenced by the two independent shRNA constructs, using shRNA against luciferase as control. We then analyzed the sphere formation ability of *Tlx* KD human BTSC. Similar to the results in mouse BTSCs, we found that *Tlx* inhibition leads to loss of sphere-forming ability and a decrease of proliferation of human BTSCs (Figures 7C, 7D, and S5), and no *Tlx*-KD cells can be passaged three times for further analysis. We then analyzed genes that are changed in *Tlx* mutant mouse cells. p21 and *PML* were found to be upregulated in *Tlx* KD cells, which explains the sphere-forming results. We also found that *Tlx* KD leads to upregulation of *TGFβR1* and *Dlx2*, suggesting an increase of neuronal differentiation, as we have seen in the mouse mutant cells (Figure 7E). We did not observe changes of *CDKN2A*, *CDKN2B*, *SMARCC1*, and *PTEN* (data not shown). These results suggest that *Tlx* is important for the regulation of human BTSCs.

Here, we demonstrate that *Tlx* is expressed in slow-dividing BTSCs of primary brain tumors by using a mouse somatic brain tumor model. Lineage tracing of single *Tlx*<sup>+</sup> cells in vivo demonstrate they are BTSCs. We also achieved BTSC-specific genetic targeting of *Tlx* in established brain tumors and showed that *Tlx* is essential for brain tumor maintenance and survival of tumor-bearing mice (Figures 7H and 7I). These results strongly support the CSCs hypothesis and provide direct evidence of improved animal survival after targeting BTSC.

## **DISCUSSION**

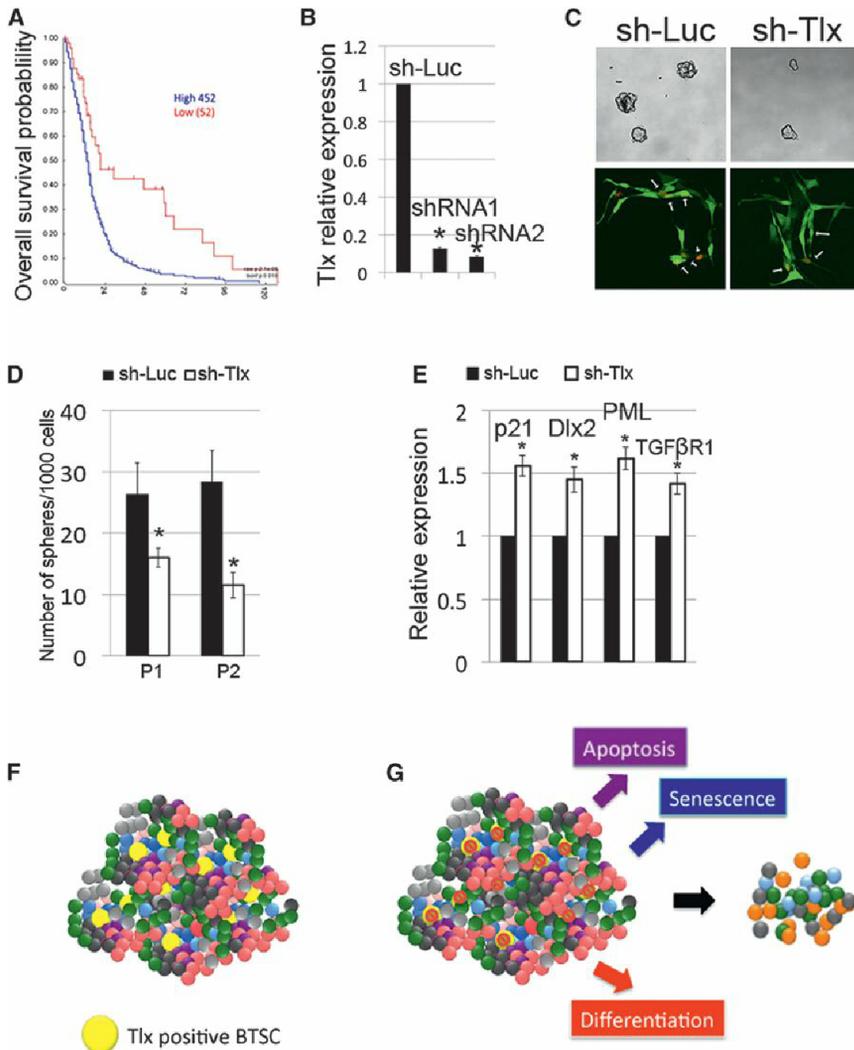
### **Cancer Stem Cells Are Suitable Therapeutic Targets**

The CSCs hypothesis is under intensive debate. One major reason is that markers used for CSCs isolation are not reproducible, which may reflect the genetic and phenotypic heterogeneity of malignant cells or transient expression of these markers. The scientific community faced similar problems when they started to identify adult stem cells in different organs and tissues. The controversies were later alleviated mostly by using animal models, which allow genetic manipulation of particular cell types and follow their behavior for a life-long time (Kretschmar and Watt, 2012). Recently, several studies were published, which demonstrated the existence of CSC-like cells in different tumor entities using different genetic tools (Chen et al., 2012; Driessens et al., 2012; Schepers et al., 2012). These studies strongly suggest the existence of a differentiation hierarchy in primary tumors. Here, we demonstrated that in a mouse glioma model, a subset of glioma cells expressing the functional NSC marker *Tlx*, and these cells are slow-dividing in vivo. The *Tlx*-GFP cells are the cells that can form long-term self-renewing spheres

(G) 4-OHT was added to adherent cultures of the *Ntv-a;Nestin-CreERT2;Tlx<sup>fllox/fllox</sup>* tumors to induce Cre recombination. RNA was collected for qRT-PCR assay, *Tlx* can be efficiently removed by 4-OHT treatment. Results show that p21, *PML*, *TGFβR1*, *SMARCC1*, and *DCX* are induced upon *Tlx* inactivation in BTSCs. (H–L) Mouse brain tumors were stained for *INK4B* (H), *INK4A* (I), *ARF* (J), and *PML* (K). All the four proteins were upregulated in the *Tlx* mutant tumors. Scale bars represent 50 μm in (H–J) and 20 μm in (D).

(L) TUNEL assay demonstrated that there are more apoptotic cells in *Tlx* mutant brain tumors ( $n = 4$ ,  $p < 0.01$ ). Scale bar represents 50 μm.

(M) 4-OHT was added to adherent cultures of the *Ntv-a;Nestin-CreERT2;Tlx<sup>fllox/fllox</sup>* tumors to induce Cre recombination, a senescence-associated *β-gal* assay (*SA-β-Gal*) was performed. Note that there are more *SA-β-Gal*<sup>+</sup> cells in *Tlx* mutant culture ( $n = 4$ ,  $p < 0.01$ ). Scale bar represents 50 μm. See also Table S2.



**Figure 7. Tlx Is a Therapeutic Target in Human BTSCs**

(A) The Kaplan-Meier survival curve of GB patients from the TCGA database. Note that patients with higher Tlx expression have a worse survival rate than patients with lower Tlx expression ( $p < 0.05$ ). (B) Knock-down (KD) of Tlx expression in human BTSCs was achieved by using two independent shRNAs against human Tlx ( $p < 0.05$ ). (C and D) KD of Tlx in human BTSCs (T269) leads to loss of sphere forming ability and decrease of cell proliferation measured by BrdU (red) assay. (D) Arrow indicates GFP/BrdU double-positive cells. BrdU was added to the medium for 3 hr before analysis, GFP indicates the shRNA-infected cells ( $p < 0.05$  for passage 1 and passage 2). (E) KD of Tlx leads to upregulation of p21, Dlx2, PML, and TGFβR1 expression in human BTSCs ( $p < 0.05$ ). (F and G) Graphic illustration of Tlx expression is restricted to the slow-dividing BTSCs (F). Removal of Tlx in the BTSCs leads to regression of brain tumors (G) through induction of differentiation, senescence, and apoptosis pathways. See also Figure S5.

in vitro. We also performed lineage tracing experiments to show that Tlx<sup>+</sup> cells can generate large tumor clones over 2 months. Most importantly, we showed that inactivation of Tlx in brain tumors leads to a prolongation of survival of tumor-bearing animals. This is, in part, due to the induction of several key pathways, which are important for the regulation of senescence, apoptosis, cell-cycle, and differentiation (Figures 7F and 7G). This suggests that the tumor maintenance is dependent on the BTSC population, and loss of this tumor-seeding population will lead to regression of the tumor mass. This directly demonstrates that BTSCs are important for tumor progression and survival.

Nevertheless, it is noteworthy that the Tlx mutant tumors do not disappear. There are several possible explanations of this result. First, the gene targeting of Tlx in BTSCs was initiated in Tlx expressing slow-dividing cells, and the proliferating cells are not affected at the beginning of the targeting (Figure 5B). These cells can still drive the tumor growth until they are exhausted. This growth is causing neurological symptoms in the Tlx mutant mice and this was further supported by the results obtained by inactivating Tlx at early time point (Figure S4).

not support this hypothesis. Thus we would favor the first explanation for this phenotype.

**Role of Tlx in Normal and Cancer Stem Cells**

It was known that Tlx is important for adult NSC maintenance in the SVZ (Liu et al., 2008) and that overexpression of Tlx leads to glioma initiation in mouse models when combined with p53 mutations (Liu et al., 2010). High expression of Tlx was found in different types of tumors in the CNS, including astrocytoma, ependymoma, and GB (Liu et al., 2010). Here, we directly demonstrated that Tlx is also important for BTSCs maintenance. These results suggest that Tlx is an essential regulator and marker for stem cell features in BTSCs. It is interesting to see that inactivation of Tlx in tumors leads to the induction of pathways involved in regulation of cell death, cell-cycle, and neuronal differentiation. We observed that the Tlx target gene *p21*, which is an essential factor mediating cell-cycle arrest and senescence (Rowland and Peeper, 2006), is upregulated both in mouse and human BTSCs upon removal of Tlx, suggesting this is a very conserved pathway. To our knowledge, it has never been shown that Tlx inhibits the INK4A/ARF pathway in normal NSCs, and we

did not see changes of INK4A/ARF expression in Tlx overexpressing NSCs (Liu et al., 2010). This suggests that the upregulation of INK4A/ARF upon inactivation of Tlx is tumor-specific. It was shown that Tlx represses expression of PTEN, which are thought to be essential for maintaining NSC proliferation (Zhang et al., 2006). Here, we did not observe PTEN upregulation in the mutant tumors from the microarray analysis. One possible explanation is that PTEN is silenced or mutated in the AKT+PDGFB-induced high-grade mouse GBs by other mechanisms, which occurs often in human GB patients as well (Furnari et al., 2007).

Taken together, we demonstrated here that Tlx is an essential indicator and regulator of BTSCs. The mouse models we generated provide additional approaches for therapeutic target validation. We also demonstrated that Tlx is a prognosis marker for poor survival of human GB patients, which together with our mouse data suggest that Tlx is a promising therapeutic target for brain tumors. Another advantage of targeting Tlx in brain tumor is that Tlx expression is restricted to the CNS system (Monaghan et al., 1995) and hence an inhibitor that is specifically blocking Tlx activity will have few side effects. As many other nuclear receptors, Tlx is very likely a druggable protein.

## EXPERIMENTAL PROCEDURES

### Animal Experiments

Mice were housed according to international standard conditions and all animal experiments complied with local and international guidelines for the use of experimental animals. The Ntv-a mouse was kindly provided by Eric Holland. Tlx<sup>fllox/fllox</sup> animals was generated as described (Belz et al., 2007). The Nestin-CreERT2 animal was generated as described elsewhere and has been successfully used to target both SVZ and SGZ NSCs via tamoxifen injection (Corsini et al., 2009). Tlx-GFP reporter animal was obtained from the GENSAT project (Gong et al., 2003). Confetti mice were from Jackson Laboratory. Tamoxifen (Sigma) was dissolved in sunflower seed oil (Sigma) with 10% EtOH<sub>abs</sub> to prepare a 10 mg ml<sup>-1</sup> solution. Intraperitoneal injections were performed with 1 mg/day for 10 days. BrdU, CidU, or IdU (Sigma) were dissolved in sterile 0.9% saline to prepare a 15 mg ml<sup>-1</sup> solution, and mice were injected intraperitoneally with 300 mg kg<sup>-1</sup> 2 hr before sacrifice or with 100 mg kg<sup>-1</sup>/day for 3 days as described in this article. TMZ (Sigma) was dissolved in DMSO and freshly diluted in 0.9% saline (5 mg/ml) and injected intraperitoneally with 100 mg kg<sup>-1</sup>/day for 5 days.

### RNA Isolation, RT-PCR, and Microarray

RNA from cultured NSCs and dissected mouse tumors and normal brain tissues (mixture of tissues from olfactory bulb, cortex, striatum, hippocampus, and hypothalamus) were isolated with RNeasy Mini Kit (QIAGEN) and transcribed into cDNA using random primers (dN6, Roche). cDNA were quantified with RT-PCR using TaqMan gene expression assays (Applied Biosystems). The levels of four housekeeping genes including ActB, GAPDH, HPRT, and PPIA were averaged and used for normalization. For the microarray experiment, the quality of total RNA was checked by gel analysis using the total RNA Nano chip assay on an Agilent 2100 Bioanalyzer (Agilent Technologies GmbH). Only samples with RNA index values >7 were selected for expression profiling. The analysis is done with R on the GeneView data produced by the scanner.

### Cell Culture

Mouse brain tumors were dissected and then digested with Accutase (Sigma-Aldrich). Dissociated cells were grown to neurospheres in DMEM/F12 medium containing 20 ng ml<sup>-1</sup> EGF (Sigma-Aldrich), 10 ng ml<sup>-1</sup> FGF2 (Sigma-Aldrich), B27 (GIBCO), and ITSS (Roche). To establish monolayer culture, neurospheres were dissociated with Accutase (Sigma-Aldrich) and were seeded on laminin (Roche) and poly-L-lysine (Sigma-Aldrich)-coated cell culture plates. For the differentiation of NSCs, NSCs were dissociated with Accutase, and

2.5 × 10<sup>4</sup> cells were seeded on laminin and poly-L-lysine-coated coverslips in one well of a 24-well plate. Cells were cultured for 14 days in NSCs medium without EGF and with 5 ng ml<sup>-1</sup> FGF2. After 14 days, cells were collected for analysis. Medium was renewed every 3 days for the whole procedure. For *in vitro* CreERT2 induction, 4-hydroxytamoxifen (4-OHT, H7904; Sigma) was dissolved in ethanol at a final stock concentration of 10 mM and kept in single-use aliquots in the dark at -20°C and added freshly to the medium. To induce Cre activity in CreERT2-expressing BTSCs, medium was replaced with the same medium but containing 5 M 4-OHT for 72 hr, then the medium was again replaced with fresh medium without 4-OH for 4 days.

### Immunohistochemistry and Western Blot

Mice were perfused with 4% paraformaldehyde, and the brains were postfixed overnight at 4°C. Vibratome sections (50 μm) or 5 μm paraffin sections were blocked in 5% normal swine serum in PBST (PBS + 0.2%, Triton X-100) and incubated overnight at 4°C with the primary antibody. The apoptosis assay was performed by using the ApopTag Plus Peroxidase In Situ Apoptosis Kit (Chemicon). The senescence analysis was carried out by using the Senescence β-Galactosidase Staining Kit (Cell Signaling). Fluorescent images were captured using a confocal laser-scanning microscope (LSM700, Zeiss). Statistically significant results were followed up with Student's *t* tests. For western blot, protein extracts cultured brain tumor stem cells were subjected to electrophoresis and transferred onto a PVDF membrane for immunoblot analysis. The following antibodies were used: Tlx (1:500; rabbit), beta tubulin (1:1,000; Cell Signaling).

### ACCESSION NUMBERS

The microarray data have been deposited in the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) and are accessible through GEO Series accession number GSE46125.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2014.04.007>.

### ACKNOWLEDGMENTS

We thank Günther Schütz for providing mice, the Imaging Facility of the Deutsches Krebsforschungszentrum (DKFZ) and the Carl Zeiss Imaging Center at the DKFZ for their support with image acquisition, the DKFZ Genomic Facility for their support in the microarray experiment, Eric Holland for providing the RCAS constructs, and Weijun Feng and Patricia Rusu for critical reading of the manuscript. This work was supported by the Helmholtz Association (VH-NG-702), the Deutsche Forschungsgemeinschaft (LI 2140/1-1), the Deutsche Krebshilfe (110226), the Helmholtz Alliance Preclinical Comprehensive Cancer Center (PCCC), and the DKFZ Intramural Grant Program. This work was also supported in part by a grant from the DKFZ-Bayer Healthcare Alliance.

Received: December 15, 2012

Revised: December 3, 2013

Accepted: April 10, 2014

Published: May 15, 2014

### REFERENCES

- Barrett, L.E., Granot, Z., Coker, C., Iavarone, A., Hambardzumyan, D., Holland, E.C., Nam, H.S., and Benezra, R. (2012). Self-renewal does not predict tumor growth potential in mouse models of high-grade glioma. *Cancer Cell* 21, 11–24.
- Belz, T., Liu, H.K., Bock, D., Takacs, A., Vogt, M., Wintermantel, T., Brandwein, C., Gass, P., Greiner, E., and Schütz, G. (2007). Inactivation of the gene for the nuclear receptor tailless in the brain preserving its function in the eye. *Eur. J. Neurosci.* 26, 2222–2227.

- Bickenbach, J.R. (1981). Identification and behavior of label-retaining cells in oral mucosa and skin. *J. Dent. Res. 60 Spec No C (Spec No)*, 1611–1620.
- Brill, M.S., Snappyan, M., Wohlfrom, H., Ninkovic, J., Jawerka, M., Mastick, G.S., Ashery-Padan, R., Saghatelian, A., Berninger, B., and Götz, M. (2008). A *dlx2*- and *pax6*-dependent transcriptional code for periglomerular neuron specification in the adult olfactory bulb. *J. Neurosci.* 28, 6439–6452.
- Chen, J., Li, Y., Yu, T.S., McKay, R.M., Burns, D.K., Kernie, S.G., and Parada, L.F. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* 488, 522–526.
- Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. *Nat. Med.* 17, 313–319.
- Corsini, N.S., Sancho-Martinez, I., Laudenklos, S., Glasgow, D., Kumar, S., Letellier, E., Koch, P., Teodorczyk, M., Kleber, S., Klussmann, S., et al. (2009). The death receptor CD95 activates adult neural stem cells for working memory formation and brain repair. *Cell Stem Cell* 5, 178–190.
- Driessens, G., Beck, B., Caauwe, A., Simons, B.D., and Blanpain, C. (2012). Defining the mode of tumour growth by clonal analysis. *Nature* 488, 527–530.
- Feng, W., Khan, M.A., Bellvis, P., Zhu, Z., Bernhardt, O., Herold-Mende, C., and Liu, H.K. (2013). The chromatin remodeler CHD7 regulates adult neurogenesis via activation of SoxC transcription factors. *Cell Stem Cell* 13, 62–72.
- Furnari, F.B., Fenton, T., Bachoo, R.M., Mukasa, A., Stommel, J.M., Stegh, A., Hahn, W.C., Ligon, K.L., Louis, D.N., Brennan, C., et al. (2007). Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes Dev.* 21, 2683–2710.
- Gil, J., and Peters, G. (2006). Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. *Nat. Rev. Mol. Cell Biol.* 7, 667–677.
- Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E., and Heintz, N. (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 425, 917–925.
- He, S., Nakada, D., and Morrison, S.J. (2009). Mechanisms of stem cell self-renewal. *Annu. Rev. Cell Dev. Biol.* 25, 377–406.
- Holland, E.C. (2001). Gliomagenesis: genetic alterations and mouse models. *Nat. Rev. Genet.* 2, 120–129.
- Holland, E.C., Hively, W.P., DePinho, R.A., and Varmus, H.E. (1998). A constitutively active epidermal growth factor receptor cooperates with disruption of G1 cell-cycle arrest pathways to induce glioma-like lesions in mice. *Genes Dev.* 12, 3675–3685.
- Kay, M.M. (1975). Mechanism of removal of senescent cells by human macrophages in situ. *Proc. Natl. Acad. Sci. USA* 72, 3521–3525.
- Kretzschmar, K., and Watt, F.M. (2012). Lineage tracing. *Cell* 148, 33–45.
- Li, L., and Clevers, H. (2010). Coexistence of quiescent and active adult stem cells in mammals. *Science* 327, 542–545.
- Liu, H.K., Belz, T., Bock, D., Takacs, A., Wu, H., Lichter, P., Chai, M., and Schütz, G. (2008). The nuclear receptor *tailless* is required for neurogenesis in the adult subventricular zone. *Genes Dev.* 22, 2473–2478.
- Liu, H.K., Wang, Y., Belz, T., Bock, D., Takacs, A., Radlwimmer, B., Barbus, S., Reifemberger, G., Lichter, P., and Schütz, G. (2010). The nuclear receptor *tailless* induces long-term neural stem cell expansion and brain tumor initiation. *Genes Dev.* 24, 683–695.
- Magee, J.A., Piskounova, E., and Morrison, S.J. (2012). Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer Cell* 21, 283–296.
- Monaghan, A.P., Grau, E., Bock, D., and Schütz, G. (1995). The mouse homolog of the orphan nuclear receptor *tailless* is expressed in the developing forebrain. *Development* 121, 839–853.
- Nam, H.S., and Benezra, R. (2009). High levels of *Id1* expression define B1 type adult neural stem cells. *Cell Stem Cell* 5, 515–526.
- Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. *Nature* 414, 105–111.
- Rowland, B.D., and Peeper, D.S. (2006). KLF4, p21 and context-dependent opposing forces in cancer. *Nat. Rev. Cancer* 6, 11–23.
- Schepers, A.G., Snippert, H.J., Stange, D.E., van den Born, M., van Es, J.H., van de Wetering, M., and Clevers, H. (2012). Lineage tracing reveals *Lgr5*+ stem cell activity in mouse intestinal adenomas. *Science* 337, 730–735.
- Shah, N.M., Groves, A.K., and Anderson, D.J. (1996). Alternative neural crest cell fates are instructively promoted by TGFβ superfamily members. *Cell* 85, 331–343.
- Sharpless, N.E., and DePinho, R.A. (2007). How stem cells age and why this makes us grow old. *Nat. Rev.* 8, 703–713.
- Shi, Y., Chichung Lie, D., Taupin, P., Nakashima, K., Ray, J., Yu, R.T., Gage, F.H., and Evans, R.M. (2004). Expression and function of orphan nuclear receptor TLX in adult neural stem cells. *Nature* 427, 78–83.
- Singec, I., Knoth, R., Meyer, R.P., Maciaczyk, J., Volk, B., Nikkhah, G., Frotscher, M., and Snyder, E.Y. (2006). Defining the actual sensitivity and specificity of the neurosphere assay in stem cell biology. *Nat. Methods* 3, 801–806.
- Singh, S.K., Clarke, I.D., Terasaki, M., Bonn, V.E., Hawkins, C., Squire, J., and Dirks, P.B. (2003). Identification of a cancer stem cell in human brain tumors. *Cancer Res.* 63, 5821–5828.
- Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. *Nature* 429, 396–401.
- Snippert, H.J., van der Flier, L.G., Sato, T., van Es, J.H., van den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., van Rheenen, J., Simons, B.D., and Clevers, H. (2010). Intestinal crypt homeostasis results from neutral competition between symmetrically dividing *Lgr5* stem cells. *Cell* 143, 134–144.
- Sun, G., Yu, R.T., Evans, R.M., and Shi, Y. (2007). Orphan nuclear receptor TLX recruits histone deacetylases to repress transcription and regulate neural stem cell proliferation. *Proc. Natl. Acad. Sci. USA* 104, 15282–15287.
- Vescovi, A.L., Galli, R., and Reynolds, B.A. (2006). Brain tumour stem cells. *Nat. Rev. Cancer* 6, 425–436.
- Zhang, C.L., Zou, Y., Yu, R.T., Gage, F.H., and Evans, R.M. (2006). Nuclear receptor TLX prevents retinal dystrophy and recruits the corepressor *atrophin1*. *Genes Dev.* 20, 1308–1320.
- Zhang, C.L., Zou, Y., He, W., Gage, F.H., and Evans, R.M. (2008). A role for adult TLX-positive neural stem cells in learning and behaviour. *Nature* 451, 1004–1007.
- Zou, Y., Niu, W., Qin, S., Downes, M., Burns, D.K., and Zhang, C.L. (2012). The nuclear receptor TLX is required for gliomagenesis within the adult neurogenic niche. *Mol. Cell. Biol.* 32, 4811–4820.

# Introducing the China Gateway

The home for Chinese scientists



The new Cell Press China Gateway is the ultimate destination for Chinese scientists. Now highlights of the latest research, regional events, and featured content from Cell Press are accessible online, on a single webpage.

The Cell Press China Gateway was designed with you in mind. We hope you'll make yourself at home.

Visit us at [www.cell.com/china](http://www.cell.com/china)

# In Vivo Direct Reprogramming of Reactive Glial Cells into Functional Neurons after Brain Injury and in an Alzheimer's Disease Model

Ziyuan Guo,<sup>1</sup> Lei Zhang,<sup>1</sup> Zheng Wu,<sup>1</sup> Yuchen Chen,<sup>1</sup> Fan Wang,<sup>1</sup> and Gong Chen<sup>1,\*</sup>

<sup>1</sup>Department of Biology, The Huck Institutes of Life Sciences, The Pennsylvania State University, University Park, PA 16802, USA

\*Correspondence: gongchen@psu.edu

<http://dx.doi.org/10.1016/j.stem.2013.12.001>

## SUMMARY

Loss of neurons after brain injury and in neurodegenerative disease is often accompanied by reactive gliosis and scarring, which are difficult to reverse with existing treatment approaches. Here, we show that reactive glial cells in the cortex of stab-injured or Alzheimer's disease (AD) model mice can be directly reprogrammed into functional neurons *in vivo* using retroviral expression of a single neural transcription factor, NeuroD1. Following expression of NeuroD1, astrocytes were reprogrammed into glutamatergic neurons, while NG2 cells were reprogrammed into glutamatergic and GABAergic neurons. Cortical slice recordings revealed both spontaneous and evoked synaptic responses in NeuroD1-converted neurons, suggesting that they integrated into local neural circuits. NeuroD1 expression was also able to reprogram cultured human cortical astrocytes into functional neurons. Our studies therefore suggest that direct reprogramming of reactive glial cells into functional neurons *in vivo* could provide an alternative approach for repair of injured or diseased brain.

## INTRODUCTION

Gliosis is a common pathological process after brain injury that involves the activation of glial cells to proliferate and become hypertrophic to occupy the injured brain areas (Pekny and Nilsson, 2005; Robel et al., 2011; Sofroniew and Vinters, 2010). Glial cells, including astrocytes, NG2 cells, and microglia, undergo reactive response to injury in order to form a defense system against the invasion of micro-organisms and cytotoxins into surrounding tissue (Pekny and Nilsson, 2005; Robel et al., 2011; Sofroniew and Vinters, 2010). However, once activated, many reactive glial cells will stay in the injury sites and secrete neuroinhibitory factors to prevent neuronal growth, eventually forming glial scar inside the brain (Sofroniew and Vinters, 2010). Reactive glial cells have also been widely reported after stroke, spinal cord injury, glioma, and neurodegenerative disorders such as Alzheimer's disease (AD) (Gwak et al., 2012; Pekny and Nilsson, 2005; Sofroniew and Vinters, 2010; Verkhratsky et al., 2010, 2012). However, despite substantial progress in

understanding the molecular pathways of reactive gliosis (Robel et al., 2011), there has been little success in efforts to reverse glial scarring after its formation.

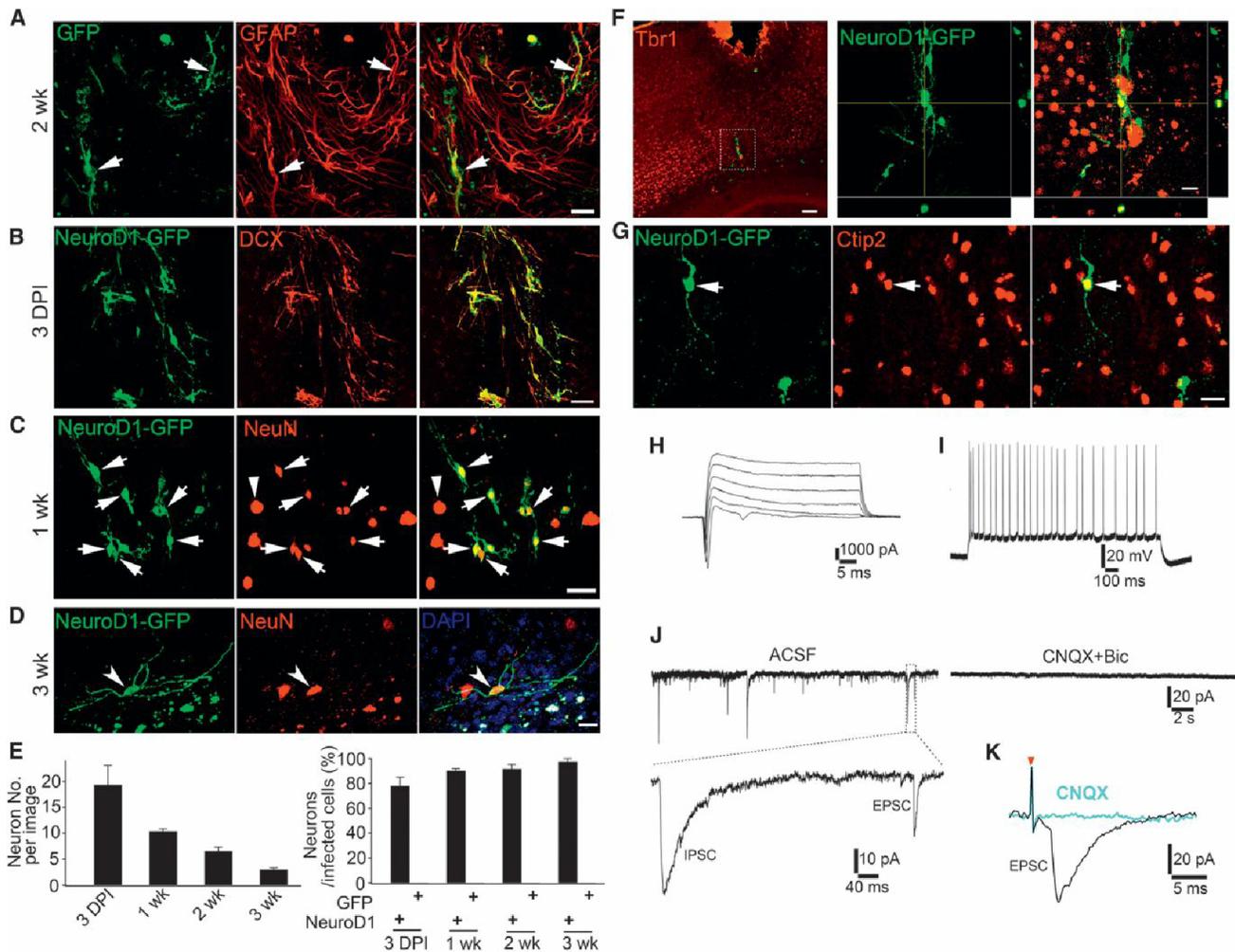
Reprogramming adult skin fibroblasts into pluripotent stem cells has opened a new field for potential stem cell therapy (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007). Many studies have since demonstrated transdifferentiation across different cell lineages, including reprogramming mouse or human fibroblasts directly into neurons (Ambasudhan et al., 2011; Caiazza et al., 2011; Kim et al., 2011; Ladewig et al., 2012; Liu et al., 2012, 2013; Meng et al., 2012; Pang et al., 2011; Pfisterer et al., 2011; Qiang et al., 2011; Son et al., 2011; Torper et al., 2013; Vierbuchen et al., 2010; Yoo et al., 2011) or oligodendroglial cells (Najm et al., 2013; Yang et al., 2013). It has also been demonstrated that astroglial cells can be transdifferentiated into neurons (Heinrich et al., 2010; Torper et al., 2013) or reprogrammed into neuroblast cells (Niu et al., 2013). However, it is unclear whether such transdifferentiation studies can be applied to brain repair after brain injury or neurodegeneration.

We demonstrate here that after brain injury, reactive glial cells including both astrocytes and NG2 cells can be reprogrammed into functional neurons in the adult mouse cortex when infected with retrovirus encoding a single transcription factor, NeuroD1. Electrophysiological recordings revealed both spontaneous and evoked synaptic responses in NeuroD1-converted neurons. Interestingly, astrocytes were mainly reprogrammed into glutamatergic neurons whereas NG2 cells were reprogrammed into both glutamatergic and GABAergic neurons after NeuroD1 expression. We also demonstrated that forced expression of NeuroD1 in a mouse model for AD was capable of reprogramming reactive glial cells into functional neurons. Furthermore, NeuroD1 was capable of reprogramming cultured human astrocytes into functional neurons efficiently. Thus, *in vivo* regeneration of functional neurons from reactive glial cells may provide a potential therapeutic approach to restore lost neuronal function in injured or diseased brain.

## RESULTS

### In Vivo Reprogramming of Reactive Glial Cells into Functional Neurons after Brain Injury

A signature of brain injury is the loss of functional neurons and the activation of glial cells. In the adult mouse cortex, astrocytes are usually quiescent and not proliferative unless activated by

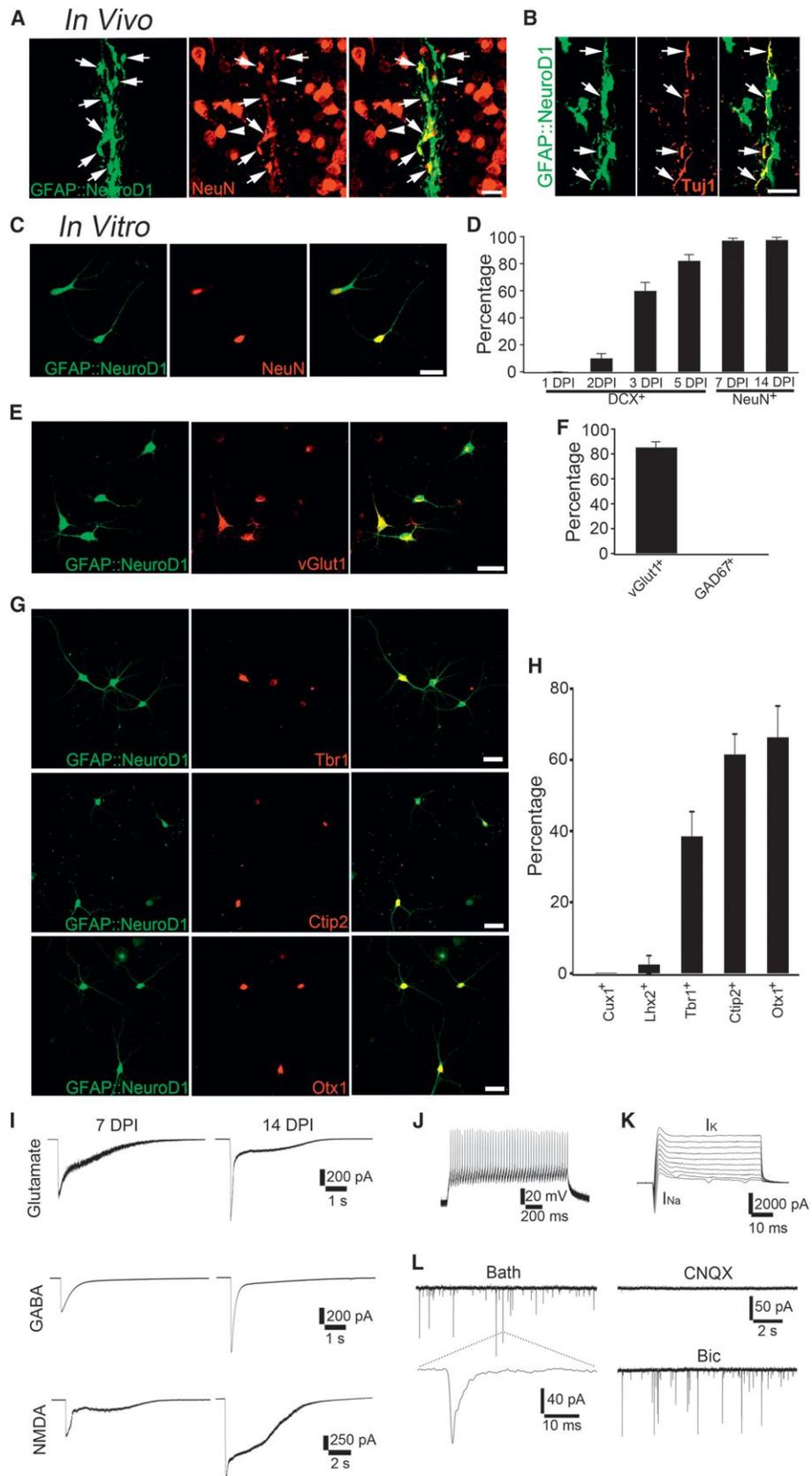


**Figure 1. In Vivo Conversion of Reactive Glial Cells into Functional Neurons after Brain Injury**

(A) Injecting control retrovirus expressing GFP (green) into mouse cortex revealed GFAP-positive reactive astrocytes (red) in the injury site (14 days postinjection, DPI).  
 (B and C) NeuroD1-IRES-GFP-infected cells (green) were immunopositive for neuronal markers DCX (B, 3 DPI) and NeuN (C, 7 DPI). Note a significant number of NeuN-positive neurons in the injury site after NeuroD1 infection.  
 (D) After 21 DPI, NeuroD1-converted neurons (NeuN-positive, arrowhead) showed extensive neurites. Scale bar, 20  $\mu$ m for (A) and (D); 40  $\mu$ m for (B) and (C).  
 (E) Quantified data showing the number of converted neurons per imaged area (40 $\times$ , 0.1 mm<sup>2</sup>) and conversion efficiency after NeuroD1 infection.  
 (F and G) NeuroD1-converted neurons were immunopositive for cortical neuron marker Tbr1 (F) and deep layer marker Ctip2 (G, 12 DPI). Scale bars: 100  $\mu$ m for low-power image, 40  $\mu$ m for high-power image.  
 (H and I) Representative traces from cortical slice recordings showing Na<sup>+</sup> and K<sup>+</sup> currents (H) and repetitive action potentials (I) in NeuroD1-converted neurons (30 DPI).  
 (J) Representative traces showing spontaneous synaptic events in a NeuroD1-converted neuron (26 DPI) in cortical slice recording (CNQX, 10  $\mu$ M; BIC, 20  $\mu$ M).  
 (K) Evoked synaptic events recorded from a converted neuron.  
 See also Figures S1–S3.

injury or diseases (Ge et al., 2012; Robel et al., 2011; Tsai et al., 2012). Besides astrocytes, NG2 cells and microglia can also be activated and proliferate rapidly in the injury sites or in diseased brain (Aguzzi et al., 2013; Hines et al., 2009; Kang et al., 2013). To test whether reactive glial cells can be reprogrammed into functional neurons for brain repair, we decided to inject retroviruses encoding neural transcription factors into adult mouse cortex in vivo. We chose retroviral delivery for in vivo injection because, unlike lentiviruses or adeno-associated viruses, retroviruses only infect dividing cells such as progenitor cells or reactive glial cells,

and do not infect nondividing cells such as neurons (Zhao et al., 2006). As a control, we first injected retroviruses expressing GFP alone under the control of CAG promoter (pCAG-GFP-IRES-GFP) (Zhao et al., 2006) into mouse cortex to examine what type of cells will be infected by the retrovirus after stab injury. As expected, many GFP-labeled cells were immunopositive for astrocytic marker GFAP (Figure 1A; 52.1%  $\pm$  4.3% were GFAP positive, n = 3 animals). We did not observe any neuronal cells infected by control retrovirus expressing GFP alone (Figure S1 available online).



(legend on next page)

Our strategy for reprogramming reactive glial cells into neurons involved construction of a retrovirus encoding NeuroD1, a bHLH proneural transcription factor that plays an important role during embryonic brain development and adult neurogenesis (Cho and Tsai, 2004; Gao et al., 2009; Kuwabara et al., 2009). We first tested the effect of NeuroD1 in a brain injury model, where reactive glial cells were induced by stab injury during stereotaxic injection of retroviruses into mouse somatosensory cortex. We limited our injection to cortical areas without penetrating the hippocampus or subventricular zone, where adult neural stem cells are known to reside. Interestingly, 3 days postinjection (DPI) of the retrovirus encoding NeuroD1 (pCAG-NeuroD1-IRES-GFP) into mouse cortex, many NeuroD1-GFP-infected cells showed bipolar morphology and were immunopositive for doublecortin (DCX), an immature neuronal marker (Figure 1B). One week after viral injection, NeuroD1-infected cells started to show staining for neuronal nuclei (NeuN), a typical neuronal marker (Figure 1C). Three weeks after viral injection, NeuroD1-infected cells showed extensive neurites and the NeuN signal reached the level of noninfected mature neurons in the same vicinity (Figure 1D). Quantitatively, we detected a large number of NeuroD1-GFP-labeled newborn neurons (DCX) at 3 DPI ( $19.3 \pm 3.7$  per  $0.1 \text{ mm}^2$ ,  $n = 5$  animals), and the number of converted neurons gradually declined during the maturation process (Figure 1E). Nevertheless, at any given time point after NeuroD1 retroviral infection, the majority of NeuroD1-infected cells were DCX- or NeuN-positive neurons, whereas control GFP viral infection resulted in no neurons at all (Figure 1E; Figure S1).

We found that NeuroD1-converted neurons were usually located in the deep cortical layer, with some exceptions in the cingulate cortex or superficial layer of the cortex, as illustrated with a general cortical neuron marker *Tbr1* (Figure 1F). To further test the neuronal properties of NeuroD1-converted neurons, we used the deep layer cortical neuron marker *Ctip2* and found that NeuroD1-converted neurons were indeed immunopositive for *Ctip2* (Figure 1G). No GFP-labeled neurons were detected in the dentate gyrus or the subventricular zone, because our viral injection was restricted to the cortical layers.

Interestingly, we found that NeuroD1-converted neurons at 3 DPI were typically localized within  $100 \mu\text{m}$  from the injection site. However, 1–2 weeks after injection, NeuroD1-infected cells were found in more broad areas, ranging from  $100\text{--}500 \mu\text{m}$  away from the injection site (Figure S2). It is possible that this gradual spread is a reflection of distant cells being exposed to a lower

level of viral infection and thus having later conversion or that the newly converted neurons migrate away from the injury core. Further investigation will be needed to distinguish between these possibilities.

To test whether NeuroD1-converted neurons have functional activity, we performed cortical slice recordings on NeuroD1-GFP-infected cells  $\sim 1$  month after they received retroviral injection. The NeuroD1-converted neurons showed large sodium currents ( $3,840 \pm 302 \text{ pA}$ ,  $n = 5$ ) and potassium currents ( $4,672 \pm 602 \text{ pA}$ ,  $n = 5$ ) (Figure 1H) and were capable of firing repetitive action potentials (Figure 1I,  $n = 4$ ). Importantly, we recorded robust spontaneous synaptic events in NeuroD1-converted neurons in cortical slice recordings (Figure 1J; frequency,  $1.96 \pm 0.43 \text{ Hz}$ ; amplitude,  $23.7 \pm 2.0 \text{ pA}$ ;  $n = 8$ ; 25–31 DPI), suggesting that these NeuroD1-converted neurons formed functional synapses with other neurons. Moreover, we placed a stimulating electrode nearby ( $50 \mu\text{m}$ ) to stimulate axon fibers and recorded evoked synaptic responses in the NeuroD1-converted neurons (Figure 1K;  $n = 3$  animals), suggesting an integration of converted neurons into local neural circuits. The NeuroD1-converted neurons can survive for a long time in mouse brain *in vivo*, at least 2 months after the retroviral injection, and they showed clear dendritic spines and large spontaneous synaptic events (Figure S3,  $n = 3$  animals). Thus, NeuroD1 can reprogram brain-injury-induced reactive glial cells into functional neurons in mouse brain *in vivo*.

### NeuroD1 Reprograms Astrocytes into Glutamatergic Neurons

After a brain injury, many glial cells are activated and become proliferative. To examine whether reactive astrocytes, a major subtype of reactive glial cells in gliosis, can be reprogrammed into neurons by NeuroD1, we generated a retrovirus expressing NeuroD1 under the control of human GFAP promoter. After injecting GFAP::NeuroD1-IRES-GFP retrovirus into the mouse cortex, we found that NeuroD1-infected cells were mostly immunopositive for the neuronal markers NeuN (Figure 2A;  $92.8\% \pm 2\%$  were NeuN positive, 8 DPI;  $n = 4$  animals) and Tuj1 (Figure 2B;  $n = 4$  animals). Therefore, reactive astrocytes induced by brain injury can be reprogrammed into neurons *in vivo* after overexpressing a single transcription factor, NeuroD1.

To further characterize the properties of astrocyte-converted neurons, we infected cultured mouse cortical astrocytes with GFAP::NeuroD1-IRES-GFP retrovirus. Our mouse astrocyte

### Figure 2. NeuroD1 Converts Astrocytes into Glutamatergic Neurons

(A and B) *In vivo* injection of GFAP-promoter-driven NeuroD1-IRES-GFP (green) retrovirus revealed astrocyte-converted neurons immunopositive for NeuN (A) and Tuj1 (B).

(C) Cultured mouse cortical astrocytes were converted into NeuN-positive neurons.

(D) Time course of GFAP::NeuroD1 conversion efficiency after infecting cultured mouse astrocytes.

(E and F) Astrocyte-converted neurons were positive for VGluT1 (E) but negative for GAD67 (F).

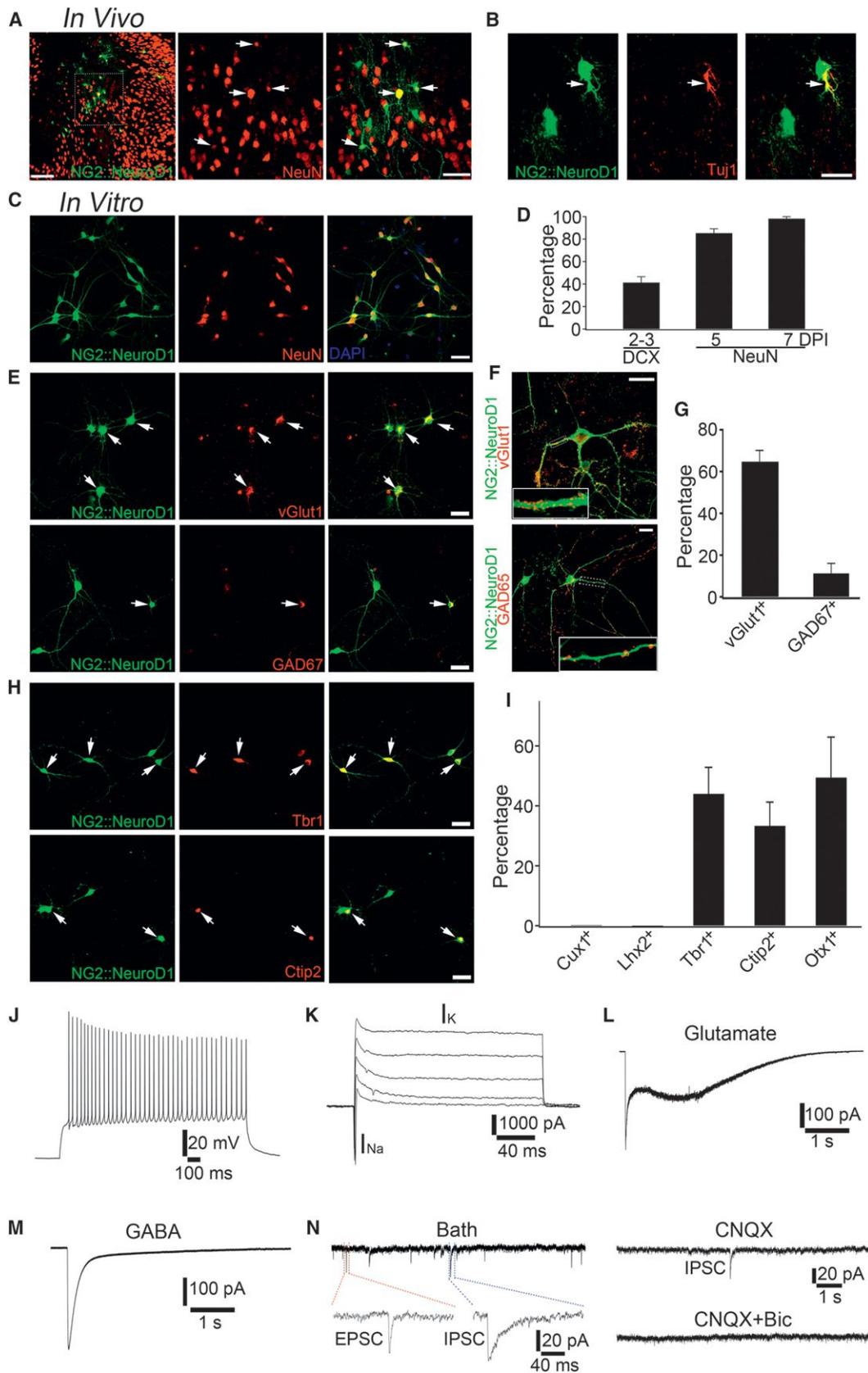
(G and H) Immunostaining with cortical layer neuronal markers showed deep layer neuronal properties (*Ctip2* and *Otx1*) after NeuroD1-induced conversion. Scale bars:  $20 \mu\text{m}$  for (A–C) and (E), and  $40 \mu\text{m}$  for (G).

(I) Mouse astrocyte-converted neurons showed large glutamate, GABA, and NMDA receptor currents within 2 weeks after NeuroD1 infection. Average GABA current, 7 DPI,  $405 \pm 97 \text{ pA}$ ,  $n = 8$ ; 14 DPI,  $861 \pm 55 \text{ pA}$ ,  $n = 13$ . Average glutamate current, 7 DPI,  $517 \pm 145 \text{ pA}$ ,  $n = 7$ ; 14 DPI,  $1,060 \pm 159 \text{ pA}$ ,  $n = 9$ . Average NMDA current, 7 DPI,  $676 \pm 118 \text{ pA}$ ,  $n = 7$ ; 14 DPI,  $1,315 \pm 95$ ,  $n = 7$ .

(J and K) Mouse astrocyte-converted neurons showed repetitive action potentials (J) and large  $I_{\text{Na}}$  and  $I_{\text{K}}$  (K).

(L) Spontaneous synaptic events recorded from mouse astrocyte-converted neurons. All events were blocked by CNQX but not BIC, suggesting that they were glutamatergic events.

Also see Figure S4.



(legend on next page)

cultures were enriched with GFAP-positive astrocytes, with little contamination from microglia or NG2 cells (Figure S4A). Similar to our *in vivo* reprogramming, cultured mouse astrocytes can be efficiently reprogrammed into NeuN-positive neurons after GFAP::NeuroD1 infection (Figure 2C). Quantitatively, we found that the NeuroD1-induced reprogramming efficiency dramatically increased in the first 3 days after GFAP::NeuroD1 infection and reached >90% after 7 DPI (Figure 2D). To examine whether GFAP::NeuroD1-converted neurons are glutamatergic or GABAergic, we performed immunostaining with VGluT1 and GAD67-specific antibodies and found that the majority of neurons were positive for VGluT1 (Figures 2E and 2F) but virtually none stained for GAD67 (Figure 2F; Figure S4B). Thus, NeuroD1 reprograms astrocytes into glutamatergic neurons. We also used superficial and deep layer cortical neuronal markers to characterize the NeuroD1-converted neurons. The GFAP::NeuroD1-converted neurons were positive for the deep layer markers Ctip2 and Otx1, as well as Tbr1 (Figures 2G and 2H), but negative for the superficial layer markers Cux1 and Lhx2 (Figure 2H; Figure S4C). This finding may explain why we observed more NeuroD1-converted neurons in deep cortical layer in mouse brain *in vivo*.

We also functionally characterized the cultured mouse astrocyte-converted neurons after NeuroD1 infection and detected large GABA (100  $\mu$ M), glutamate (100  $\mu$ M), and NMDA (100  $\mu$ M) currents (Figure 2I). The mouse astrocyte-converted neurons also showed repetitive action potentials (Figure 2J,  $n = 18$ ) and large  $\text{Na}^+$  and  $\text{K}^+$  currents (Figure 2K;  $I_{\text{Na}^+}$ ,  $2,979 \pm 626$  pA,  $n = 11$ ;  $I_{\text{K}}$ ,  $5,136 \pm 1,181$  pA,  $n = 11$ , 14 DPI). More importantly, we recorded robust synaptic events in astrocyte-converted neurons, which were blocked by glutamate receptor antagonist CNQX (10  $\mu$ M), but not by GABA<sub>A</sub> receptor antagonist bicuculline (BIC, 20  $\mu$ M) (Figure 2L; frequency,  $1.15 \pm 0.71$  Hz; amplitude,  $21.5 \pm 0.71$  pA,  $n = 13$ ; 14 DPI), confirming that astrocyte-converted neurons are glutamatergic neurons.

### NeuroD1 Reprograms NG2 Cells into Glutamatergic and GABAergic Neurons

In addition to reactive astrocytes, NG2 cells also proliferate significantly in response to brain injury and accounted for about  $19.0\% \pm 3.9\%$  of control retrovirus-infected cells in our stab injury model, whereas only 6% of infected cells were microglia. To investigate whether NG2 cells can be reprogrammed into neurons by NeuroD1, we generated a retrovirus expressing NeuroD1 under the control of the human NG2 promoter (NG2::NeuroD1-IRES-GFP). Interestingly, we discovered that injecting NG2::NeuroD1 retrovirus into mouse cortex also re-

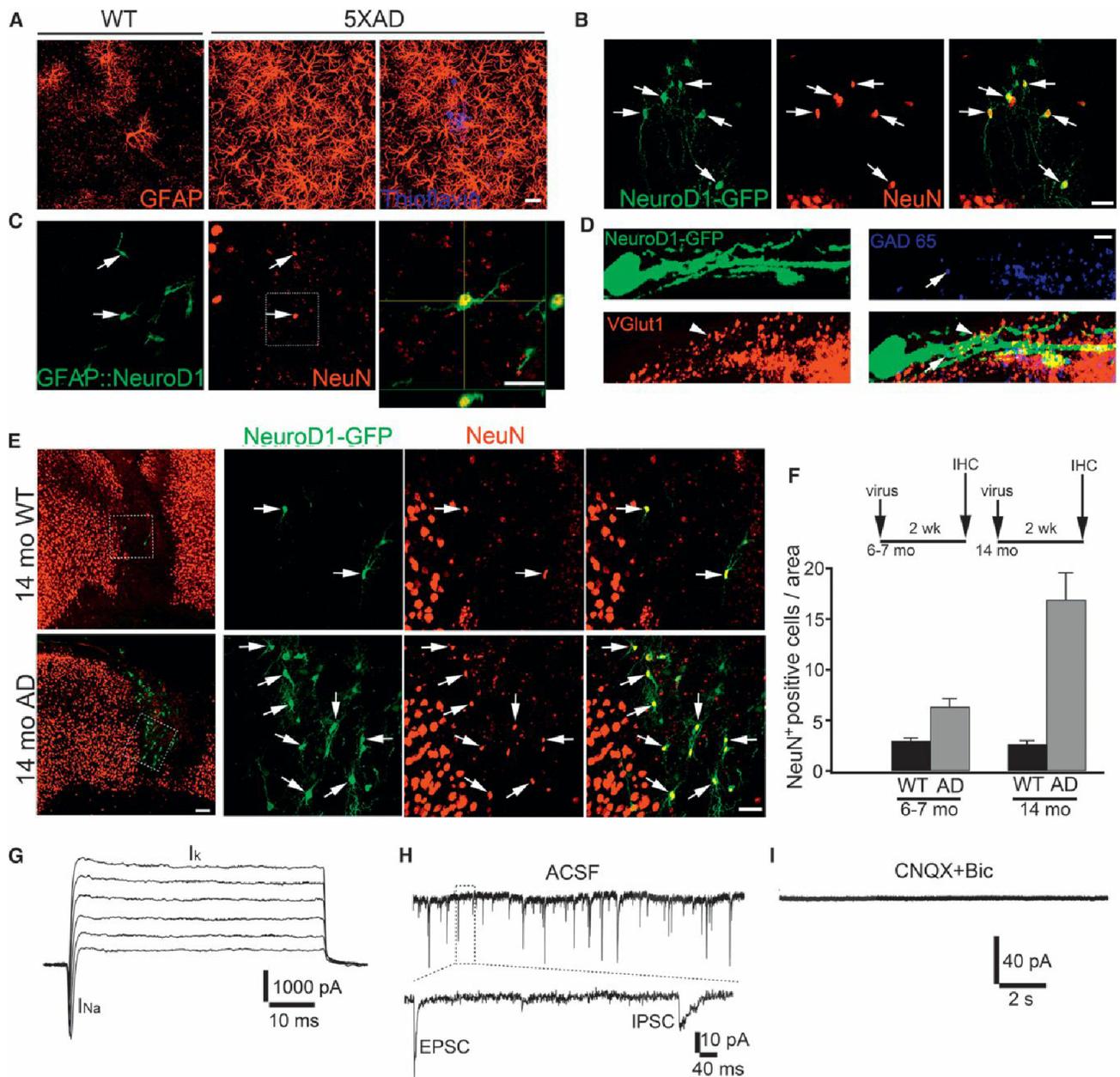
programmed NG2 cells into NeuN- and Tuj1-positive neurons (Figures 3A and 3B;  $42.5\% \pm 6.6\%$  GFP-labeled cells were NeuN positive, 8 DPI,  $n = 3$  animals). We then characterized NG2-converted neurons using cultured NG2 cells dissociated from mouse cortex. In our mouse NG2 cultures, the majority of cells were NG2 positive ( $79.2\% \pm 3.2\%$ ,  $n = 3$  repeats of cultures) (Figure S4D). Consistent with our *in vivo* study, we found that NG2::NeuroD1 also efficiently reprogrammed cultured mouse NG2 cells into neurons (Figures 3C and 3D; 7 DPI,  $98.2\% \pm 1.8\%$ ,  $n = 484$ , 4 repeats). While the majority of NG2::NeuroD1-converted neurons were also glutamatergic (VGluT1 positive), about 10% NG2-converted neurons were immunopositive for GAD67 and presynaptic GABAergic terminals (GAD65) were found on neuronal dendrites (Figures 3E–3G). Therefore, it appears that NG2 cells can be reprogrammed into both glutamatergic and GABAergic neurons after expressing NeuroD1. Immunostaining with cortical layer markers revealed that NG2-converted neurons also stained mainly for the deep layer markers Ctip2 and Otx1, but rarely for Cux1 and Lhx2 (Figures 3H and 3I; Figures S4E and S4F). Patch-clamp recordings demonstrated that NG2-converted neurons generated after NeuroD1 infection were able to fire repetitive action potentials (Figure 3J) and showed large  $\text{Na}^+$  and  $\text{K}^+$  currents (Figure 3K) and large glutamate- and GABA-evoked receptor currents (Figures 3L and 3M;  $I_{\text{Glu}} = 438 \pm 78$  pA,  $n = 7$ ;  $I_{\text{GABA}} = 496 \pm 32$  pA,  $n = 7$ ). Moreover, we detected both glutamatergic and GABAergic events in NG2-converted neurons (Figure 3N), confirming that NG2 cells can be reprogrammed into both glutamatergic and GABAergic neurons. Therefore, a single transcription factor, NeuroD1, not only reprograms astrocytes into glutamatergic neurons, but also reprograms NG2 cells into glutamatergic and GABAergic neurons.

### Reactive Glia-Neuron Conversion in an AD Mouse Model

Besides activation by mechanical injury, reactive astrocytes have been widely reported in the cortex of AD patients or animal models (Rodríguez et al., 2009; Steele and Robinson, 2010). We employed a transgenic mouse model with AD (5xFAD) (Oakley et al., 2006) to test whether reactive astrocytes in the AD brain can be reprogrammed into functional neurons. We first confirmed that there were indeed many reactive astrocytes in the cortex of 5xFAD mice compared to WT (Figure 4A). Next, we injected NeuroD1-GFP retrovirus (CAG promoter) into the cortex of 5xFAD mice and observed NeuN-positive neuron-like cells (Figure 4B, 14–16 DPI). To further confirm that the reactive astrocytes in the AD model mouse brain can be reprogrammed into neurons, we injected GFAP::NeuroD1-GFP retrovirus to infect cortical astrocytes specifically, and indeed observed

### Figure 3. NeuroD1 Converts NG2 Cells into Glutamatergic and GABAergic Neurons

(A and B) *In vivo* injection of NG2::NeuroD1-GFP retrovirus revealed the conversion of NG2 cells into neuronal cells positive for NeuN (A) or Tuj1(B) (8 DPI). (C and D) Cultured NG2 cells were converted into NeuN-positive neurons within 1 week after infection by NG2::NeuroD1. (E–G) NG2 cell-converted neurons after NeuroD1 infection were immunopositive for both VGluT1 (>60%) and GAD67 (10%). VGluT1 and GAD65 immunostaining also showed glutamatergic and GABAergic puncta on converted neural dendrites (F). (H and I) Cortical layer neuronal marker immunostaining showed deep layer neuronal properties (Ctip2 and Otx1) after NeuroD1-induced conversion of NG2 cells. Scale bars: 100  $\mu$ m for left panel in (A); 40  $\mu$ m for (A, right two panels), (C), (E), and (H); 20  $\mu$ m for (B) and (F). (J and K) NG2-converted neurons showed repetitive action potentials (J;  $n = 9$ ) and large sodium and potassium currents (K;  $n = 10$ ). (L and M) NG2-converted neurons showed large glutamate-evoked current (L;  $n = 7$ ) and GABA-evoked current (M;  $n = 7$ ). (N) Spontaneous synaptic events recorded from NG2-converted neurons showed both glutamatergic and GABAergic events, confirming that NeuroD1 can convert NG2 cells into both excitatory and inhibitory neurons.



**Figure 4. NeuroD1 Converts Reactive Glial Cells into Functional Neurons in AD Mouse Brain In Vivo**

(A) Reactive astrocytes (labeled by GFAP, red) in 5xFAD mouse cortex (5 months old) were significantly increased compared to that in WT cortex. A $\beta$  plaques were labeled by thioflavin-S (blue).

(B) NeuroD1-infected cells (16 DPI) in AD mouse cortex (7 months old) showed clear neuron-like morphology (green) and NeuN staining (red).

(C) Injecting GFAP::NeuroD1 retrovirus into AD cortex also converted astrocytes into NeuN-positive neurons (7 DPI).

(D) NeuroD1-converted neurons in the AD brain were innervated by glutamatergic (VGLUT1, red) and GABAergic terminals (GAD65, blue). Scale bars: 20  $\mu$ m for (A) and (C); 40  $\mu$ m for (B); 5  $\mu$ m for (D).

(E) Efficient induction of many new neurons in 14-month-old AD animals after NeuroD1-GFP retroviral infection. Scale bar: 100  $\mu$ m for low-power image, 40  $\mu$ m for high-power image.

(F) Quantified data showing enhanced neural conversion in AD animals compared to WT animals, likely due to more reactive glial cells in old AD brain.

(G) Representative traces of sodium and potassium currents recorded from NeuroD1-infected cells in AD cortical slices.

(H) Spontaneous synaptic events recorded from NeuroD1-converted neurons (28 DPI) in AD cortical slices.

(I) All synaptic events were blocked by CNQX (10  $\mu$ M) and BIC (20  $\mu$ M).

NeuroD1-converted neurons labeled by NeuN (Figure 4C). Immunostaining with VGLUT1 and GAD65 revealed glutamatergic and GABAergic terminals innervating NeuroD1-converted neu-

rons in the 5xFAD mouse brain (Figure 4D). Because AD is a progressive neurodegenerative disorder, we wondered whether in vivo reprogramming could occur in very old animals.

Remarkably, when we injected NeuroD1-GFP retrovirus into 14-month-old 5xFAD animals, we observed many NeuroD1-GFP infected cells immunopositive for NeuN (Figure 4E). Interestingly, we found that the number of NeuroD1-converted neurons was higher in 5xFAD than in WT mouse brain and even higher in 14-month-old animals than in 7-month-old animals (Figure 4F). This increase might occur because there are more reactive glial cells in older diseased animals that can be reprogrammed into neurons. Using cortical slice recordings, we further demonstrated that the NeuroD1-converted neurons in 5xFAD mouse brain were functional, with the peak amplitude of  $I_{Na}$  at  $2,270 \pm 282$  pA ( $n = 5$ ) and  $I_K$  at  $5,498 \pm 706$  pA ( $n = 5$ ) (Figure 4G). More importantly, we recorded robust synaptic events in NeuroD1-converted neurons in cortical slice recordings (Figures 4H and 4I; frequency,  $2.80 \pm 0.95$  Hz; amplitude,  $20.5 \pm 2.7$  pA;  $n = 7$ ), suggesting that these newly reprogrammed neurons are functionally connected with surrounding neurons in 5xFAD mouse brain. Therefore, overexpression of NeuroD1 in reactive glial cells has the potential to regenerate functional neurons in an AD model brain.

### Reprogramming Cultured Human Astrocytes into Functional Neurons

We next investigated whether NeuroD1 can reprogram human astrocytes into functional neurons using a human cortical astrocyte cell line (ScienCell, San Diego, CA). The majority of our cultured human astrocytes were immunopositive for GFAP and S100 $\beta$  (Figure 5A; Figure S5), but with very low level of neuroprogenitor marker Sox2 or Musashi (Figures S5A and S5B). We infected human astrocytes with GFAP::NeuroD1-IRES-GFP retrovirus and found that the majority of NeuroD1-infected cells were immunopositive for NeuN (Figure 5B), suggesting that human astrocytes can also be efficiently reprogrammed into neurons by expressing NeuroD1. We then examined the time course of NeuroD1-induced astrocyte-neuron conversion using a series of neuronal markers including DCX, NeuN, and MAP2. We found that the conversion efficiency increased dramatically between 3–5 DPI, with 90% of NeuroD1-infected human astrocytes becoming neurons by 5 DPI (Figures 5C–5F). Infection by NeuroD1 significantly changed the cell morphology from astrocytes to neurons, as shown by the phase contrast images in Figure 5G. To investigate whether NeuroD1-induced reprogramming involved a transient neuroprogenitor stage, we monitored the transdifferentiation process from 24 hr until 5 days after NeuroD1 infection of human astrocytes (Figures S5C and S5D). No transient increase in the expression level of the neural stem cell markers Sox2 (Figure S5C) or Musashi (Figure S5D) occurred during the early conversion period. In fact, after only 3 days of infection by NeuroD1, some astrocytes already became neuron-like cells with clearly extended neurites (Figures S5C and S5D). Therefore, NeuroD1 appears to reprogram astrocytes directly into neurons without transition through neuroprogenitor stage.

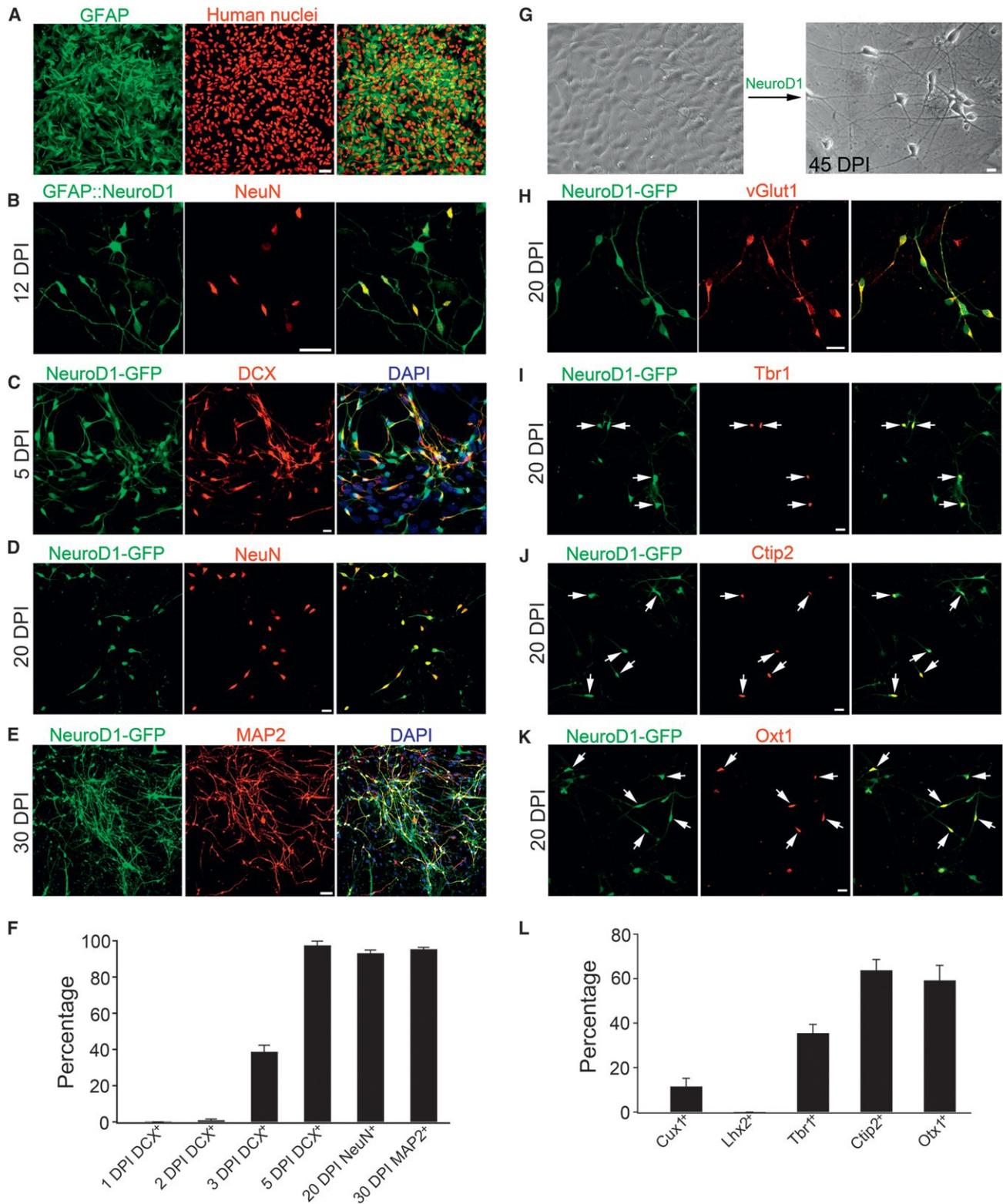
Next, we investigated what types of neurons were reprogrammed from human astrocytes. Immunostaining with VGluT1 and GAD67 revealed that human astrocytes infected by NeuroD1 were mainly reprogrammed into glutamatergic neurons as shown by immunopositivity for VGluT1 (Figure 5H), but not GAD67 (Figure S6B), consistent with our observations

for mouse astrocyte conversion. Using cortical layer markers, we found that, as for mouse astrocytes, human astrocyte-converted neurons also stained positive for the cortical neuron marker Tbr1 and the deep layer markers Ctip2 and Otx1 (Figures 5I–5L), but much less for the superficial layer markers Cux1 and Lhx2 (Figure 5L; Figures S6C and S6D). To investigate whether human microglia can be reprogrammed into neurons, we cultured human microglia and infected them with NeuroD1-GFP retrovirus, but did not detect any DCX-positive neurons (Figures S6E–S6H; 0 DCX+ neurons out of 33 NeuroD1-GFP infected microglial cells). However, this apparently different result might be influenced by the low infection efficiency of microglia by retrovirus ( $9.9\% \pm 0.8\%$ ,  $n = 3$  batches of culture), compared to the high infection efficiency of astrocytes ( $51.6\% \pm 2.9\%$ ,  $n = 3$  batches) or NG2 cells ( $57.8\% \pm 5\%$ ,  $n = 3$  batches).

To examine whether NeuroD1-converted human neurons are functionally connected, we performed immunostaining with the synaptic marker SV2 and the glutamatergic synapse marker VGluT1 (Figures 6A and 6B). After NeuroD1-induced conversion, we observed numerous SV2 puncta on MAP2-labeled neuronal dendrites (Figure 6A, 45 DPI). Some neurons even showed mushroom-like mature spines, which were colocalized with VGluT1 puncta (Figure 6B). Next, we employed patch-clamp recordings to test the function of reprogrammed human neurons. Human astrocyte-converted neurons started to show detectable NMDA receptor currents at 20 DPI, but very small GABA or glutamate receptor currents at this stage (Figure 6C). However, after 30–40 DPI, we detected large glutamate receptor currents ( $548 \pm 138$  pA,  $n = 7$ ; 31–35 DPI), GABA<sub>A</sub> receptor currents ( $599 \pm 114$  pA,  $n = 8$ ; 31–35 DPI), and NMDA receptor currents ( $966 \pm 101$  pA,  $n = 8$ ; 40 DPI) (Figures 6C and 6D). At 20 DPI, we also detected clear sodium ( $I_{Na}$ ) and potassium currents ( $I_K$ ) (Figure 6E), which increased dramatically by 40 DPI (Figures 6E and 6F). Accordingly, we recorded repetitive action potential firing in human astrocyte-converted neurons (Figure 6G,  $n = 15$ ). Furthermore, we detected functional synaptic events in NeuroD1-converted human neurons (frequency,  $1.6 \pm 0.3$  Hz; amplitude,  $23.2 \pm 0.8$  pA;  $n = 13$ ), which were blocked by CNQX (10  $\mu$ M) but not by BIC (20  $\mu$ M) (Figure 6H). Thus, it appears that NeuroD1 can reprogram human astrocytes into functional glutamatergic neurons.

### DISCUSSION

We show here that reactive glial cells generated after brain injury or in an AD model can be directly reprogrammed into functional neurons by a single transcription factor, NeuroD1, in mouse brain in vivo. Interestingly, after expressing the same transcription factor, astrocytes are mainly reprogrammed into glutamatergic neurons, whereas NG2 cells can be reprogrammed into both glutamatergic and GABAergic neurons. Such different cell fates after reprogramming by the same transcription factor may provide important clues regarding the lineage relationship between neurons and glial cells. We also show that human astrocytes in culture can be reprogrammed into functional neurons after expressing NeuroD1. The in vivo reprogramming of reactive glial cells into functional neurons after brain injury or in diseased mouse brain could potentially provide a therapeutic approach



**Figure 5. Conversion of Cultured Human Astrocytes into Functional Neurons**

(A) The majority of cultured human astrocytes were labeled by GFAP (green).

(B) Infection by GFAP::NeuroD1 retrovirus converted human astrocytes into NeuN-positive neurons.

(C–E) NeuroD1-induced conversion of human astrocytes into neurons as shown by a series of neuronal markers: DCX (C), NeuN (D), and MAP2 (E).

(F) Quantified data showing a significant increase of conversion efficiency during 3–5 DPI.

(legend continued on next page)

for treating reactive gliosis, which is widely associated with nerve injury and neurodegenerative disorders.

### Direct Reprogramming of Reactive Astrocytes into Functional Neurons

During mammalian brain development, neurogenesis typically occurs in embryonic stage before birth, while gliogenesis usually occurs in postnatal stage (Miller and Gauthier, 2007). In the central nervous system, neurons, astrocytes, oligodendrocytes, and NG2 cells (oligodendrocyte precursor cells) are all generated from neural stem cells, whereas microglial cells are generated from hematopoietic stem cells. Astrocytes maintain proliferative properties after differentiation from neural stem cells, and the majority of astrocytes in the cortex are generated locally by astrocytes themselves in the postnatal stage (Ge et al., 2012). Astrocyte proliferation largely stops after 1 month of age in rodents (Ge et al., 2012; Tsai et al., 2012). However, after brain injury or neurodegenerative disorders, astrocytes become activated and start to proliferate again (Robel et al., 2011; Sofroniew and Vinters, 2010). Previous studies reported that reactive astrocytes after injury may have certain stem cell properties, since isolating and culturing reactive astrocytes *in vitro* can generate neurospheres (Buffo et al., 2008; Lang et al., 2004; Shimada et al., 2012; Sirko et al., 2009, 2013). However, a number of studies have pointed out that the reactive astrocytes cannot generate neurons *in vivo*, although they may be able to generate glial cells (Buffo et al., 2008; Shimada et al., 2012). Therefore, injury itself can activate astrocytes to proliferate and even dedifferentiate to acquire certain properties of progenitor cells, but reactive astrocytes genetically remain within glial lineages. Our forced expression of NeuroD1 in reactive astrocytes may help them to overcome translineage barriers that cannot be surmounted by injury alone. We selected NeuroD1 for *in vivo* astrocyte-neuron conversion because NeuroD1 has been reported to be essential for adult neurogenesis (Gao et al., 2009; Kuwabara et al., 2009). NeuroD1 has also been shown to induce terminal neuronal differentiation (Boutin et al., 2010) and help reprogram human fibroblast cells into induced neurons when combined with Brn2, Ascl1, and Myt1l (Pang et al., 2011). Other transcription factors such as neurogenin-2 and Dlx2 have been shown to reprogram cultured mouse astrocytes into neurons (Heinrich et al., 2010). Therefore, it only takes a single neural transcription factor to change glial fate into neuronal fate both *in vitro* and *in vivo*.

Reactive astrocytes activated under different pathological conditions seem to have different proliferation rates. In particular, stab-injury- and ischemic-stroke-induced reactive astrocytes can be highly proliferative, whereas reactive astrocytes in APPPS1 or CK/p25 mice have lower rates of proliferation (Sirko et al., 2013). Sonic hedgehog (SHH) also plays a critical role in regulating the proliferative rate (Sirko et al., 2013). Although stab-injury-induced reactive astrocytes were found to express nestin, they were not reported to express Sox2 or Musashi. Therefore, it is possible that SHH alone can promote the pro-

liferation of reactive astrocytes, but it may not be sufficient to reverse reactive astrocytes into genuine neuroprogenitor cells. Interestingly, forced expression of Sox2 has been shown to dedifferentiate astrocytes into neuroblast cells (Niu et al., 2013). Sox2 is a critical marker for neural stem cells, and expression of Sox2 in fibroblast cells has been shown to induce neural stem cells (Ring et al., 2012). It is important to note that after Sox2-induced astrocyte-neuroblast conversion, addition of BDNF and noggin is required to further induce differentiation of neuroblasts into neurons (Niu et al., 2013). In contrast, our NeuroD1 expression reprograms reactive astrocytes directly into functional neurons *in vivo* without them going through a neuroprogenitor stage. Thus, our NeuroD1 reprogramming strategy can produce functional neurons rapidly after injury. Furthermore, our application of retroviral vectors targets proliferative glial cells that are typically activated by injury or diseases in the adult brain without affecting quiescent glial cells. Thus, NeuroD1-induced reprogramming may be particularly well-suited for therapeutic intervention and may not interfere significantly with normal astrocyte functions. Moreover, we demonstrate that NeuroD1-induced reactive glia-neuron conversion can occur in very old animals and even old animals with a model of AD. Therefore, NeuroD1-induced *in vivo* reactive astrocyte-neuron conversion could potentially be useful for regeneration of new neurons in the aging brain. In addition, the reprogramming of human astrocytes into functional neurons suggests that such a reactive glia-neuron conversion approach is potentially applicable to human patients.

### NeuroD1 Reprograms NG2 Cells into Glutamatergic and GABAergic Neurons

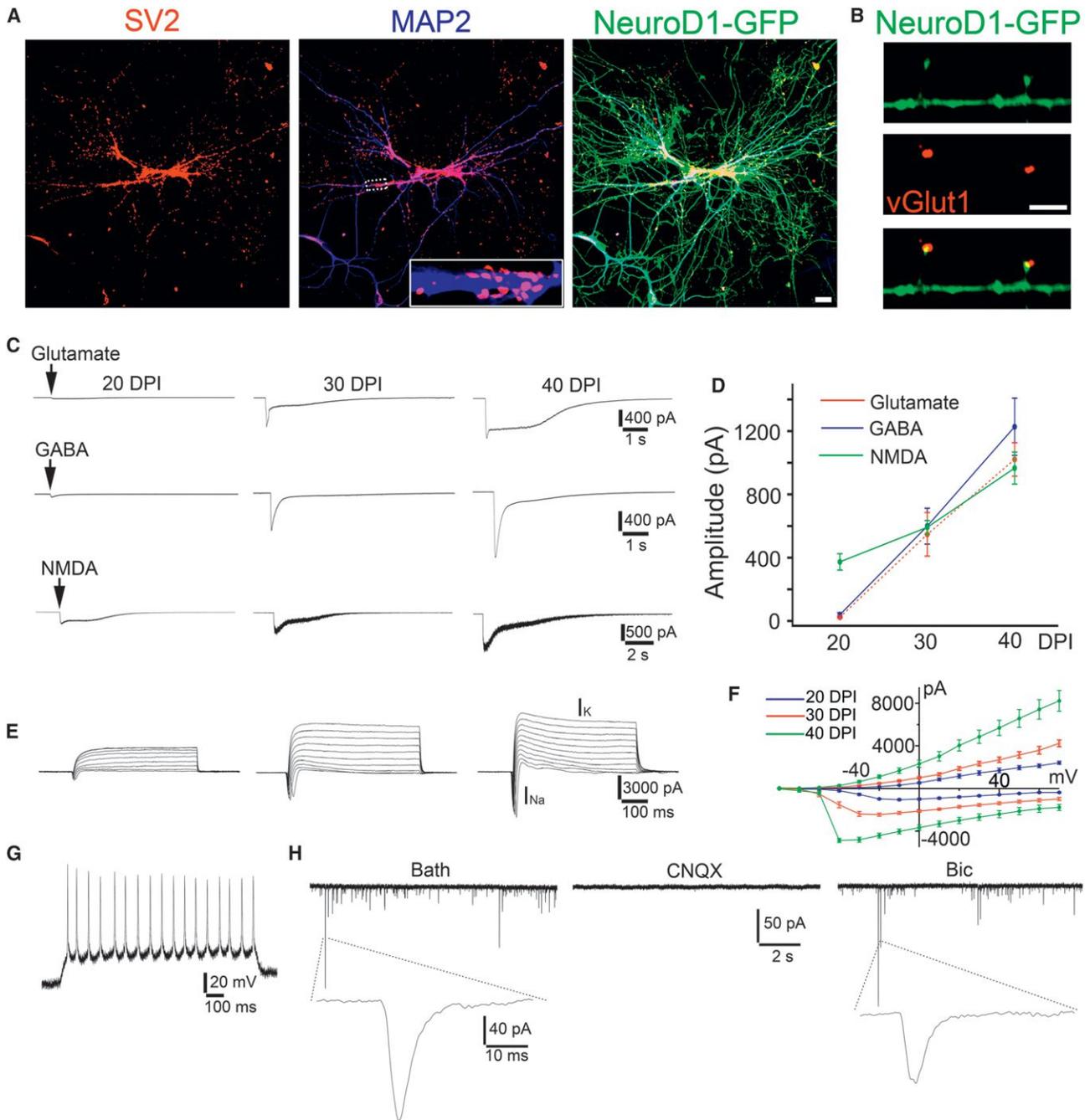
NG2 cells are the major proliferative glial cells in the adult brain under normal physiological conditions (Buffo et al., 2008; Kang et al., 2010). NG2 cells can receive synaptic inputs from neurons although the function of such neuron-glia synapses is not well understood (Bergles et al., 2010). In our stab injury model, we found that our CAG-GFP retrovirus-infected cells are mainly GFAP positive cells, and NG2 cells only account for about 20% of total infected cells. This bias might be due to the preferential infection of astrocytes by the retrovirus we used or the higher proliferation rate of reactive astrocytes than NG2 cells in our stab injury model. An unexpected finding in our study is that NeuroD1 not only reprograms astrocytes into functional neurons, but also reprograms NG2 cells into functional neurons. More interestingly, NeuroD1 reprograms astrocytes into glutamatergic neurons but reprograms NG2 cells into both glutamatergic and GABAergic neurons, suggesting that different glial cells may be associated with different neuronal fate in terms of lineage differentiation. Since glutamatergic and GABAergic neurons are the two major subtypes of neurons in the cortex, our finding that NeuroD1 can reprogram astrocytes and NG2 cells into glutamatergic and GABAergic neurons may have important functional implications. The simultaneous generation of both

(G) Phase contrast images showing NeuroD1-induced morphological change from astrocytes (left) to neurons (right, 45 DPI).

(H) Human astrocyte-converted neurons were immunopositive for VGluT1.

(I–K) Cortical layer neuronal markers revealed that human astrocyte-converted neurons were immunopositive for Tbr1 (I), Ctip2 (J), and Otx1 (K).

(L) Quantitative analysis of human astrocyte-converted neurons labeled by superficial (Cux1 and Lhx2) or deep layer (Ctip2 and Otx1) neuronal markers. Scale bars: 50  $\mu$ m for (A) and (E); 20  $\mu$ m for panels (C), (D), and (G–K); 40  $\mu$ m for panel (B). See also Figures S5 and S6.



**Figure 6. Functional Characterization of Human Astrocyte-Converted Neurons**

(A) Synaptic puncta (SV2, red) on the dendrites (MAP2, blue) of human astrocyte-converted neurons (green, 45 DPI) after NeuroD1 infection.  
 (B) High-power image showing vGlut1 puncta (red) colocalized with dendritic spines on NeuroD1-converted neurons. Scale bars: 20  $\mu$ m for panel (A); 10  $\mu$ m for panel (B).  
 (C–D) Representative traces (C) and quantitative analysis (D) of the receptor currents induced by bath application of glutamate (100  $\mu$ M), GABA (100  $\mu$ M), and NMDA (100  $\mu$ M).  
 (E and F) Representative traces of Na<sup>+</sup> and K<sup>+</sup> currents (E) and their I–V curve (F) recorded from NeuroD1-converted neurons.  
 (G) Representative trace of repetitive action potentials in NeuroD1-converted neurons (20 DPI).  
 (H) Representative traces of spontaneous synaptic events in NeuroD1-converted human neurons (40 DPI). Note that all synaptic events were blocked by CNQX (10  $\mu$ M) but not by BIC (20  $\mu$ M), suggesting that human astrocyte-converted neurons induced by NeuroD1 expression were glutamatergic neurons.

excitatory and inhibitory neurons by NeuroD1 alone could potentially make it possible to balance excitation and inhibition in the cortex after reprogramming.

### Conclusion

Our findings suggest that in situ reprogramming of reactive astrocytes and NG2 cells into functional neurons may offer a new approach to use internally reprogrammed neurons for brain repair. One outstanding question is whether it is possible to use a virus-free or small-molecule strategy to effect the reprogramming in vivo (Bayart and Cohen-Haguener, 2013; Chambers et al., 2012; Hou et al., 2013; Kaji et al., 2009; Li et al., 2013; Shi et al., 2008). An equally challenging question is whether the in vivo reprogramming can ultimately rescue behavioral deficits, such as cognitive impairment, in a diseased brain. Nevertheless, our in situ reprogramming of reactive glial cells into functional neurons suggests that it may be possible to replace neurons lost after nerve injury or diseases by direct reprogramming as a first step toward brain repair.

### EXPERIMENTAL PROCEDURES

#### Animals and In Vivo Assays

In vivo experiments were conducted on wild-type C57/BL6 and AD transgenic mice (5xFAD). AD transgenic mice were purchased from The Jackson Laboratory (B6SJL-Tg (APP<sup>SwFLon</sup>, PSEN1<sup>M146L</sup>\*L286V) 6799Vas/Mmjax) (Oakley et al., 2006) and mated with C57/BL6 mice. Mice were housed in a 12 hr light/dark cycle and supplied with enough food and water. Experimental protocols were approved by The Pennsylvania State University IACUC and in accordance with guidelines of the National Institutes of Health.

#### Stereotaxic Viral Injection

Surgeries were performed on 1- to 14-month-old WT and AD mice for virus injection. The mice were anesthetized by being injected with 20 ml/kg 2.5% Avertin (a mixture of 25 mg/ml of Tribromoethanol and 25  $\mu$ l/ml T-amylalcohol) into the peritoneum and then placed in a stereotaxic setup. Artificial eye ointment was applied to cover and protect the eye. The animals were operated upon with a midline scalp incision and a drilling hole on the skulls above somatosensory cortex. Each mouse received an injection (position: AP  $-1.25$  mm, ML 1.4 mm, DV  $-1.5$  mm) of virus with a 5  $\mu$ l syringe and a 34G needle. The injection volume and flow rate were controlled as 3  $\mu$ l at 0.2  $\mu$ l/min, and the needle was moved up during the injection at a speed of 0.1 mm/min. After injection, the needle was kept in place for at least 5 additional minutes and then slowly withdrawn. The needle injection itself was used as a stab injury model.

#### Mouse Cortical Astrocyte and NG2 Culture

For astrocyte culture, postnatal (P3–P5) mouse cortical tissue was dissociated and plated onto 25 cm<sup>2</sup> flasks (Wu et al., 2012). Cells were cultured for 5–6 days, and flasks were rigorously shaken daily to remove neurons and nonastrocytic cells. After reaching confluence, astrocytes were centrifuged for 5 min at 1,000 rpm, resuspended, and plated on poly-D-lysine (Sigma) -coated coverslips (12 mm). Astrocyte culture medium contained DMEM/F12 (GIBCO), 10% fetal bovine serum (GIBCO), penicillin/streptomycin (GIBCO), and 3.5 mM glucose (Sigma), supplemented with B27 (GIBCO), 10 ng/ml epidermal growth factor (EGF, Invitrogen), and 10 ng/ml fibroblast growth factor 2 (FGF2, Invitrogen). For mouse NG2 culture, the cortical tissue of postnatal mice (P3–P5) was dissociated and plated in 25 cm<sup>2</sup> flasks coated with poly-D-lysine (Sigma). The cells were maintained in DMEM/F12 (GIBCO) with 10% fetal bovine serum (GIBCO) for 9 days, with a medium change every 3 days. On the ninth day, the flasks were shaken rigorously and the supernatant was collected and centrifuged to enable the harvest of NG2 cells with a small number of neurons and microglia cells. After centrifuge, cells were resuspended and seeded on poly-D-lysine (Sigma) -coated coverslips (12 mm). The cells were cultured in serum-free DMEM medium (GIBCO) with N2 supplements (STEMCELL) and

10 ng/ml platelet-derived growth factor (PDGF, Invitrogen), 10 ng/ml EGF (Invitrogen), and 10 ng/ml FGF2 (Invitrogen) for 3 days. Cells were maintained at 37°C in humidified air with 5% CO<sub>2</sub>.

#### Human Cortical Astrocyte and Microglia Culture

Human cortical astrocytes (HA1800) were purchased from ScienCell (California). Cells were subcultured when they were over 90% confluent. For subculture, cells were trypsinized by TrypLE Select (Invitrogen), centrifuged for 5 min at 1,000 rpm, resuspended, and plated in a medium consisting of DMEM/F12 (GIBCO), 10% fetal bovine serum (GIBCO), penicillin/streptomycin (GIBCO), and 3.5 mM glucose (Sigma), supplemented with B27 (GIBCO), 10 ng/ml EGF (Invitrogen), and 10 ng/ml FGF2 (Invitrogen). The astrocytes were cultured on poly-D-lysine (Sigma) -coated coverslips (12 mm) at a density of 50,000 cells per coverslip in 24-well plates (BD Biosciences). Human primary microglial cells were obtained from Clonexpress, Inc. (MD). The cells were cultured in DMEM/F12 (GIBCO) supplemented with 5% FBS, 10 ng/ml of macrophage colony-stimulating factor (M-CSF, Invitrogen), 10 ng/ml EGF (Invitrogen), and 10 ng/ml FGF2 (Invitrogen). Cells were maintained at 37°C in humidified air with 5% CO<sub>2</sub>.

#### Retrovirus Production

The mouse NeuroD1 plasmid was constructed from our PCR product according to a template of the pAd NeuroD1-nGFP (Zhou et al., 2008) (Addgene) and inserted into a pCAG-GFP-IRES-GFP retroviral vector (Zhao et al., 2006) (gift of Dr. Fred Gage) to generate pCAG-NeuroD1-IRES-GFP. The human GFAP promoter gene was subcloned from hGFAP Promoter-Cre-MP-1 (Addgene) and replaced the CAG promoter to generate pGFAP-NeuroD1-IRES-GFP or pGFAP-GFP-IRES-GFP retroviral vector. The human NG2 promoter gene was subcloned from hNG2 Promoter-GLuc (GeneCopoeia) and replaced the CAG promoter to generate pNG2-NeuroD1-IRES-GFP or pNG2-GFP-IRES-GFP retroviral vector. Viral particles were packaged in gpg helperfree human embryonic kidney (HEK) cells to generate vesicular stomatitis virus glycoprotein (VSV-G)-pseudotyped retroviruses encoding neurogenic factors in CellMax hollow fiber cell culture system (Spectrum Laboratories). The titer of viral particles was about 10<sup>8</sup> particles/ $\mu$ l, determined after transduction of HEK cells.

#### Transdifferentiation of Glial Cells into Neurons

Twenty-four hours after infection of astrocytes, NG2 cells, or microglia with GFP or NeuroD1 retrovirus, the culture medium was completely replaced by a differentiation medium that included DMEM/F12 (GIBCO), 0.5% FBS (GIBCO), 3.5 mM glucose (Sigma), penicillin/streptomycin (GIBCO), and N2 supplement (GIBCO). Brain-derived neurotrophic factor (BDNF, 20 ng/ml, Invitrogen) was added to the cultures every 4 days during the differentiation to promote synaptic maturation (Song et al., 2002). Due to the morphological change from astrocytes or NG2 cells to neurons during conversion, we filled the empty space with additional human or mouse astrocytes to support the functional development of converted neurons.

#### Immunocytochemistry

For brain section staining, the mice were anesthetized with 2.5% Avertin and then sequentially perfused, first with saline solution (0.9% NaCl) to wash the blood off and then with 4% paraformaldehyde (PFA) to fix the brain. The brains were removed and postfixed in 4% PFA overnight at 4°C, and then cut at 45  $\mu$ m sections by a vibratome (Leica). Coronal brain sections were first pretreated in 0.3% Triton X-100 in phosphate-buffered saline (PBS, pH 7.4) for 1 hr, followed by incubation in 3% normal goat serum, 2% normal donkey serum, and 0.1% Triton X-100 in PBS for 1 hr.

For cell culture staining, the cultures were fixed in 4% PFA in PBS for 15 min at room temperature. Cells were first washed three times by PBS and then pretreated in 0.1% Triton X-100 in PBS for 30 min, followed by incubation in 3% normal goat serum, 2% normal donkey serum, and 0.1% Triton X-100 in PBS for 1 hr. Primary antibodies were incubated with either brain slices or cultures overnight at 4°C in 3% normal goat serum, 2% normal donkey serum, and 0.1% Triton X-100 in PBS. After additional washing in PBS, the samples were incubated with appropriate secondary antibodies conjugated to Alexa Fluor 488, Alexa Fluor 546, Alexa Fluor

647 (1:300, Molecular Probes), or Dylight (1:500, Jackson ImmunoResearch) for 1 hr at room temperature, followed by extensive washing in PBS. Coverslips were finally mounted onto a glass slide with an antifading mounting solution with DAPI (Invitrogen). Slides were first examined with an epifluorescent microscope (Nikon TE-2000-S) and further analyzed with a confocal microscope (Olympus FV1000). Z-stacks of digital images, which can either release single confocal images or collapse as one resulting picture, were acquired and analyzed using FV10-ASW 3.0 Viewer software (Olympus). For a detailed antibodies list, please see the Supplemental Information.

### Patch-Clamp Recordings in Cell Cultures

For glial cell-converted neurons, whole-cell recordings were performed using Multiclamp 700A patch-clamp amplifier (Molecular Devices, Palo Alto, CA) as described before (Deng et al., 2007), and the chamber was constantly perfused with a bath solution consisting of 128 mM NaCl, 30 mM glucose, 25 mM HEPES, 5 mM KCl, 2 mM CaCl<sub>2</sub>, and 1 mM MgCl<sub>2</sub>. The pH of bath solution was adjusted to 7.3 with NaOH, and osmolarity was at 315–325 mOsm/l. Patch pipettes were pulled from borosilicate glass (3–5 M $\Omega$ ) and filled with a pipette solution consisting of 135 mM KCl, 5 mM Na-phosphocreatine, 10 mM HEPES, 2 mM EGTA, 4 mM MgATP, and 0.5 mM Na<sub>2</sub>GTP (pH 7.3, adjusted with KOH). The series resistance was typically 10–30 M $\Omega$ . For voltage-clamp experiments, the membrane potential was typically held at –70 or –80 mV. Drugs were applied through a gravity-driven drug delivery system (VC-6, Warner Hamden, CT). NMDA currents were recorded in Mg<sup>2+</sup> free bath solution (128 mM NaCl, 30 mM D-glucose, 25 mM HEPES, 5 mM KCl, and 2 mM CaCl<sub>2</sub> [pH 7.3, adjusted with NaOH]) plus 10  $\mu$ M glycine, 0.5  $\mu$ M TTX, and 20  $\mu$ M BIC. Data were acquired using pClamp 9 software (Molecular Devices, Palo Alto, CA), sampled at 10 kHz, and filtered at 1 kHz. Na<sup>+</sup> and K<sup>+</sup> currents and action potentials were analyzed using pClamp 9 Clampfit software. Spontaneous synaptic events were analyzed using MiniAnalysis software (Synaptosoft, Decatur, GA). All experiments were conducted at room temperature.

### Brain Slice Recordings

Cortical slices were prepared typically ~1 month after virus injection and cut at 300  $\mu$ m thick coronal slices with a Leica vibratome in ice cold cutting solution (containing 75 mM sucrose, 87 mM NaCl, 2.5 mM KCl, 0.5 mM CaCl<sub>2</sub>, 7 mM MgCl<sub>2</sub>, 25 mM NaHCO<sub>3</sub>, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub> and 20 mM glucose). Slices were maintained in artificial cerebral spinal fluid (ACSF) containing 119 mM NaCl, 2.5 mM KCl, 1.25 mM NaH<sub>2</sub>PO<sub>4</sub>, 26 mM NaHCO<sub>3</sub>, 1.3 mM MgCl<sub>2</sub>, 2.5 mM CaCl<sub>2</sub>, and 10 mM glucose. Slices were incubated in ACSF and continuously bubbled with 95% O<sub>2</sub> and 5% CO<sub>2</sub>, first at 34°C for 30 min, and then at room temperature. Whole-cell recordings were performed using a pipette solution containing 135 mM K-Gluconate, 10 mM KCl, 5 mM Na-phosphocreatine, 10 mM HEPES, 2 mM EGTA, 4 mM MgATP, and 0.5 mM Na<sub>2</sub>GTP (pH 7.3, adjusted with KOH, 290 mOsm/l). Pipette resistance was 3–5 M $\Omega$ , and series resistance was typically 20–40 M $\Omega$ . The holding potential for voltage-clamp experiments was –70 mV. Data were collected using pClamp 9 software (Molecular Devices, Palo Alto, CA), sampled at 10 kHz, and filtered at 1 kHz, then analyzed with Clampfit and Synaptosoft softwares.

### SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes Supplemental Experimental Procedures and six figures and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2013.12.001>.

### AUTHOR CONTRIBUTIONS

Z.G. performed the major part of the experiments and data analysis. L.Z. contributed significantly on brain injury experiments. Z.W. performed cortical slice recordings. Y.C. made the GFAP promoter constructs and contributed to immunostaining. F.W. made viruses and contributed to immunostaining. G.C. conceived and supervised the entire project, analyzed the data, and wrote the manuscript.

### ACKNOWLEDGMENTS

We would like to thank Drs. Bernhard Luscher, Richard Ordway, and Timothy Jegla, as well as Chen lab members for vigorous discussion and helpful comments on the manuscript. This work was supported by a stem cell fund from Pennsylvania State University Eberly College of Science and grants from National Institutes of Health (MH083911 and MH092740) to G.C.

Received: June 19, 2013

Revised: October 12, 2013

Accepted: December 5, 2013

Published: December 19, 2013

### REFERENCES

- Aguzzi, A., Barres, B.A., and Bennett, M.L. (2013). Microglia: scapegoat, saboteur, or something else? *Science* 339, 156–161.
- Ambasudhan, R., Talantova, M., Coleman, R., Yuan, X., Zhu, S., Lipton, S.A., and Ding, S. (2011). Direct reprogramming of adult human fibroblasts to functional neurons under defined conditions. *Cell Stem Cell* 9, 113–118.
- Bayart, E., and Cohen-Haguener, O. (2013). Technological overview of iPS induction from human adult somatic cells. *Curr. Gene Ther.* 13, 73–92.
- Bergles, D.E., Jabs, R., and Steinhäuser, C. (2010). Neuron-glia synapses in the brain. *Brain Res. Brain Res. Rev.* 63, 130–137.
- Boutin, C., Hardt, O., de Chevigny, A., Coré, N., Goebbels, S., Seidenfaden, R., Bosio, A., and Cremer, H. (2010). NeuroD1 induces terminal neuronal differentiation in olfactory neurogenesis. *Proc. Natl. Acad. Sci. USA* 107, 1201–1206.
- Buffo, A., Rite, I., Tripathi, P., Lepier, A., Colak, D., Horn, A.P., Mori, T., and Götz, M. (2008). Origin and progeny of reactive gliosis: A source of multipotent cells in the injured brain. *Proc. Natl. Acad. Sci. USA* 105, 3581–3586.
- Caiazza, M., Dell'Anno, M.T., Dvoretzkova, E., Lazarevic, D., Taverna, S., Leo, D., Sotnikova, T.D., Menegon, A., Roncaglia, P., Colciago, G., et al. (2011). Direct generation of functional dopaminergic neurons from mouse and human fibroblasts. *Nature* 476, 224–227.
- Chambers, S.M., Qi, Y., Mica, Y., Lee, G., Zhang, X.J., Niu, L., Bilisland, J., Cao, L., Stevens, E., Whiting, P., et al. (2012). Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nat. Biotechnol.* 30, 715–720.
- Cho, J.H., and Tsai, M.J. (2004). The role of BETA2/NeuroD1 in the development of the nervous system. *Mol. Neurobiol.* 30, 35–47.
- Deng, L., Yao, J., Fang, C., Dong, N., Luscher, B., and Chen, G. (2007). Sequential postsynaptic maturation governs the temporal order of GABAergic and glutamatergic synaptogenesis in rat embryonic cultures. *J. Neurosci.* 27, 10860–10869.
- Gao, Z., Ure, K., Ales, J.L., Lagace, D.C., Nave, K.A., Goebbels, S., Eisch, A.J., and Hsieh, J. (2009). NeuroD1 is essential for the survival and maturation of adult-born neurons. *Nat. Neurosci.* 12, 1090–1092.
- Ge, W.P., Miyawaki, A., Gage, F.H., Jan, Y.N., and Jan, L.Y. (2012). Local generation of glia is a major astrocyte source in postnatal cortex. *Nature* 484, 376–380.
- Gwak, Y.S., Kang, J., Unabia, G.C., and Hulsebosch, C.E. (2012). Spatial and temporal activation of spinal glial cells: role of gliopathy in central neuropathic pain following spinal cord injury in rats. *Exp. Neurol.* 234, 362–372.
- Heinrich, C., Blum, R., Gascón, S., Masserdotti, G., Tripathi, P., Sánchez, R., Tiedt, S., Schroeder, T., Götz, M., and Berninger, B. (2010). Directing astroglia from the cerebral cortex into subtype specific functional neurons. *PLoS Biol.* 8, e1000373.
- Hines, D.J., Hines, R.M., Mulligan, S.J., and Macvicar, B.A. (2009). Microglia processes block the spread of damage in the brain and require functional chloride channels. *Glia* 57, 1610–1618.
- Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., et al. (2013). Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 341, 651–654.

- Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 458, 771–775.
- Kang, S.H., Fukaya, M., Yang, J.K., Rothstein, J.D., and Bergles, D.E. (2010). NG2+ CNS glial progenitors remain committed to the oligodendrocyte lineage in postnatal life and following neurodegeneration. *Neuron* 68, 668–681.
- Kang, S.H., Li, Y., Fukaya, M., Lorenzini, I., Cleveland, D.W., Ostrow, L.W., Rothstein, J.D., and Bergles, D.E. (2013). Degeneration and impaired regeneration of gray matter oligodendrocytes in amyotrophic lateral sclerosis. *Nat. Neurosci.* 16, 571–579.
- Kim, J., Su, S.C., Wang, H., Cheng, A.W., Cassady, J.P., Lodato, M.A., Lengner, C.J., Chung, C.Y., Dawlaty, M.M., Tsai, L.H., and Jaenisch, R. (2011). Functional integration of dopaminergic neurons directly converted from mouse fibroblasts. *Cell Stem Cell* 9, 413–419.
- Kuwabara, T., Hsieh, J., Muotri, A., Yeo, G., Warashina, M., Lie, D.C., Moore, L., Nakashima, K., Asashima, M., and Gage, F.H. (2009). Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. *Nat. Neurosci.* 12, 1097–1105.
- 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.
- Lang, B., Liu, H.L., Liu, R., Feng, G.D., Jiao, X.Y., and Ju, G. (2004). Astrocytes in injured adult rat spinal cord may acquire the potential of neural stem cells. *Neuroscience* 128, 775–783.
- Li, W., Li, K., Wei, W., and Ding, S. (2013). Chemical approaches to stem cell biology and therapeutics. *Cell Stem Cell* 13, 270–283.
- Liu, X., Li, F., Stubblefield, E.A., Blanchard, B., Richards, T.L., Larson, G.A., He, Y., Huang, Q., Tan, A.C., Zhang, D., et al. (2012). Direct reprogramming of human fibroblasts into dopaminergic neuron-like cells. *Cell Res.* 22, 321–332.
- Liu, M.L., Zang, T., Zou, Y., Chang, J.C., Gibson, J.R., Huber, K.M., and Zhang, C.L. (2013). Small molecules enable neurogenin 2 to efficiently convert human fibroblasts into cholinergic neurons. *Nat. Commun.* 4, 2183.
- Meng, F., Chen, S., Miao, Q., Zhou, K., Lao, Q., Zhang, X., Guo, W., and Jiao, J. (2012). Induction of fibroblasts to neurons through adenoviral gene delivery. *Cell Res.* 22, 436–440.
- Miller, F.D., and Gauthier, A.S. (2007). Timing is everything: making neurons versus glia in the developing cortex. *Neuron* 54, 357–369.
- Najm, F.J., Lager, A.M., Zaremba, A., Wyatt, K., Capriariello, A.V., Factor, D.C., Karl, R.T., Maeda, T., Miller, R.H., and Tesar, P.J. (2013). Transcription factor-mediated reprogramming of fibroblasts to expandable, myelinogenic oligodendrocyte progenitor cells. *Nat. Biotechnol.* 31, 426–433.
- Niu, W., Zang, T., Zou, Y., Fang, S., Smith, D.K., Bachoo, R., and Zhang, C.L. (2013). In vivo reprogramming of astrocytes to neuroblasts in the adult brain. *Nat. Cell Biol.* 15, 1164–1175.
- Oakley, H., Cole, S.L., Logan, S., Maus, E., Shao, P., Craft, J., Guillozet-Bongaarts, A., Ohno, M., Disterhoft, J., Van Eldik, L., et al. (2006). Intraneuronal beta-amyloid aggregates, neurodegeneration, and neuron loss in transgenic mice with five familial Alzheimer's disease mutations: potential factors in amyloid plaque formation. *J. Neurosci.* 26, 10129–10140.
- Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Südhof, T.C., and Wernig, M. (2011). Induction of human neuronal cells by defined transcription factors. *Nature* 476, 220–223.
- Pekny, M., and Nilsson, M. (2005). Astrocyte activation and reactive gliosis. *Glia* 50, 427–434.
- Pfisterer, U., Kirkeby, A., Torper, O., Wood, J., Nelander, J., Dufour, A., Björklund, A., Lindvall, O., Jakobsson, J., and Parmar, M. (2011). Direct conversion of human fibroblasts to dopaminergic neurons. *Proc. Natl. Acad. Sci. USA* 108, 10343–10348.
- Qiang, L., Fujita, R., Yamashita, T., Angulo, S., Rhinn, H., Rhee, D., Doege, C., Chau, L., Aubry, L., Vanti, W.B., et al. (2011). Directed conversion of Alzheimer's disease patient skin fibroblasts into functional neurons. *Cell* 146, 359–371.
- Ring, K.L., Tong, L.M., Balestra, M.E., Javier, R., Andrews-Zwilling, Y., Li, G., Walker, D., Zhang, W.R., Kreitzer, A.C., and Huang, Y. (2012). Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. *Cell Stem Cell* 11, 100–109.
- Robel, S., Berninger, B., and Götz, M. (2011). The stem cell potential of glia: lessons from reactive gliosis. *Nat. Rev. Neurosci.* 12, 88–104.
- Rodriguez, J.J., Olabarria, M., Chvatal, A., and Verkhratsky, A. (2009). Astroglia in dementia and Alzheimer's disease. *Cell Death Differ.* 16, 378–385.
- Shi, Y., Despons, C., Do, J.T., Hahm, H.S., Schöler, H.R., and Ding, S. (2008). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3, 568–574.
- Shimada, I.S., LeComte, M.D., Granger, J.C., Quinlan, N.J., and Spees, J.L. (2012). Self-renewal and differentiation of reactive astrocyte-derived neural stem/progenitor cells isolated from the cortical peri-infarct area after stroke. *J. Neurosci.* 32, 7926–7940.
- Sirko, S., Neitz, A., Mittmann, T., Horvat-Brocker, A., von Holst, A., Eysel, U.T., and Faissner, A. (2009). Focal laser-lesions activate an endogenous population of neural stem/progenitor cells in the adult visual cortex. *Brain* 132, 2252–2264.
- Sirko, S., Behrendt, G., Johansson, P.A., Tripathi, P., Costa, M., Bek, S., Heinrich, C., Tiedt, S., Colak, D., Dichgans, M., et al. (2013). Reactive glia in the injured brain acquire stem cell properties in response to sonic hedgehog. [corrected]. *Cell Stem Cell* 12, 426–439.
- Sofroniew, M.V., and Vinters, H.V. (2010). Astrocytes: biology and pathology. *Acta Neuropathol.* 119, 7–35.
- Son, E.Y., Ichida, J.K., Wainger, B.J., Toma, J.S., Rafuse, V.F., Woolf, C.J., and Eggan, K. (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. *Cell Stem Cell* 9, 205–218.
- Song, H., Stevens, C.F., and Gage, F.H. (2002). Astroglia induce neurogenesis from adult neural stem cells. *Nature* 417, 39–44.
- Steele, M.L., and Robinson, S.R. (2010). Reactive astrocytes give neurons less support: implications for Alzheimer's disease. *Neurobiol. Aging* 33, 423, e421–413.
- 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.
- 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.
- Torper, O., Pfisterer, U., Wolf, D.A., Pereira, M., Lau, S., Jakobsson, J., Björklund, A., Grealish, S., and Parmar, M. (2013). Generation of induced neurons via direct conversion in vivo. *Proc. Natl. Acad. Sci. USA* 110, 7038–7043.
- Tsai, H.H., Li, H., Fuentealba, L.C., Molofsky, A.V., Taveira-Marques, R., Zhuang, H., Tenney, A., Murnen, A.T., Fancy, S.P., Merkle, F., et al. (2012). Regional astrocyte allocation regulates CNS synaptogenesis and repair. *Science* 337, 358–362.
- Verkhratsky, A., Olabarria, M., Noristani, H.N., Yeh, C.Y., and Rodriguez, J.J. (2010). Astrocytes in Alzheimer's disease. *Neurotherapeutics* 7, 399–412.
- Verkhratsky, A., Sofroniew, M.V., Messing, A., deLanerolle, N.C., Rempe, D., Rodriguez, J.J., and Nedergaard, M. (2012). Neurological diseases as primary gliopathies: a reassessment of neurocentrism. *ASN Neuro* 4, 4.
- 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.
- Wu, X., Wu, Z., Ning, G., Guo, Y., Ali, R., Macdonald, R.L., De Blas, A.L., Luscher, B., and Chen, G. (2012).  $\gamma$ -Aminobutyric acid type A (GABAA) receptor  $\alpha$  subunits play a direct role in synaptic versus extrasynaptic targeting. *J. Biol. Chem.* 287, 27417–27430.

- Yang, N., Zuchero, J.B., Ahlenius, H., Marro, S., Ng, Y.H., Vierbuchen, T., Hawkins, J.S., Geissler, R., Barres, B.A., and Wernig, M. (2013). Generation of oligodendroglial cells by direct lineage conversion. *Nat. Biotechnol.* 31, 434–439.
- Yoo, A.S., Sun, A.X., Li, L., Shcheglovitov, A., Portmann, T., Li, Y., Lee-Messer, C., Dolmetsch, R.E., Tsien, R.W., and Crabtree, G.R. (2011). MicroRNA-mediated conversion of human fibroblasts to neurons. *Nature* 476, 228–231.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920.
- Zhao, C., Teng, E.M., Summers, R.G., Jr., Ming, G.L., and Gage, F.H. (2006). Distinct morphological stages of dentate granule neuron maturation in the adult mouse hippocampus. *J. Neurosci.* 26, 3–11.
- Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J., and Melton, D.A. (2008). In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* 455, 627–632.

# The Developmental Potential of iPSCs Is Greatly Influenced by Reprogramming Factor Selection

Yosef Buganim,<sup>2,8,\*</sup> Styliani Markoulaki,<sup>1,8</sup> Niek van Wietmarschen,<sup>3</sup> Heather Hoke,<sup>1,4</sup> Tao Wu,<sup>5</sup> Kibibi Ganz,<sup>1</sup> Batool Akhtar-Zaidi,<sup>1</sup> Yupeng He,<sup>6</sup> Brian J. Abraham,<sup>1</sup> David Porubsky,<sup>3</sup> Elisabeth Kulenkampff,<sup>1</sup> Dina A. Faddah,<sup>1,4</sup> Linyu Shi,<sup>1</sup> Qing Gao,<sup>1</sup> Sovan Sarkar,<sup>1</sup> Malkiel Cohen,<sup>1</sup> Johanna Goldmann,<sup>1</sup> Joseph R. Nery,<sup>6</sup> Matthew D. Schultz,<sup>6</sup> Joseph R. Ecker,<sup>6</sup> Andrew Xiao,<sup>5</sup> Richard A. Young,<sup>1,4,7</sup> Peter M. Lansdorp,<sup>3,7</sup> and Rudolf Jaenisch<sup>1,4,7,\*</sup>

<sup>1</sup>Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

<sup>2</sup>Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel-Canada, Hebrew University-Hadassah Medical School, Jerusalem 91120, Israel

<sup>3</sup>European Research Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, Antonius Deusinglaan 1, AV Groningen 9713, the Netherlands

<sup>4</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>5</sup>Yale Stem Cell Center, Yale University, New Haven, CT 06520, USA

<sup>6</sup>Genomic Analysis Laboratory, Howard Hughes Medical Institute, Salk Institute for Biological Studies, La Jolla, CA 92037, USA

<sup>7</sup>Skolkovo Institute of Science and Technology (Skoltech), Novaya strazha 100, Skolkovo Moscow Region 143025, Russia

<sup>8</sup>Co-first author

\*Correspondence: yossibug@ekmd.huji.ac.il (Y.B.), jaenisch@wi.mit.edu (R.J.)

<http://dx.doi.org/10.1016/j.stem.2014.07.003>

## SUMMARY

Induced pluripotent stem cells (iPSCs) are commonly generated by transduction of Oct4, Sox2, Klf4, and Myc (OSKM) into cells. Although iPSCs are pluripotent, they frequently exhibit high variation in terms of quality, as measured in mice by chimera contribution and tetraploid complementation. Reliably high-quality iPSCs will be needed for future therapeutic applications. Here, we show that one major determinant of iPSC quality is the combination of reprogramming factors used. Based on tetraploid complementation, we found that ectopic expression of Sall4, Nanog, Esrrb, and Lin28 (SNEL) in mouse embryonic fibroblasts (MEFs) generated high-quality iPSCs more efficiently than other combinations of factors including OSKM. Although differentially methylated regions, transcript number of master regulators, establishment of specific superenhancers, and global aneuploidy were comparable between high- and low-quality lines, aberrant gene expression, trisomy of chromosome 8, and abnormal H2A.X deposition were distinguishing features that could potentially also be applicable to human.

## INTRODUCTION

Recent reports indicate that the majority of OSKM-derived iPSCs may have reduced differentiation potential as compared to embryonic stem cells (ESCs) derived by somatic cell nuclear transfer (SCNT), which are equivalent in their developmental potential to ESCs derived from the fertilized egg (Boland et al.,

2009; Brambrink et al., 2006; Jiang et al., 2011, 2013; Kang et al., 2009; Kim et al., 2010; Pera, 2011; Polo et al., 2010; Zhao et al., 2009). In addition, it has been suggested that OSKM-derived iPSCs exhibit genetic and epigenetic aberrations throughout the genome that are distinct from ESCs (Bar-Nur et al., 2011; Chin et al., 2009; Doi et al., 2009; Gore et al., 2011; Hussein et al., 2011; Kim et al., 2010, 2011; Laurent et al., 2011; Lister et al., 2011; Mayshar et al., 2010; Ohi et al., 2011; Phanstiel et al., 2011; Polo et al., 2010). These data are consistent with the prevailing current reprogramming method affecting the quality of the resulting pluripotent cells. Several parameters have been shown to affect the quality of iPSCs, such as factor stoichiometry (Carey et al., 2011), culture condition, and supplements used to derive the cells (Chen et al., 2011). For example, by comparing two genetically defined transgenic systems to identify parameters affecting reprogramming, it has been shown that high levels of Oct4 and Klf4, together with low levels of Sox2 and Myc, are favorable with respect to the quality of the iPSCs even though a much lower reprogramming efficiency was observed when compared to high levels of Sox2 and Myc and low levels of Oct4 and Klf4 (Carey et al., 2011). Also, derivation of iPSCs in the absence of serum but in the presence of vitamin C improved the quality of the cells and generated tetraploid complementation-competent iPSCs even when a sub-optimal factor stoichiometry was used to induce pluripotency (Esteban and Pei, 2012; Stadtfeld et al., 2012). In summary, the available data suggest that factor stoichiometry, as well as specific culture conditions, affect the quality of iPSCs. Here, we show that the quality of iPSCs is dramatically affected by the specific choice of reprogramming factors. Reprogramming by Sall4, Nanog, Esrrb, and Lin28 (SNEL) generated a very low number of iPSC colonies, the majority of which were of high quality as defined by their capacity to produce healthy “all-iPSC” mice, as determined by 4n complementation, the most stringent test for pluripotency. In stark contrast, OSKM produced a large number

of iPSC colonies, the majority of which, using the same assay, exhibited low developmental potential. Removing Myc from the cocktail (OSK) yielded a higher number of high-quality iPSCs, indicating that the presence of Myc in the reprogramming factors combination has a negative effect on iPSC quality. Surprisingly, a combination of Oct4, Sox2, Sall4, Nanog, and Esrrb (OSSNE), although lacking potent oncogenes like Myc and Lin28, yielded the highest number of poor quality iPSCs, suggesting that the interplay between the reprogramming factors plays a critical role in the reprogramming process as well. To shed light on the elements that dictate successful reprogramming events, we performed a large number of genomic and epigenomic analyses. While whole genome transcriptional profile, methylome analysis, establishment of superenhancers, or single-cell analysis of key master regulator transcript number and global aneuploidy did not distinguish between poor- and high-quality iPSCs, aberrant expression of 1,765 genes, trisomy of chromosome 8, and abnormal H2A.X deposition were frequently observed in poor-quality iPSCs that were derived by OSKM or OSK. Our results demonstrate that the selection of the reprogramming factor combination is an important determinant for retaining genomic integrity, appropriate transcriptional resetting, and functional pluripotency of iPSCs.

## RESULTS

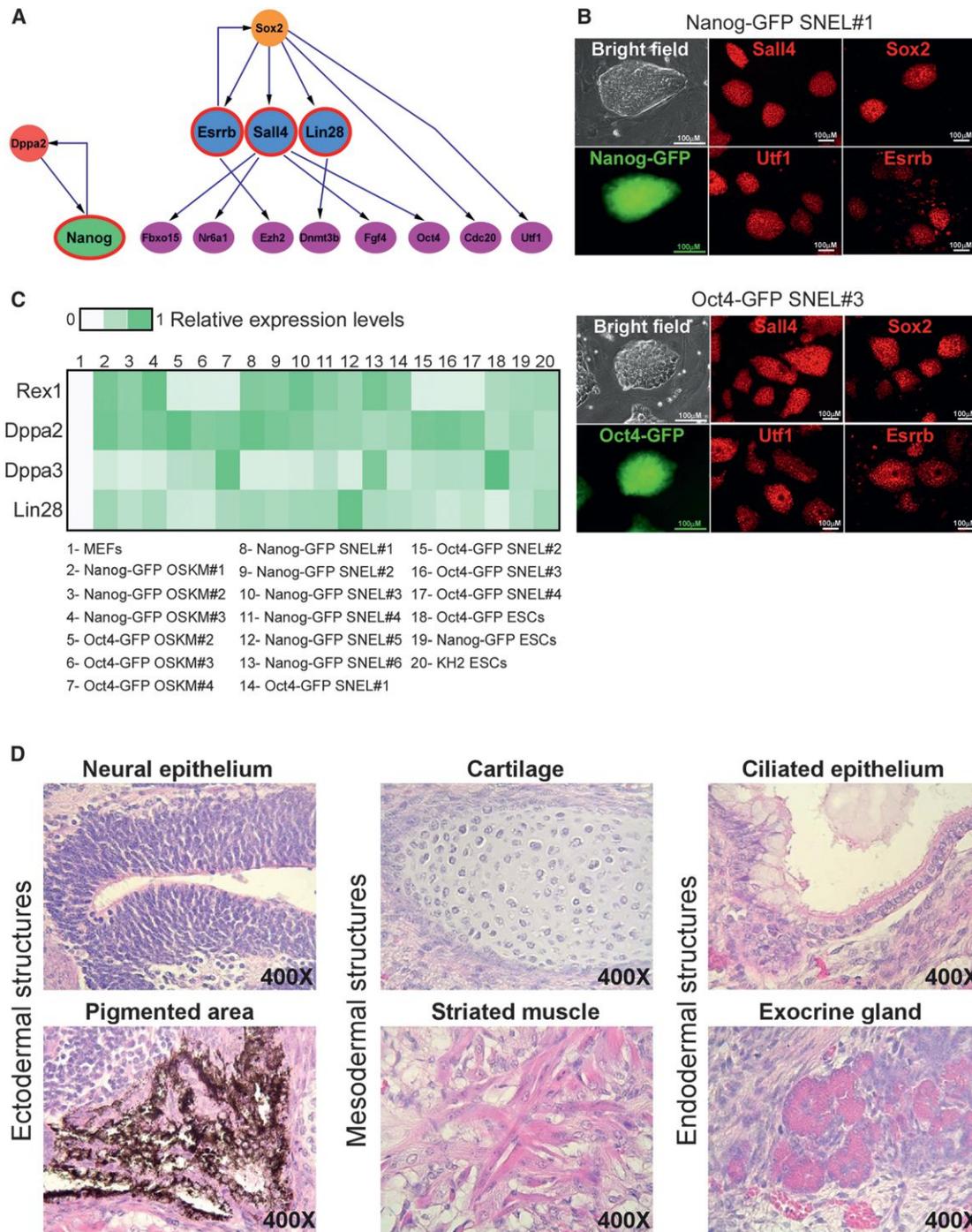
### Ectopic Expression of Sall4, Nanog, Esrrb, and Lin28 Activates the Endogenous Pluripotency Circuitry

Recently, using two complementary single-cell technologies, we have demonstrated that the reprogramming process involves a late hierarchical/deterministic phase that starts with the activation of the Sox2 locus and continues with a series of gene activation events that lead to a stable and transgene-independent pluripotency state (Figure 1A) (Buganim et al., 2012, 2013; Klemm et al., 2014; Pan and Pei, 2012). We reasoned that a combination of key factors derived from this later phase would reprogram cells in a more controlled manner and potentially would then uniformly yield iPSCs of high quality. We focused mainly on one specific combination of factors, Sall4, Nanog, Esrrb, and Lin28 (SNEL), because we wished to avoid ectopic expression of the key master regulators Oct4 and Sox2. We hypothesized that exogenous expression of strong key master regulators such as Oct4 and Sox2, without their endogenous regulators present in ESCs or during nuclear transfer-mediated reprogramming in the oocyte, might yield aberrant activation of various loci in the somatic cell genome that interfere with normal reprogramming events as is apparent in partially reprogrammed cells (Buganim et al., 2012, 2013). Sall4 and Esrrb were chosen based on a Bayesian network analysis prediction that they could activate the endogenous Oct4 and Sox2 genes, respectively. We selected Lin28 because it has been shown to act as a global mRNA regulator (Cho et al., 2012) and to activate the de novo DNA methyltransferase, *Dnmt3b*, and Nanog that was predicted to have a separate role during reprogramming (Figure 1A). Nanog-GFP or Oct4-GFP mouse embryonic fibroblasts (MEFs) were introduced with doxycycline (dox)-inducible lentiviruses encoding for the SNEL reprogramming factors and cultured until the formation of iPSC colonies. The efficiency of the reprogramming process was very low, producing one to five colonies per

$1 \times 10^5$  plated cells with a latency that ranged between 14–60 days. In total, we isolated ten SNEL-iPSC colonies (six from Nanog-GFP and four from Oct4-GFP MEFs). The resulting iPSC colonies expressed a bright GFP signal from both the Oct4 and the Nanog locus and showed upregulation of key pluripotency markers such as Sox2, endogenous Sall4, Utf1, and endogenous Esrrb, as assessed by immunostaining (Figure 1B). A comparable mRNA level to OSKM-iPSCs and several ESC lines of *Dppa2*, *Dppa3*, endogenous *Lin28*, and *Rex1* was noted as well by quantitative real-time PCR (qRT-PCR) (Figure 1C). When injected into NOD/SCID mice, the cells formed well-differentiated teratomas with structures from all three germ layers (representative images can be seen in Figure 1D).

### SNEL-iPSCs Contribute to High-Grade Chimeras and Frequently Produce “All-iPSC” Mice

The potential of SNEL-iPSCs to generate chimeras was tested by injecting cells from all ten clones into host blastocysts that were subsequently transferred into pseudopregnant recipient females. All lines gave rise to chimeras, with 8/10 (80%) generating high-grade chimeras (50%–95%) as assessed by coat color (Figures S1A and S1B available online). Germline transmission was noted in all five tested lines (Figures S1A and S1C). Chimeric mice from one of the iPSC clones (Oct4-GFP SNEL#2) suffered from an eye problem and one adult mouse developed a tumor. However, these isolated events might be explained by leaky expression of Esrrb and Lin28, which have been linked to similar phenotypes (Audo et al., 2012; Viswanathan et al., 2009; West et al., 2009). All other chimeras generated from independent clones grew to old age without any obvious evidence of tumorigenicity (Figure S1C). To stringently compare the developmental potential of SNEL and OSKM-derived iPSCs, we performed 4n complementation assay. Utilizing identical infection and culture conditions as used for the derivation of the SNEL-iPSCs, ten iPSC lines were derived by infection of MEFs with OSKM lentiviruses all of which expressed high levels of GFP and pluripotency markers (6/10 of the colonies are presented in Figure 1C). Cells from the ten SNEL-iPSC and ten OSKM-iPSC lines were injected into 4n blastocysts and transferred into pseudopregnant recipient females. The SNEL-iPSCs produced approximately five times as many live 4n pups with the majority surviving postnatally ( $p = 3.46 \times 10^{-12}$  by  $\chi^2$  test) (Figures 2A and 2B; Table S1) as compared to OSKM-iPSCs. Out of a total of 1,495 OSKM-iPSC-injected blastocysts, only 21 (1.4%) were delivered, 11 (0.7%) of which sustained normal breathing and were foster nursed. In contrast, out of 2,138 blastocysts injected with SNEL iPSCs, 149 (7%) survived to birth, 109 (5%) of which were breathing normally and were fostered nursed (Table S1). In total, ~40% of the OSKM-iPSC lines gave rise to live pups, compared to 80% of the SNEL-iPSC lines. In general, the adult “all-SNEL-iPSC” mice were healthy and fertile (Figures 2B and S2C), although some mice exhibited some phenotypes that are purely related to leaky expression of Lin28, such as long tail and ears and flattened nose as described in (Zhu et al., 2010, 2011), and several individuals died prematurely after 1 year. To exploit the maximum potential of the cells and determine whether the developmental differences between these two types of iPSCs would be further exacerbated, we cultured the 20 iPSC lines in 2i medium (LIF, GSK3 $\beta$ , and Mek 1/2 inhibitors



**Figure 1. Characterization of SNEL-iPSC Lines**

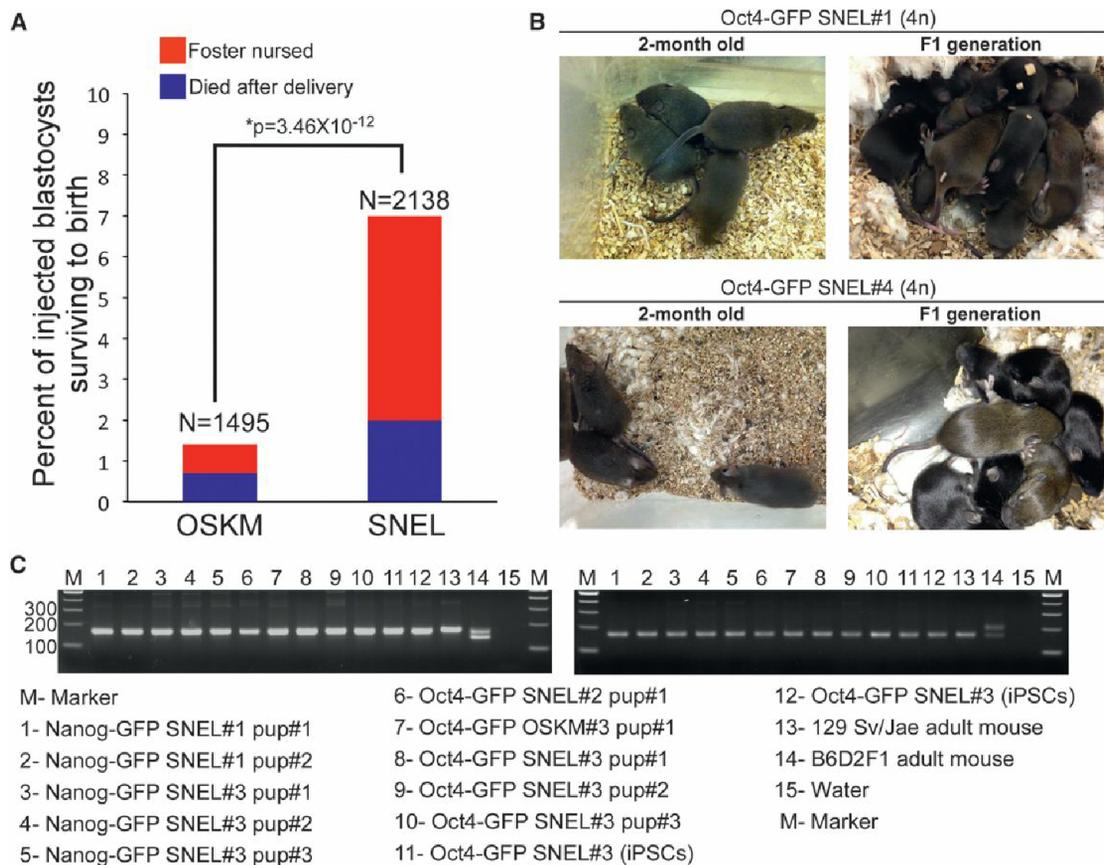
(A) Schematic presentation of Bayesian network demonstrates the hierarchy of a subset of pluripotent genes that leads to a stable and transgene independent pluripotency state (Buganim et al., 2012). Sall4, Nanog, Esrrb, and Nanog (SNEL) are marked by a red circle.

(B) Representative images of two stable dox-independent, GFP-positive colonies (Nanog-GFP SNEL#1 and Oct4-GFP SNEL#3) and immunostaining for Sall4, Sox2, Utf1, and Esrrb.

(C) Heatmap demonstrating the relative expression levels of *Dppa3*, *Dppa2*, *Zfp42* (*Rex1*), and *Lin28* normalized to the *Hprt* housekeeping control gene in the indicated samples.

(D) Hematoxylin and eosin staining of teratoma sections generated from Oct4-GFP SNEL#1 showing structures from all three layers.

See also Figure S1.



**Figure 2. SNEL-iPSCs Produce “All-iPSC” Mice with High Success Rates as Compared to OSKM**

(A) Percent of injected blastocysts surviving to birth are plotted for OSKM and SNEL lines, with the number of blastocysts noted on the x axis. Blue represents the number of pups that merely survived delivery, red the number of pups additionally foster-nursed. Percentages were compared by  $\chi^2$  test to compute significance. (B) Representative images of 4n adult mice produced from Oct4-GFP SNEL#1 and Oct4-GFP SNEL#4 lines and their F1 generation. (C) Confirmation of origin of “all-iPSC” mice by PCR for strain-specific polymorphisms. Two different simple sequence polymorphism (SSLP) markers were tested using genomic DNA isolated from tissues of “all-iPSC” mice. Genomic DNA from the parental iPSCs (donor cells), a 129 Sv/Jae mouse (donor strain), and a B6D2F1 mouse (host blastocyst strain) served as controls. See also Figure S2.

containing medium) for two passages and then injected each line into 60 4n blastocysts. The percentage of live-born pups in the SNEL combination was significantly higher, reaching 23%–25% in some SNEL-iPSC lines (Table 1 and Figures S2A and S2B). From a total of 600 OSKM-iPSC-injected blastocysts only 13 (2.1%) were delivered, eight (1.7%) of which sustained normal breathing and were foster nursed. In contrast, out of 600 blastocysts injected with SNEL-iPSCs, 64 (10.7%) survived to birth, 51 (8.5%) of which were breathing normally and were fostered nursed (Table 1). Germline transmission was noted in all examined lines. Simple sequence length polymorphism (SSLP) analysis for ten randomly selected 4n embryos (PCR-based assay for two loci) confirmed that the embryos were solely derived from the injected iPSCs (Figure 2C). Our data suggest that reprogramming with SNEL, in contrast to reprogramming using OSKM under the same conditions, produces high-quality iPSCs at high rates as assessed by the most stringent test of 4n complementation.

Myc is a potent oncogene that affects genomic integrity when deregulated (Barlow et al., 2013) and enhances the transcription

at all active promoters when overexpressed (Lin et al., 2012; Nie et al., 2012). To determine whether overexpression of Myc is the cause for the high number of poor quality OSKM-iPSC colonies, we derived ten OSK-iPSC colonies under identical conditions and assessed their quality by 4n complementation. Indeed, a larger number of colonies (8/10 in OSK-iPSC colonies compared to 4/10 in OSKM-iPSC colonies) passed the 4n complementation test. However, a significantly lower number of pups were delivered per line (30 live pups, 5%) in colonies derived from OSK compared to colonies derived from SNEL (64 live pups, 10.7%). These results indicate that Myc is partially responsible for the high number of poor quality colonies in the OSKM combination but suggest that other parameters influence the quality, as well. We then sought to determine whether potent oncogenes are the main reason for the generation of poor quality iPSCs. To test this hypothesis, we isolated 6 colonies, generated by ectopic expression of Oct4, Sox2, Sall4, Nanog, and Esrrb (OSSNE) but without the potent oncogenes Myc and Lin28. Surprisingly, OSSNE combination yielded the highest number of poor quality iPSCs (five live pups, 1.3%), suggesting that the

interplay between the reprogramming factors is a crucial element during the reprogramming process.

### Comparative Transcriptome Profiling of Poor- and High-Quality iPSCs Reveals a Distinct Signature of 1,765 Genes that Robustly Distinguish Lines by 4n Proficiency

To reveal a gene expression signature associated with developmental competence, we selected the following groups of iPSC lines for microarray analysis: (1) “poor quality” iPSCs: this group included the three OSKM-iPSC lines Nanog-GFP OSKM#2, Oct4-GFP OSKM#2, and KH2 OSKM (Stadtfield et al., 2010), that either did not produce fully developed pups or produced very low number of pups; (2) “good quality” iPSCs: this group included BC\_2 OSKM (Carey et al., 2011) and Nanog-GFP SNEL#3, both of which gave rise to live, normal pups that survived only a few hours; (3) “high quality” iPSCs consisting of Nanog-GFP SNEL#2 and Oct4-GFP SNEL#1, both of which generated live pups that survived postnatally (representative pups from each iPSC group are shown in Figure S3A). We used Nanog-GFP, Oct4-GFP, and KH2 (Beard et al., 2006) ESCs as controls.

Whole genome transcriptional analysis did not distinguish between the groups as assessed by hierarchical clustering and principle component analysis (PCA), consistent with their common identities as pluripotent cells (Figures 3A and S3B). In contrast, unbiased hierarchical clustering and PCA analysis of differentially expressed genes between all groups (F test,  $p < 0.01$ , Table S2) revealed a list of 1,765 genes that separated perfectly the different groups and clustered the “poor quality” group away from the other three groups (Figures 3B and 3C). qRT-PCR for two representative differentially expressed genes, *Col6a1* and *Thsb1*, from the 1,765 gene signature demonstrated a trend of expression that is highly correlated with 4n competence (Figure 3D).

Gene ontologies and pathways (GeneDecks) (Stelzer et al., 2009) for the 1,765 differentially expressed genes revealed enrichment not only for categories associated with the control of cellular growth and division, but also for more refined and specific developmental pathways and phenotypes: respiratory, immune, musculature, and aortic integrity phenotypes; hypoxia, myocardial infarction, and pulmonary disease; abnormal limb/digit/tail morphology; genes involved in extracellular matrix composition and TGF $\beta$  signaling; and defective embryogenesis (the  $p$  value and number of genes of representatives categories are presented in Figure 3E). Standard motif enrichment analysis on the 1,765 gene promoters revealed a strong enrichment for transcription factors vital for early embryonic development and ESC self-renewal (Hanna et al., 2002; Liu and Labosky, 2008; Shah et al., 2012; Weinhold et al., 2000) such as Foxd3, HMG-I/Y, and Srf (Figure S3C), lending support to the role of the 1,765 genes in development and ESC maintenance.

### Differentially Methylated Regions and the Establishment of ESC-Specific Superenhancers Cannot Distinguish between Poor- and High-Quality iPSCs

To assess whether this gene expression pattern may be associated with epigenetic alterations, we profiled the methylomes of these samples by whole genome bisulfite sequencing. Although over 2,500 differentially methylated regions (DMRs) were identi-

fied, these were largely specific to individual iPSC lines (i.e., Nanog-GFP versus Oct4-GFP lines, Figure 3F). Thus, the exclusively intronic and intergenic genomic distribution of DMRs precluded accurate assessment of any contribution of DNA methylation to the observed gene expression pattern.

To test whether variations in the transcript levels of ESC key master regulators could explain the differences in the gene expression pattern between poor- and high-quality iPSCs, we employed the single-molecule-mRNA fluorescent in situ hybridization (sm-mRNA-FISH) technique, which allows the quantification of mRNA transcripts of up to three genes in individual cells (Raj et al., 2010). The transcript number of three ESC key master regulators, Oct4, Sox2, and Esrrb, was quantified in two poor quality OSKM-iPSC lines, Nanog-GFP OSKM#3 and Oct4-GFP OSKM#4, and two high-quality SNEL-iPSC lines, Nanog-GFP SNEL#2 and Oct4-GFP SNEL#3. As depicted in Figures 4A and S4A, the transcript count of all three factors was comparable between the different lines. It should be noted that the transcript count of Esrrb was lower in the Nanog-GFP lines as compared to Oct4-GFP lines (Figures 4A and S4B). This result is consistent with the previous observation that *Esrrb* is a direct target of Nanog (Festuccia et al., 2012), as in our system GFP was introduced into the *Nanog* locus by replacing the coding region of the endogenous *Nanog* gene, thereby creating a mutant allele which reduces the total level of endogenous Nanog protein.

Superenhancers are regulatory elements that are associated with genes that determine cell identity (Hnisz et al., 2013). Therefore, the establishment of ESC-specific superenhancers during the reprogramming process is crucial to allow proper transcription and function. It has been shown that the binding of key master regulators, H3K27ac and the Mediator complex rigorously mark superenhancers (Whyte et al., 2013). To test whether poor quality iPSCs acquire aberrant establishment of superenhancers during the reprogramming process, we measured the enrichment of the Mediator complex on ESC-specific superenhancers by chromatin immunoprecipitation sequencing (ChIP-seq) in several lines. ChIP-seq for Med1, a component of the Mediator complex, was performed on two ESC lines, V6.5 and ZHBcT4, as positive controls, the parental MEFs and two previously described partially reprogrammed lines (i.e., cells that initiated the reprogramming process but never gave rise to stable iPSCs, 23 and 44) as negative controls (Buganim et al., 2012), three poor quality iPSC lines, Nanog-GFP OSKM#2, Oct4-GFP OSKM#4, and Oct4-GFP OSKM#8, and two high-quality iPSC lines, Oct4-GFP SNEL#3 and Oct4-GFP SNEL#4. The distribution of the Mediator complex throughout the genome was compared to the distribution of the binding of Oct4, Sox2, and Nanog (OSN), three ESC key master regulators. The recruitment of Med1 protein throughout the genome was comparable in all iPSC lines tested but was absent in the negative control cells as can be seen in three representative ESC-specific superenhancers for Sox2, Nanog, and miR290-295, (Figures 4B and S4C) and following hierarchical clustering analysis of Med1 densities in all superenhancers from all lines (Figure 4C). The results so far indicate that both poor- and high-quality iPSCs reach a complete activation of the endogenous pluripotency circuitry as measured by transcript count of ESC key master regulators and by establishment of ESC-specific superenhancers.

**Table 1. Generation of “All-iPSC” Mice Produced by Various Combinations of Reprogramming Factors**

Cell Line	Injected Embryos	Recipient Females	Implantation Sites	Dead Fetuses and Pups <sup>a</sup> (%)	Pups Born Alive <sup>b</sup> (%)	Fostered Pups <sup>c</sup> (%)
Oct4-GFP SNFI #1	60	3	6	1 (1.6)	2 (3.3)	1 (1.6)
Oct4-GFP SNFI #2	60	3	9	3 (5.0)	9 (15.0)	8 (13.3)
Oct4-GFP SNFI #3	60	3	9	2 (3.3)	14 (23.3)	13 (21.6)
Oct4-GFP SNFI #4	60	3	15	3 (5.0)	11 (18.3) <sup>d</sup>	6 (10.0)
Nanog-GFP SNFI #1	60	3	18	1 (1.6)	9 (15.0)	8 (13.3)
Nanog-GFP SNFI #2	60	3	12	0 (0.0)	15 (25.0)	14 (23.3)
Nanog-GFP SNFI #3	60	3	22	2 (3.3)	2 (3.3)	1 (1.6)
Nanog-GFP SNFI #4	60	3	3	0 (0.0)	0 (0.0)	0 (0.0)
Nanog-GFP SNFI #5	60	3	7	0 (0.0)	0 (0.0)	0 (0.0)
Nanog-GFP SNFI #6	60	3	22	1 (1.6)	2 (3.3)	0 (0.0)
Total (Average)	60	3	13.3	1.3 (2.1)	6.4 (10.7)	5.1 (8.5)
Oct4-GFP OSKM#2	60	3	18	0 (0.0)	4 (6.6) <sup>d</sup>	2 (3.3)
Oct4-GFP OSKM#3	60	3	16	0 (0.0)	5 (8.3) <sup>e</sup>	3 (5.0)
Oct4-GFP OSKM#4	60	3	37	1 (1.6)	0 (0.0)	0 (0.0)
Oct4-GFP OSKM#5	60	3	12	0 (0.0)	0 (0.0)	0 (0.0)
Oct4-GFP OSKM#6	60	3	29	1 (1.6)	1 (1.6)	1 (1.6)
Oct4-GFP OSKM#7	60	3	25	1 (1.6)	3 (5.0)	2 (3.3)
Oct4-GFP OSKM#8	60	3	16	0 (0.0)	0 (0.0)	0 (0.0)
Nanog-GFP OSKM#1	60	3	12	0 (0.0)	0 (0.0)	0 (0.0)
Nanog-GFP OSKM#2	60	3	24	0 (0.0)	0 (0.0)	0 (0.0)
Nanog-GFP OSKM#3	60	3	11	0 (0.0)	0 (0.0)	0 (0.0)
Total (average)	60	3	20	0.3 (0.48)	1.3 (2.1)	0.8 (1.7)
Oct4-GFP OSK#1	60	3	14	0 (0.0)	3 (5.0)	3 (5.0)
Oct4-GFP OSK#2	60	3	18	2 (3.3)	2 (3.3)	2 (3.3)
Oct4-GFP OSK#3	60	3	13	0 (0.0)	2 (3.3)	2 (3.3)
Oct4-GFP OSK#4	60	3	3	2 (3.3)	13 (21.6)	13 (21.6)
Oct4-GFP OSK#5	60	3	0	5 (8.3)	5 (8.3) <sup>d</sup>	4 (6.6)
Oct4-GFP OSK#6	60	3	6	0 (0.0)	1 (1.6)	1 (1.6)
Oct4-GFP OSK#7	60	3	10	0 (0.0)	1 (1.6)	1 (1.6)
Oct4-GFP OSK#8	60	3	3	0 (0.0)	3 (5.0)	3 (5.0)
Oct4-GFP OSK#9	60	3	14	0 (0.0)	0 (0.0)	0 (0.0)
Oct4-GFP OSK#10	60	3	28	0 (0.0)	0 (0.0)	0 (0.0)
Total (Average)	60	3	11	0.9 (1.49)	3 (5.0)	2.9 (4.8)
Oct4-GFP OSSNF#1	60	3	31	0 (0.0)	0 (0.0)	0 (0.0)
Oct4-GFP OSSNF#2	60	3	4	0 (0.0)	2 (3.3)	2 (3.3)
Oct4-GFP OSSNF#3	60	3	5	0 (0.0)	3 (5.0)	0 (0.0)

(Continued on next page)

**Table 1. Continued**

Cell Line	Injected Embryos	Recipient Females	Implantation Sites	Dead Fetuses and Pups <sup>a</sup> (%)	Pups Born Alive <sup>b</sup> (%)	Fostered Pups <sup>c</sup> (%)
Oct4-GFP OSSNF#4	60	3	23	0 (0.0)	0 (0.0)	0 (0.0)
Oct4-GFP OSSNF#5	60	3	11	0 (0.0)	0 (0.0)	0 (0.0)
Oct4-GFP OSSNE#6	60	3	27	0 (0.0)	0 (0.0)	0 (0.0)
Total (Average)	60	3	13	0 (0.0)	0.8 (1.3)	0.3 (0.6)
Nanog-GFP F5Cs	60	3	9	1 (1.6)	3 (5.0)	3 (5.0)
Oct4-GFP F5Cs	60	3	20	0 (0.0)	7 (1.6)	5 (8.3)
KH2 ESCs	60	3	21	1 (1.6)	6 (10.0)	5 (8.3)
Total (Average)	60	3	13	0.6 (1.0)	5.3 (8.8)	4.3 (7.2)

Summary of the production of "all-iPSC" mice from SNEL-iPSC, OSKM-iPSC, OSK-iPSC, and OSSNE-iPSC lines grown in 2i medium by tetraploid complementation. SNEL, Sall4, Nanog, Esrrb, and Lin28; OSKM, Oct4, Sox2, Klf4, and Myc; OSK, Oct4, Sox2, and Klf4; OSSNE, Oct4, Sox2, Sall4, Nanog, and Esrrb.

<sup>a</sup>Number of fetuses found dead in utero and pups found dead at the time of Caesarean section or immediately following natural birth. It also includes pups born with a hernia that were sacrificed immediately.

<sup>b</sup>Pups in this category exhibited assisted breathing at the time of birth that ceased shortly after birth.

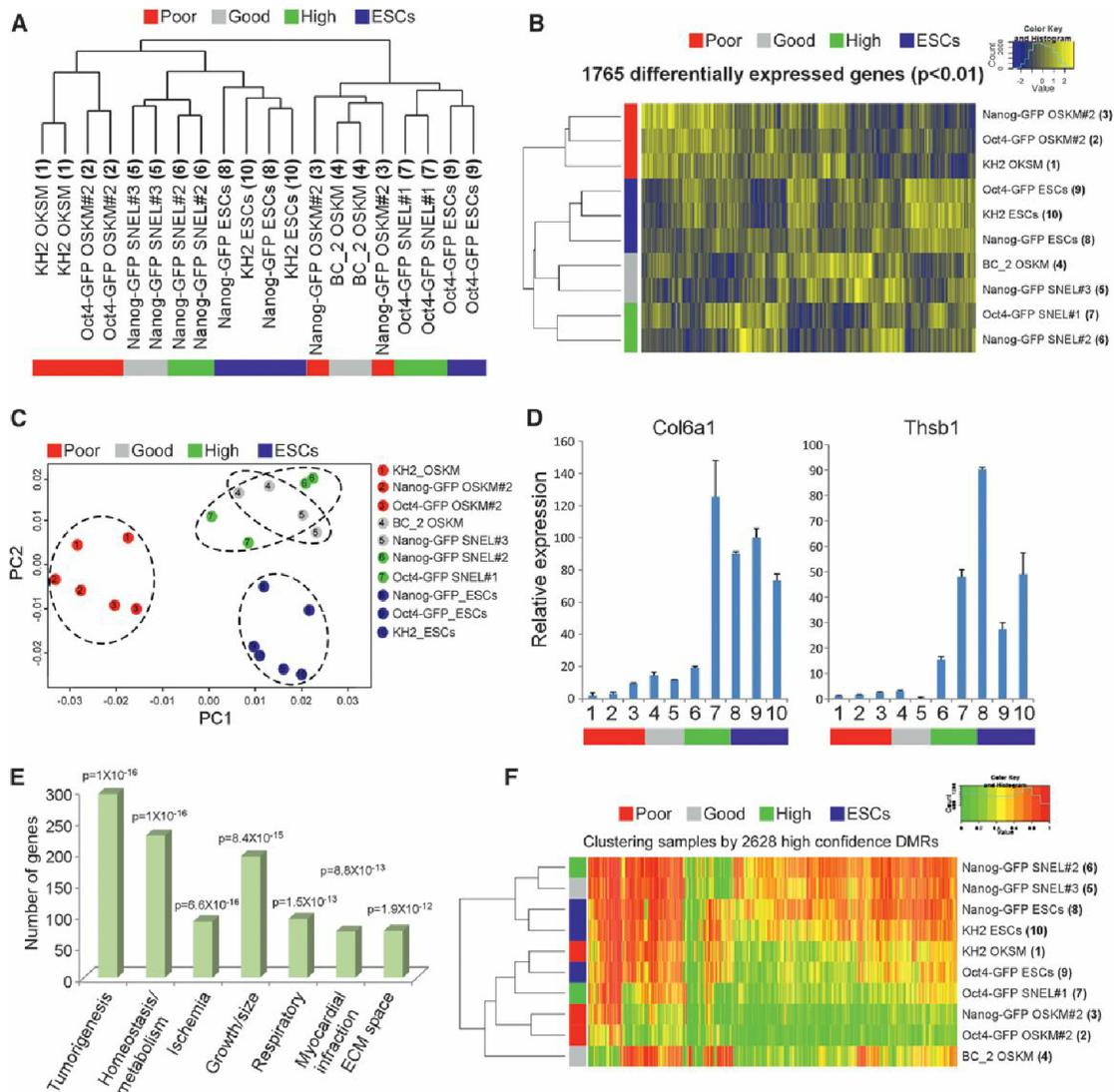
<sup>c</sup>Pups in this category exhibited independent breathing followed by fostering with lactating moms.

<sup>d</sup>Hernia.

<sup>e</sup>Developmental defect.

### Trisomy of Chromosome 8 Is a Frequent Genomic Aberration in OSKM/OSK-iPSC Lines

It has been demonstrated that oncogenic stress induced by Myc leads to DNA damage that promotes genomic instability and tumor progression (Vafa et al., 2002). To test whether OSKM elicit stronger DNA damage than SNEL when overexpressed, we infected MEFs with either OSKM, SNEL, or an empty vector as control and measured the levels of a well-known DNA damage sensor,  $\gamma$ -H2A.X phosphorylation (Bonner et al., 2008) 7 days postdox exposure. Consistent with Myc being a very potent oncogene (González et al., 2013; Marión et al., 2009; Müller et al., 2012), we observed a significant increase in  $\gamma$ -H2A.X phosphorylation level in OSKM-infected MEFs as compared to MEFs that were infected with empty vector or SNEL (Figure 5A). Consistent with that, a large fraction of propidium iodide-positive cells (10.9%) was observed in OSKM-infected MEFs, but not in SNEL or empty vector-infected MEFs, suggesting that OSKM expression leads to severe DNA damage that induces cell death (Figure 5B). This result led us to investigate whether OSKM-iPSCs acquire genomic aberrations throughout the reprogramming process, which might explain the premature death observed in 4n embryos produced by the poor quality OSKM-iPSCs. DNA rearrangements such as sister chromatid exchanges (SCEs) are sensitive indicators of genomic stress and instability, thus, we mapped SCEs using Strand-seq (Falconer et al., 2012). It has been shown that, at most, three ESCs eventually contribute to the formation of a 2n or 4n embryo out of 8–12 injected cells (Wang and Jaenisch, 2004). Thus, an iPSC clone that contains even only a few cells with genomic aberrations might still exhibit poor developmental potential. To overcome the heterogeneity observed within an iPSC colony, we measured SCEs at the single-cell level as opposed to the majority of the studies in the field that measured genomic instability of iPSCs by employing population-based analyses. Single cell sequencing libraries were made from MEFs, ESCs, two poor quality iPSC lines, Oct4-GFP OSKM#4 and Oct4-GFP OSKM#8, and two high-quality iPSC lines, Oct4-GFP SNEL#3 and Oct4-GFP SNEL#4, and SCE frequency was examined. As shown in Figures 5C, S5A, and S5B, both SNEL lines fall between the two OSKM lines, revealing no difference in SCE counts between the OSKM- and SNEL-iPSCs. We then hypothesized that long transient genomic instability during reprogramming could generate cells with higher susceptibility to acquire chromosomal aberrations. To that end, we also examined the chromosome content by single cell sequencing. Interestingly, as observed previously for poor quality ESCs, trisomy of chromosome 8 was detected in the majority of the Oct4-GFP OSKM#8 iPSCs. To test whether trisomy 8 is a common aberration in iPSC lines, single cell sequencing libraries were made from 22 cell lines (one ESC line, eight OSKM lines, six OSK lines, and seven SNEL lines) that were cultured for only five passages. For each colony, single cells were sorted and sequenced. Out of the eight OSKM lines analyzed, three showed cells with trisomy 8: 6% of the Nanog-GFP OSKM#3 cells, 91% of the Nanog-GFP OSKM#2, and 95% of the Oct4-GFP OSKM#8 (Figure 5D). Out of the six OSK lines, four showed cells with trisomy 8: 4% of the Oct4-GFP OSK#8 cells, 8% of Oct4-GFP OSK#7, 8% of Oct4-GFP OSK#10, and 95% of Oct4-GFP OSK#9. No trisomy of chromosome 8 was detected in any of the other lines tested (Figures 5D,



**Figure 3. Unbiased Comparative Transcriptome Analyses Distinguish iPSCs According to Their 4n Proficiency**

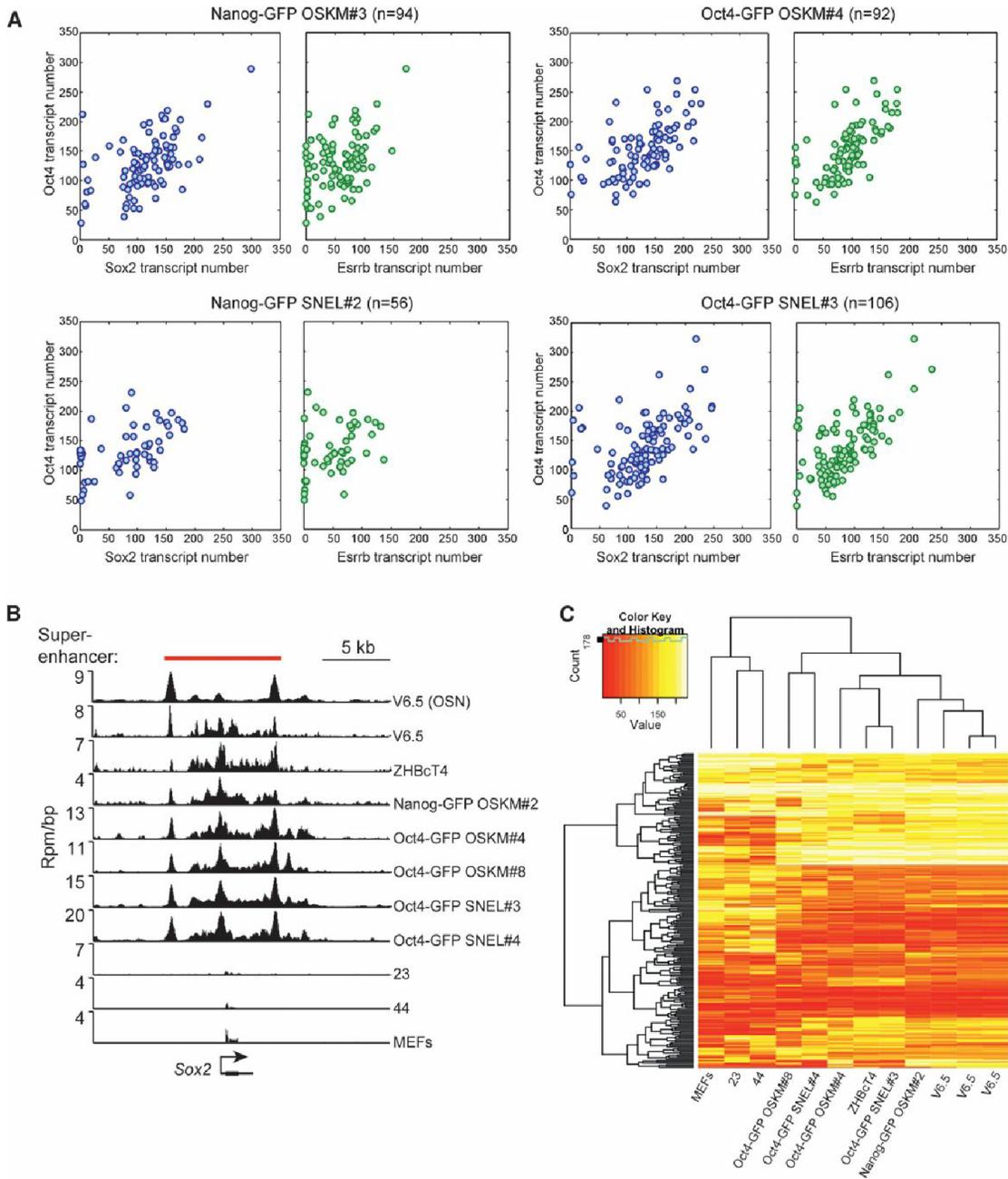
(A) Hierarchical clustering of global gene expression profiles for two microarray technical replicates for every iPSC and ESC (reference) line. Replicate pairs were assigned a shared numerical value.  
 (B) Hierarchical clustering of all genes ( $n = 1,765$ ) exhibiting significant variation ( $p < 0.01$  by F test) across all ESC and iPSC samples.  
 (C) Principle component analysis for genes from (B). Each of the iPSC and ESC groups is marked by specific color and is surrounded by a circle. The numbers inside the circles correspond to the numbers in (A).  
 (D) qRT-PCR of the *Col6a1* and *Thsb1* normalized to the *Hprt* housekeeping control gene in the indicated samples. Error bars are presented as a mean  $\pm$  SD of two duplicate runs from a typical experiment. The numbers on the x axis correspond to the numbers in (A).  
 (E) Gene ontology analysis using the GeneDecks (Stelzer et al., 2009) algorithm of genes from (A).  
 (F) Hierarchical clustering of 2,628 differentially methylated regions (DMRs) derived from whole genome bisulfite sequencing does not segregate samples by either reprogramming factor combinations or ESC versus iPSC status. Each group (poor, good, high, and ESCs) is marked by a different color. Sample numbers correspond to the numbers in (A).  
 See also Figure S3.

S5C, and S5D). On average, 21% of the OSKM cells and 19% of the OSK cells had a trisomy 8, versus 0% in the SNEL cells and the ESCs ( $p = 0.002$  and  $p = 0.005$ , respectively). We also examined whether we could detect any clear difference in the frequency of aneuploidy between the lines once trisomy 8 was excluded from the analysis. On average, OSKM lines attained a 10% frequency of aneuploidy compared with 13% for the OSK lines, 9% for the SNEL lines, and 9% for the ESC line (differences

are not significant). These results suggest that trisomy 8 is a common aberration in iPSCs-derived from a combination of factors that includes the key factors, Oct4, Sox2, and Klf4.

**High-Quality SNEL-iPSCs Are Correlated with Faithful H2A.X Deposition Patterns**

Defective H2A.X deposition frequently occurs in iPSC clones generated by OSKM factors that failed to support “all-iPS”



**Figure 4. The Transcript Number of Key Master Regulators and the Establishment of ESC-Specific Superenhancers Are Comparable between Poor- and High-Quality iPSCs**

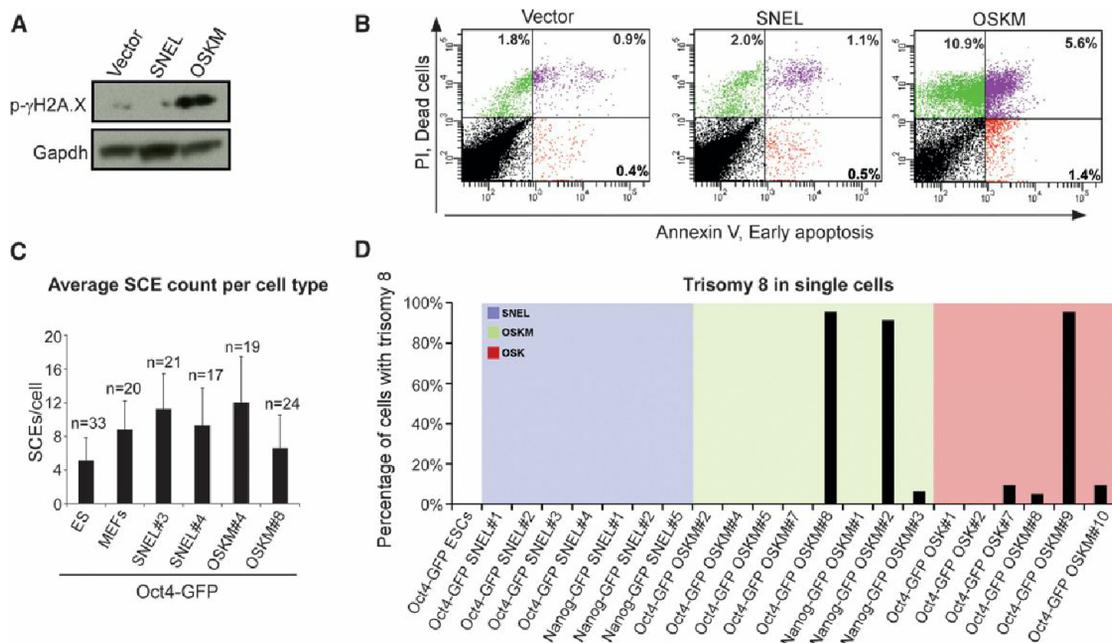
(A) sm-mRNA-FISH experiments depict the transcript number of Oct4 versus Sox2 and Oct4 versus Esrrb in single cells from the indicated iPSC lines. n, represents the number of single cells analyzed.

(B) ChIP-seq binding profiles for Oct4, Sox2, and Nanog (merged, OSN) in V6.5 mESCs and Med1 for the indicated cell lines at the Sox2 locus. Location of the superenhancer, as defined in V6.5 mESCs (Whyte et al., 2013), is indicated by the red bar. Rpm/bp, reads per million per base pair.

(C) Hierarchical clustering of Med1 densities in superenhancers recapitulates phylogeny of cell types. ChIP-seq read densities for Med1 were calculated in mES superenhancers. Clustering these densities indicates that cell types of similar origin have similar signal of Med1 in superenhancers. All ChIP-seq was performed with a Bethyl Laboratories antibody (A300-793A, lot A300-783A), except for the farthest right V6.5 ChIP, which was performed with a Santa Cruz Biotechnology antibody (SC-5334X, lot A1112). See also Figure S4.

mice development in tetraploid complementation experiments (Wu et al., 2014). In addition, our results suggest that ectopic expression of OSKM in MEFs induces a strong DNA damage

response that leads to the accumulation of  $\gamma$ -H2A.X phosphorylation (Figure 5A). We therefore hypothesized that the genome-wide H2A.X deposition patterns might be intact in



**Figure 5. Trisomy of Chromosome 8 Is a Frequent Genomic Aberration in Poor-Quality OSK/OSKM-iPSCs**

(A) Western blot analysis for the DNA damage sensor,  $\gamma$ -H2A.X phosphorylation, and the housekeeping control protein Gapdh, 7 days postdox exposure in MEFs infected with the indicated dox-inducible reprogramming factors.

(B) FACS analyses demonstrating the percentage of cells from (A) that initiated the apoptotic process as assessed by PI and Annexin V.

(C) Graph summarizing the average number of sister chromatid exchanges (SCEs) occurring at the single cell level in the indicated lines using the Strand-seq technique. n, represents the number of single cells tested. Error bars are presented as a mean  $\pm$  SD of the indicated "n" examined single cells.

(D) Graph summarizing the frequency of trisomy 8 observed in the indicated colonies.

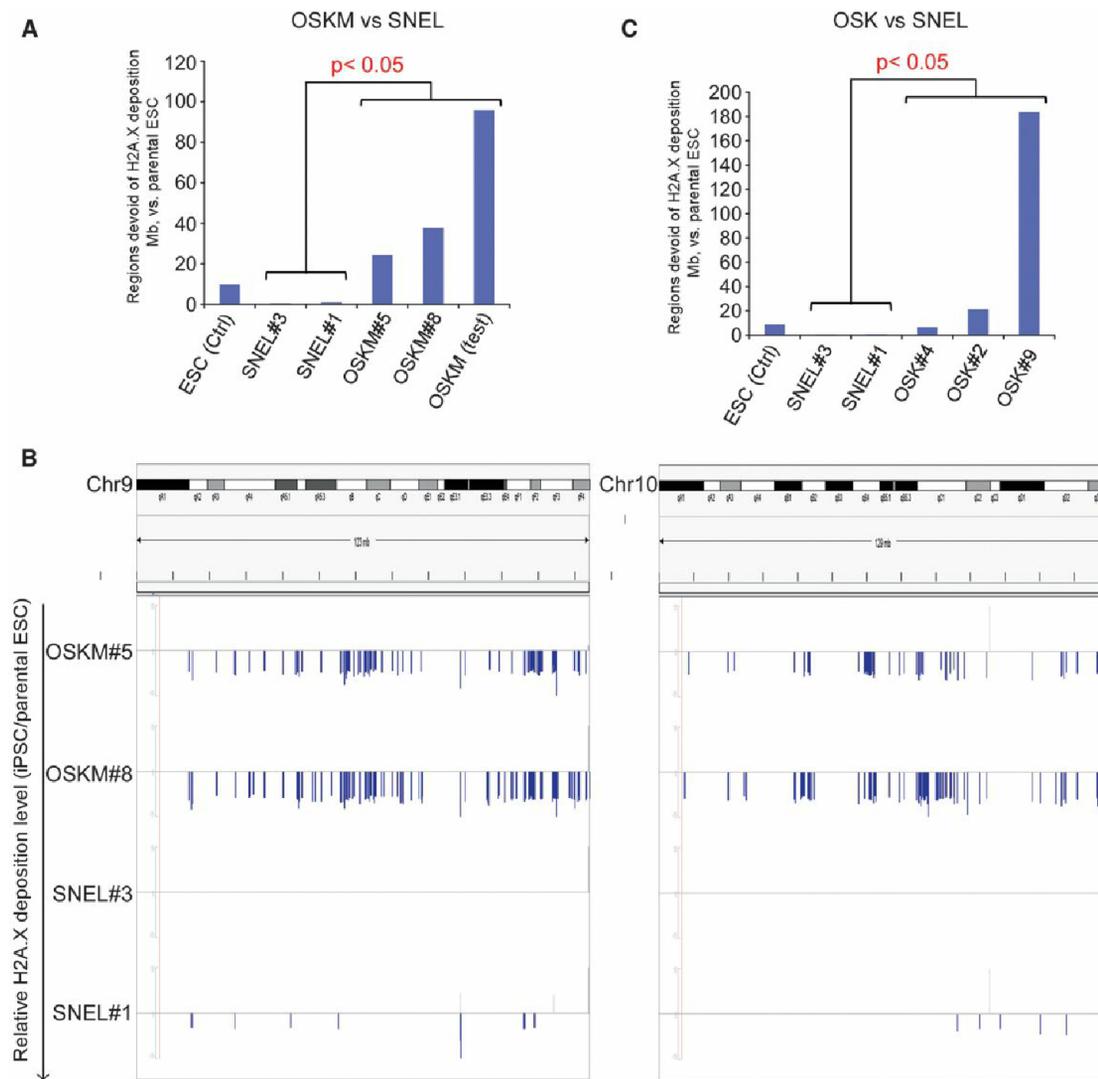
See also Figure S5.

high-quality SNEL-iPSCs, while abnormal in poor quality OSKM-iPSCs and OSK-iPSCs. To test this hypothesis, we performed ChIP-seq for H2A.X on two poor quality OSKM-iPSC lines, Oct4-GFP OSKM#5 and Oct4-GFP OSKM#8, two high-quality SNEL-iPSC lines, Oct4-GFP SNEL#1 and Oct4-GFP SNEL#3, two high-quality OSK-iPSC lines, Oct4-GFP OSK#4 and Oct4-GFP OSK#2, and one poor quality OSK-iPSC line, Oct4-GFP OSK#9, and compared them to the parental Oct4-GFP ESCs. We used an established Hidden-Markov-Model (HMM) algorithm (Song and Smith, 2011) to inspect the differential H2A.X deposition regions in these cells (Wu et al., 2014). H2A.X deposition patterns in the high-quality Oct4-GFP SNEL#1 and SNEL#3 iPSC lines, were almost identical to the parental Oct4-GFP ESC line. Moreover, the number of H2A.X devoid regions was significantly lower in SNEL-iPSCs even compared to the control ESC line; devoid regions that mark the natural variations between ESC (e.g., the parental Oct4-GFP ESC versus 129SVj/c57 ESC control line) (Figures 6A and 6C). In contrast, H2A.X deposition patterns were greatly different in poor quality lines, Oct4-GFP OSKM#5, Oct4-GFP OSKM#8, and another iPSC cell line (OSKM\_test) generated by the secondary inducible system that failed tetraploid complementation assays (Stadtfield et al., 2010); these clones had significantly more differential H2A.X deposition regions than the high-quality SNEL-iPSC clones in terms of the genome coverage (100- to 400-fold, p value < 0.05, Wilcoxon test, Figures 6A and 6B). The same trend was observed in the OSK-

iPSC lines. H2A.X deposition patterns were also greatly different in the poor quality line (p value < 0.05, Wilcoxon test, Figure 6C) and comparable to control ESCs in the high-quality OSK-iPSC lines. However, the differences between the high-quality OSK-iPSC lines and the parental Oct4-GFP ESC line were still greater than those between SNEL-iPSCs and the same parental ESCs. These results demonstrate a strong correlation between the capability to produce "all-iPSC" mice and the capability to faithfully recapitulate an intact H2A.X deposition. These data raise the possibility that SNEL reprogramming factors can generate frequently high-quality iPSCs by retaining intact H2A.X deposition throughout the genome.

## DISCUSSION

Tetraploid complementation is considered to be the most stringent assay for pluripotency and is frequently used to assess the quality of iPSCs. Even though iPSCs have been shown by several laboratories to be 4n competent, only a small fraction of the tested lines passed this most stringent test (Boland et al., 2009; Jiang et al., 2011, 2013; Kang et al., 2009; Pera, 2011; Zhao et al., 2009). Our experiments show that the rate of high-quality iPSCs as assessed by 4n competence is significantly influenced by the choice of factors used to induce reprogramming. We demonstrate that the SNEL factors, which are downstream targets of the late pluripotency factor Sox2



**Figure 6. H2A.X Deposition Pattern Can Distinguish SNEL and OSK or OSKM iPSC Lines**

(A) The bar chart illustrates that the genome coverage of defective H2A.X deposition regions in OSKM-iPSC lines ( $n = 3$ ) are significantly higher than those in 4N+SNEL-iPSC lines ( $n = 2$ ) or the ESC control line ( $p$  value  $< 0.05$ , Wilcoxon rank sum test).

(B) Comparative H2A.X depositions in SNEL iPSC and OSKM iPSC at two chromosomes (left: Chr9; right: Chr10). y axis: relative H2A.X deposition level (RSEG score, as compared to the ESC parental line, see Experimental Procedures). Positive value: regions enriched for H2A.X deposition over control (gray bars); negative values: regions devoid of H2A.X deposition over control (blue bars).

(C) The bar chart illustrates that the genome coverage of H2A.X defective deposition regions in OSK-iPSC lines are significantly higher than those in SNEL-iPSC lines ( $n = 2$ ) ( $p$  value  $< 0.05$ , Wilcoxon rank sum test).

(Buganim et al., 2012), produce iPSCs that have a considerably higher competence to generate “all-iPSC” mice by 4n complementation than iPSCs produced by the conventional OSKM factors or other combinations such as OSK or OSSNE. The various iPSC lines that were produced throughout this study were generated, isolated and tested under identical conditions to rule out effects caused by variations in cell culture, method of factor delivery, the passage number of the lines, or blastocyst injections. While reprogramming by OSKM produced many colonies with shorter latency but with variable quality, reprogramming by SNEL yielded significantly fewer iPSCs the majority of which, however, were of high quality. This is reminiscent of previous studies in which similar transgenic OSKM systems producing

iPSCs with different efficiencies resulted in different quality iPSCs, with the more efficient system (Stadtfeld et al., 2010) producing iPSCs of lower quality—as assessed by 4n complementation—than the less efficient system (Carey et al., 2011). These two studies suggest that high reprogramming efficiency may result in a low fraction of high-quality iPSCs.

To define molecular signature of 4n competency, we compared the transcriptional profile of poor, good and high-quality iPSCs. Genes involved in “Respiratory,” “Ischemia,” and “Myocardial infarction” separated high, good, and poor quality iPSCs consistent with the observation that poor quality iPSC lines produced “all-iPSC” pups that were either retarded in development or died in utero. Importantly, neither DNA

methylation patterns, as revealed by whole genome bisulfite sequencing, nor the transcript level of key master regulators such as Oct4, Sox2, and Esrrb and the establishment of ESC-specific superenhancers could explain the difference in the transcriptional profile between the groups or could separate the groups by 4n proficiency. Instead, we demonstrated that a large number of OSKM/OSK-iPSCs acquire a unique genomic aberration at least in some fraction of cells. While the overall incidence of aneuploidy between MEFs, ESCs, and poor- and high-quality iPSC colonies was comparable, trisomy of chromosome 8 was solely observed in iPSCs that were produced with a combination of factors that includes Oct4, Sox2, and Klf4. As shown previously, trisomy 8 ESCs have a proliferative advantage but poor developmental potential (Liu et al., 1997). Consistent with the notion that expression of OSKM or OSK leads to genome instability, we and others (González et al., 2013) demonstrated that expression of OSKM induces DNA damage as assessed by the accumulation of  $\gamma$ -H2A.X phosphorylation. This DNA damage led to cell death in ~17% (a sum of an early and late apoptotic cells) of the OSKM-infected cells, showing that intense and, in some cases, irreversible DNA damage is caused by this combination of factors. In accordance with the accumulation of  $\gamma$ -H2A.X phosphorylation, the deposition of H2A.X throughout the genome was abnormal in poor quality OSKM and OSK-iPSCs, as well. More importantly, high-quality SNEL-iPSC lines exhibited a significantly lower number of regions devoid of H2A.X deposition as compared to control ESC line or high-quality OSK-iPSC lines. Because histones regulate transcription, it is tempting to speculate that the aberrant deposition of H2A.X in poor quality iPSC lines might explain the mild difference in the gene expression between poor- and high-quality iPSCs.

In summary, our study provides a proof of principle that different combinations of reprogramming factors do not equally affect the biological characteristics of iPSCs, with some combinations consistently resulting in high-quality cells, whereas others generate cells of variable quality. While genomic instability and the deposition of H2A.X partially explain why poor quality iPSCs fail in the 4n complementation test, the reason why most colonies from other combinations of factors, that lack potent oncogenes such as OSSNE failed in this test, requires further investigation. Our results indicate that the interplay between the reprogramming factors is a crucial determinant for reprogramming efficiency and quality. This is consistent with the observation that conversion of cells into other cell types depends on the composition of reprogramming factors (Graf and Enver, 2009). Attempts to reprogram human cells to pluripotency using SNEL reprogramming factors have so far failed (data not shown), suggesting that the activation of the core pluripotency circuitry of human cells might be different than that of the mouse or that the reprogramming efficiency of SNEL in human cells is drastically low so that one needs to infect a large number of fibroblasts in order to achieve reprogramming. Based on these results it will be important to define the most optimal factor combinations for reprogramming and to assess how different factor combinations might affect the quality of human iPSCs as a step forward toward transplantation therapy.

## EXPERIMENTAL PROCEDURES

### Cell Culture and Mice

Mouse embryonic fibroblasts (MEFs) were grown in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and antibiotics. ESCs and iPSCs and were grown in DMEM supplemented with 15% fetal bovine serum, 1% nonessential amino acids, 2 mM L-glutamine,  $2 \times 10^6$  units mouse leukemia inhibitory factor (mLif), 0.1 mM  $\beta$ -mercaptoethanol (Sigma), and antibiotics or in 2i medium. 2i medium (500 ml) was generated by including: 230 ml DMEM/F12 (Invitrogen; 11320), 230 ml Neurobasal medium (Invitrogen; 21103), 5 ml N2 supplement (Invitrogen; 17502048), 10 ml B27 supplement (Invitrogen; 17504044), 10 ml (2%) fetal bovine serum,  $2 \times 10^6$  units mLif, 1 mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM  $\beta$ -mercaptoethanol (Sigma), penicillin-streptomycin (Invitrogen), 5  $\mu$ g/ml BSA (Sigma), PD0325901 (PD, 1  $\mu$ M), and CHIR99021 (CH, 3  $\mu$ M). All the cells were maintained in a humidified incubator at 37°C and 5% CO<sub>2</sub>. For the primary infection, MEFs were isolated from mice heterozygous for the reverse tetracycline-dependent transactivator (M2rtTA) that resides in the ubiquitously expressed *Gt(ROSA)26Sor* locus (Beard et al., 2006) and either with GFP that was knocked-in the *Nanog* or the *Oct4* locus. All infections were performed on MEFs (passage 0) that were seeded at 70% confluence 2 days before the first infection. Animal care was in accordance with institutional guidelines and was approved by the Committee on Animal Care, Department of Comparative Medicine, Massachusetts Institute of Technology.

### Tetraploid Embryo Complementation and Chimera Formation

Blastocyst injections were performed using (C57/Bl<sub>6</sub>xDBA) B<sub>6</sub>D<sub>2</sub>F<sub>2</sub> host embryos. All injected iPSC lines were derived from crosses of 129Sv/Jae to C57/Bl<sub>6</sub> mice and could be identified by agouti coat color. Embryos were obtained 24 hr (1 cell stage) or 40 hr (2 cell stage) posthuman chorionic gonadotropin (hCG) hormone priming. To obtain tetraploid (4n) blastocysts, electrofusion was performed at ~44–47 hr post-hCG using a BEX LF-101 or LF-301 cell fusion apparatus (Protech International). Both fused and diploid embryos were cultured in EmbryoMax KSOM (Millipore) or Evolve KSOMaa (Zenith Biotech) until they formed blastocysts (94–98 hr after hCG injection) at which point they were placed in a drop of Evolve w/HEPES KSOMaa (Zenith) medium under mineral oil. A flat tip microinjection pipette with an internal diameter of 16  $\mu$ m (Origio) was used for iPSC injections. Each blastocyst received 10–12 iPSCs. Shortly after injection, blastocysts were transferred to day 2.5 recipient CD1 females (20 blastocysts per female). Pups, when not born naturally, were recovered at day 19.5 by cesarean section and fostered to lactating Balb/c mothers.

### SSLP Assay

PCR reactions were set up as previously described (Stadtfeld et al., 2012) using genomic DNA from the 4n pups and primers reported to detect polymorphisms in the genome of inbred mouse strains (D8Mit94-F: GTTGGG GCTCTGCTCTCTC; D8Mit94-R: CACATATGCATAC ATATACATACAGT; D2Mit102-F: TTCCCTGTCACTCTCC; D2Mit102-R: TGTCTTT ATGCTCA GACATACACA). As controls, genomic DNA was analyzed from cultured iPSCs used for injections and adult mice (B<sub>6</sub>D<sub>2</sub>F<sub>1</sub>) that served as hosts for the blastocysts. The reactions were performed with 100 ng DNA for 30 cycles of 30 s at 94°C, 30 s at 60°C, and 60 s at 72°C. The products were analyzed on a 3% agarose gel.

### Gene Expression Microarrays

Two micrograms of RNA extracted using the RNeasy kit (QIAGEN) from each iPSC and ESC were tested by Agilent BioAnalyzer assays to ensure sample integrity. Technical duplicates of each sample were run on Agilent SurePrint G3 8x60K microarrays per manufacturer's instructions. Raw expression values were quantile normalized and binary log values of technical duplicate averages used for downstream analyses.

### Gene Expression Analyses

Standardized scoring was used to normalize microarray expression values. To generate heatmaps, expression values were used as input to perform clustering by samples in R, using the heatmap.2 function of the gplots library package with the default Euclidean distance parameter. In parallel, normalized expression values were used to group samples by principle component

analysis using the R princomp function. Gene ontology analysis was performed using the publicly available GeneDecks V3 software suite available at <http://www.genecards.org/>.

#### Whole-Genome Bisulfite Library Preparation and Sequencing

Bisulfite treatment and sequencing library preparation were performed as previously described (Lister et al., 2011). Briefly, genomic DNA from each iPSC and ESC was fragmented by sonication prior to ligation of methylated sequencing adaptors. Fragments were size-selected by gel electrophoresis and purification, treated with sodium-bisulfite and amplified with four cycles of PCR to generate libraries that were sequenced on the Illumina HiSeq platform. Fastq output files were aligned using bowtie (Langmead et al., 2009).

#### Illumina Sequencing

Libraries were pooled for sequencing and 200 to 400 bp size range fragments were purified using a 2% E-Gel Agarose Gel (Invitrogen). DNA quality was assessed and quantified on a High Sensitivity dsDNA kit (Agilent) on the Agilent 2100 Bio-Analyzer and on the Qubit 2.0 Fluorometer (Life Technologies). For sequencing, clusters were generated on the CBot (HiSeq2500), and single-end 50 bp reads were generated using the HiSeq2500 sequencing platform (Illumina).

#### Bioinformatic Analysis

Demultiplexed fastq files were aligned to the mouse reference genome assembly (GRCm38/mm10) using short read aligner Bowtie2 (version 2.0.5) (Langmead and Salzberg, 2012) with default settings. Sorted and indexed bam files were processed as previously described (Falconer et al., 2012) using the BAIT software package (Hills et al., 2013). Briefly, diploid and aneuploid chromosomes were identified by calculating the average numbers of reads/Mb for each chromosome in the library. Monosomies and trisomies were classified when chromosomes had an average read count 0.66× lower or 1.33× higher, respectively than the average reads per Mb for diploid chromosomes in the library.

#### Statistics

The average frequency of trisomy 8 and aneuploidy per cell type was determined and p values were calculated using binomial distribution models.

#### Chromatin Immunoprecipitation for Med1 Distribution

Cells were crosslinked for 10 min at room temperature by the addition of one-tenth of the volume of 11% formaldehyde solution (11% formaldehyde, 50 mM HEPES [pH 7.3], 100 mM NaCl, 1 mM EDTA [pH 8.0], 0.5 mM EGTA [pH 8.0]) to the growth media. Cells were washed twice with PBS, then the supernatant was aspirated and the cell pellet was flash frozen in liquid nitrogen. Frozen crosslinked cells were stored at  $-80^{\circ}\text{C}$ . Dynal magnetic beads (100  $\mu\text{l}$ ) (Sigma) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were bound with 10  $\mu\text{g}$  of the indicated antibody. For Med1 (CRSP1/TRAP220) occupied genomic regions, we performed ChIP-seq experiments using a Bethyl Laboratories antibody (A300-793A, lot A300-783A-2) or a Santa Cruz Biotechnology antibody (SC-5334X, lot A1112).

Crosslinked cells were lysed with lysis buffer 1 (50 mM HEPES [pH 7.3], 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100) and resuspended and sonicated in sonication buffer (20 mM Tris-HCl [pH 8.0], 150 mM NaCl, 2 mM EDTA [pH 8.0], 0.1% SDS, and 1% Triton X-100). Cells were sonicated for ten cycles at 30 s each on ice (18–21 W) with 60 s on ice between cycles. Sonicated lysates were cleared and incubated overnight at  $4^{\circ}\text{C}$  with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Beads were washed two times with sonication buffer, one time with sonication buffer with 500 mM NaCl, one time with LiCl wash buffer (10 mM Tris [pH 8.0], 1 mM EDTA, 250 mM LiCl, 1% NP-40), and one time with TE with 50 mM NaCl. DNA was eluted in elution buffer (50 mM Tris-HCl [pH 8.0], 10 mM EDTA, 1% SDS). Crosslinks were reversed overnight. RNA and protein were digested using RNase A and Proteinase K, respectively, and DNA was purified with phenol chloroform extraction and ethanol precipitation.

#### Illumina Sequencing and Library Generation

Purified ChIP DNA was used to prepare Illumina multiplexed sequencing libraries. Libraries for Illumina sequencing were prepared following the Illu-

mina TruSeq DNA Sample Preparation v2 kit protocol with the following exceptions. After end-repair and A-tailing, immunoprecipitated DNA ( $\sim 10$ –50 ng) or Whole Cell Extract DNA (50 ng) was ligated to a 1:50 dilution of Illumina Adaptor Oligo Mix assigning one of 24 unique indexes in the kit to each sample. Following ligation, libraries were amplified by 18 cycles of PCR using the HiFi NGS Library Amplification kit from KAPA Biosystems. Amplified libraries were then size-selected using a 2% gel cassette in the Pippin Prep system from Sage Science set to capture fragments between 200 and 400 bp. Libraries were quantified by qPCR using the KAPA Biosystems Illumina Library Quantification kit according to kit protocols. Libraries with distinct TruSeq indexes were multiplexed by mixing at equimolar ratios and running together in a lane on the Illumina HiSeq 2000 for 40 bases in single read mode.

#### ChIP-seq Analysis

Mouse ESC superenhancers were downloaded from Whyte et al. (2013). Briefly, these superenhancers were identified by (1) intersecting regions enriched in Oct4, Sox2, and Nanog ChIP-seq to locate constituent enhancers, (2) stitching proximal enhancers into domains, and (3) separating superenhancers from typical enhancers by signal of Med1.

ChIP-seq reads for Med1 in all cell types were aligned to the mouse mm9 reference genome using bowtie (Langmead et al., 2009) with command-line parameters `-k 1 -m 1 -best -n 2`. Duplicate (multiple reads per position) reads were removed. Reads were artificially extended 200 bp downstream and their reads-per-million normalized density was calculated in mES superenhancers as described in (Lin et al., 2012). These densities were then hierarchically clustered in two dimensions using heatmap.2.

#### ACCESSION NUMBERS

The accession number for the data resulting from whole-genome bisulfite sequencing as discussed in the text is GSE59696; for Mediator ChIP Sequencing, GSE59569; for whole-transcriptome analysis, GSE45173; and for H2A.X ChIP sequencing, GSE55731.

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2014.07.003>.

#### AUTHOR CONTRIBUTIONS

Y.B. and R.J. conceived the study. Y.B., S.M., and R.J. wrote the manuscript and prepared the figures. Y.B. designed the experiments and performed cloning of the various factors, MEF infection, iPSC line isolation from the various combinations of factors, embryo C-section, FACS for apoptotic cells, real-time PCR for pluripotency markers, teratoma assay, and immunostaining for pluripotency markers. S.M. performed iPSC injections, tetraploid complementation assay, determination of chimeric contributions, C-sections, and SSLP PCR. N.V.W., D.P., and P.M.L. performed and analyzed the SCE and Trisomy 8 experiments. H.H., B.J.A., and R.A. performed and analyzed the superenhancer activation experiments. T.W. and A.X. performed and analyzed the H2A.X deposition experiments. K.G. and L.S. performed part of the tetraploid complementation experiments. B.A.Z. and Y.H. analyzed the mRNA microarray. E.K. and M.C. performed part of the C-sections and isolated several iPSC colonies. D.A.F. performed and analyzed the sm-mRNA-FISH experiments. Q.G. analyzed the teratoma experiments. S.S. performed the gamma-H2A.X western blotting. J.G. performed the human SNEL reprogramming experiments. Y.H., M.D.S., J.R.N., and J.R.E. performed and analyzed the methylome experiments.

#### ACKNOWLEDGMENTS

We thank Biology and Research Computing (BaRC), especially Prathapan Thiru at the Whitehead Institute's Center for Microarray Technology, for computational and technical assistance. Y.B. was supported by an NIH

Kirschstein National Research Service Award (1 F32GM099153-01A1). D.A.F. is a Vertex Scholar and was supported by a National Science Foundation (NSF) Graduate Research Fellowship and Jerome and Florence Brill Graduate Student Fellowship. R.J. is an adviser to Stemgent and a cofounder of Fate Therapeutics. J.R.E. is an investigator of the Howard Hughes Medical Institute. This research project was supported by NIH grants HD 045022 and R37CA084198 to R.J., the Gordon and Betty Moore Foundation (GMBF3034) and Chapman Foundation grants to J.R.E., and the Israeli Centers of Research Excellence (I-CORE) program (Center No. 41/11) grant to Y.B.

Received: May 1, 2013

Revised: March 9, 2014

Accepted: July 15, 2014

Published: September 4, 2014

## REFERENCES

- Audo, I., Bujakowska, K.M., Léveillard, T., Mohand-Saïd, S., Lancelot, M.E., Germain, A., Antonio, A., Michiels, C., Saraiva, J.P., Letexier, M., et al. (2012). Development and application of a next-generation-sequencing (NGS) approach to detect known and novel gene defects underlying retinal diseases. *Orphanet J. Rare Dis.* 7, 8.
- Bar-Nur, O., Russ, H.A., Efrat, S., and Benvenisty, N. (2011). Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell* 9, 17–23.
- Barlow, J.H., Faryabi, R.B., Callén, E., Wong, N., Malhowski, A., Chen, H.T., Gutierrez-Cruz, G., Sun, H.W., McKinnon, P., Wright, G., et al. (2013). Identification of early replicating fragile sites that contribute to genome instability. *Cell* 152, 620–632.
- Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* 44, 23–28.
- Boland, M.J., Hazen, J.L., Nazor, K.L., Rodriguez, A.R., Gifford, W., Martin, G., Kupriyanov, S., and Baldwin, K.K. (2009). Adult mice generated from induced pluripotent stem cells. *Nature* 461, 91–94.
- Bonner, W.M., Redon, C.E., Dickey, J.S., Nakamura, A.J., Sedelnikova, O.A., Solier, S., and Pommier, Y. (2008). GammaH2AX and cancer. *Nat. Rev. Cancer* 8, 957–967.
- Brambrink, T., Hochedlinger, K., Bell, G., and Jaenisch, R. (2006). ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable. *Proc. Natl. Acad. Sci. USA* 103, 933–938.
- Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150, 1209–1222.
- Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet.* 14, 427–439.
- Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creighton, M.P., et al. (2011). Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell* 9, 588–598.
- Chen, J., Liu, J., Chen, Y., Yang, J., Chen, J., Liu, H., Zhao, X., Mo, K., Song, H., Guo, L., et al. (2011). Rational optimization of reprogramming culture conditions for the generation of induced pluripotent stem cells with ultra-high efficiency and fast kinetics. *Cell Res.* 21, 884–894.
- Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiwu, O., Richter, L., Zhang, J., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5, 111–123.
- Cho, J., Chang, H., Kwon, S.C., Kim, B., Kim, Y., Choe, J., Ha, M., Kim, Y.K., and Kim, V.N. (2012). LIN28A is a suppressor of ER-associated translation in embryonic stem cells. *Cell* 151, 765–777.
- Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., et al. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* 41, 1350–1353.
- Esteban, M.A., and Pei, D. (2012). Vitamin C improves the quality of somatic cell reprogramming. *Nat. Genet.* 44, 366–367.
- Falconer, E., Hills, M., Naumann, U., Poon, S.S., Chavez, E.A., Sanders, A.D., Zhao, Y., Hirst, M., and Lansdorp, P.M. (2012). DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat. Methods* 9, 1107–1112.
- Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* 11, 477–490.
- González, F., Georgieva, D., Vanoli, F., Shi, Z.D., Stadtfeld, M., Ludwig, T., Jasin, M., and Huangfu, D. (2013). Homologous recombination DNA repair genes play a critical role in reprogramming to a pluripotent state. *Cell Reports* 3, 651–660.
- Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. *Nature* 471, 63–67.
- Graf, T., and Enver, T. (2009). Forcing cells to change lineages. *Nature* 462, 587–594.
- Hanna, L.A., Foreman, R.K., Tarasenko, I.A., Kessler, D.S., and Labosky, P.A. (2002). Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. *Genes Dev.* 16, 2650–2661.
- Hills, M., O'Neill, K., Falconer, E., Brinkman, R., and Lansdorp, P.M. (2013). BAIT: organizing genomes and mapping rearrangements in single cells. *Genome Med.* 5, 82.
- Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947.
- Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Närvä, E., Ng, S., Sourour, M., Hämmäläinen, R., Olsson, C., et al. (2011). Copy number variation and selection during reprogramming to pluripotency. *Nature* 471, 58–62.
- Jiang, J., Ding, G., Lin, J., Zhang, M., Shi, L., Lv, W., Yang, H., Xiao, H., Pei, G., Li, Y., et al. (2011). Different developmental potential of pluripotent stem cells generated by different reprogramming strategies. *J. Mol. Cell Biol.* 3, 197–199.
- Jiang, J., Lv, W., Ye, X., Wang, L., Zhang, M., Yang, H., Okuka, M., Zhou, C., Zhang, X., Liu, L., et al. (2013). Zscan4 promotes genomic stability during reprogramming and dramatically improves the quality of iPSC cells as demonstrated by tetraploid complementation. *Cell Res.* 23 (1), 92–106.
- Kang, L., Wang, J., Zhang, Y., Kou, Z., and Gao, S. (2009). iPSC cells can support full-term development of tetraploid blastocyst-complemented embryos. *Cell Stem Cell* 5, 135–138.
- 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.
- Kim, K., Zhao, R., Doi, A., Ng, K., Unteraehrer, J., Cahan, P., Huo, H., Loh, Y.H., Aryee, M.J., Lensch, M.W., et al. (2011). Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat. Biotechnol.* 29, 1117–1119.
- Klemm, S., Semrau, S., Wiebrands, K., Mooijman, D., Faddah, D.A., Jaenisch, R., and van Oudenaarden, A. (2014). Transcriptional profiling of cells sorted by RNA abundance. *Nat. Methods* 11, 549–551.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.
- Laurent, L.C., Ulitsky, I., Slavina, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J.V., Lee, S., Barrero, M.J., et al. (2011). Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell* 8, 106–118.

- Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**, 56–67.
- Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73.
- Liu, Y., and Labosky, P.A. (2008). Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. *Stem Cells* **26**, 2475–2484.
- Liu, X., Wu, H., Loring, J., Hormuzdi, S., Disteché, C.M., Bornstein, P., and Jaenisch, R. (1997). Trisomy eight in ES cells is a common potential problem in gene targeting and interferes with germ line transmission. *Dev. Dyn.* **209**, 85–91.
- Marión, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature* **460**, 1149–1153.
- Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., and Benvenisty, N. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* **7**, 521–531.
- Müller, L.U., Milsom, M.D., Harris, C.E., Vyas, R., Brumme, K.M., Parmar, K., Moreau, L.A., Schambach, A., Park, I.H., London, W.B., et al. (2012). Overcoming reprogramming resistance of Fanconi anemia cells. *Blood* **119**, 5449–5457.
- Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., et al. (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* **151**, 68–79.
- Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J.M., Guo, T., Qi, Z., Downey, S.L., Manos, P.D., Rossi, D.J., et al. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat. Cell Biol.* **13**, 541–549.
- Pan, G., and Pei, D. (2012). Order from chaos: single cell reprogramming in two phases. *Cell Stem Cell* **11**, 445–447.
- Pera, M.F. (2011). Stem cells: The dark side of induced pluripotency. *Nature* **471**, 46–47.
- Phanstiel, D.H., Brumbaugh, J., Wenger, C.D., Tian, S., Probasco, M.D., Bailey, D.J., Swaney, D.L., Tervo, M.A., Bolin, J.M., Ruotti, V., et al. (2011). Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nat. Methods* **8**, 821–827.
- Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* **28**, 848–855.
- Raj, A., Rifkin, S.A., Andersen, E., and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913–918.
- Shah, S.N., Kerr, C., Cope, L., Zambidis, E., Liu, C., Hillion, J., Belton, A., Huso, D.L., and Resar, L.M. (2012). HMGA1 reprograms somatic cells into pluripotent stem cells by inducing stem cell transcriptional networks. *PLoS ONE* **7**, e48533.
- Song, Q., and Smith, A.D. (2011). Identifying dispersed epigenomic domains from ChIP-Seq data. *Bioinformatics* **27**, 870–871.
- Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., and Hochedlinger, K. (2010). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465**, 175–181.
- Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R.M., Chen, T., Ooi, S.S., Kim, S.Y., Bestor, T.H., Shioda, T., et al. (2012). Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nat. Genet.* **44**, 398–405, S1–S2.
- Stelzer, G., Inger, A., Olender, T., Ilyin-Stein, T., Dalah, I., Harel, A., Safran, M., and Lancet, D. (2009). GeneDecks: paralog hunting and gene-set distillation with GeneCards annotation. *OMICS* **13**, 477–487.
- Vafa, O., Wade, M., Kern, S., Beeche, M., Pandita, T.K., Hampton, G.M., and Wahl, G.M. (2002). c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. *Mol. Cell* **9**, 1031–1044.
- Viswanathan, S.R., Powers, J.T., Einhorn, W., Hoshida, Y., Ng, T.L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L.A., Lockhart, V.L., et al. (2009). Lin28 promotes transformation and is associated with advanced human malignancies. *Nat. Genet.* **41**, 843–848.
- Wang, Z., and Jaenisch, R. (2004). At most three ES cells contribute to the somatic lineages of chimeric mice and of mice produced by ES-tetraploid complementation. *Dev. Biol.* **275**, 192–201.
- Weinhold, B., Schratt, G., Arsenian, S., Berger, J., Kamino, K., Schwarz, H., Rütger, U., and Nordheim, A. (2000). Srf(-/-) ES cells display non-cell-autonomous impairment in mesodermal differentiation. *EMBO J.* **19**, 5835–5844.
- West, J.A., Viswanathan, S.R., Yabuuchi, A., Cunniff, K., Takeuchi, A., Park, I.H., Sero, J.E., Zhu, H., Perez-Atayde, A., Frazier, A.L., et al. (2009). A role for Lin28 in primordial germ-cell development and germ-cell malignancy. *Nature* **460**, 909–913.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319.
- Wu, T., Liu, Y., Wen, D., Tseng, Z., Tahmasian, M., Zhong, M., Rafii, S., Stadtfeld, M., Hochedlinger, K., and Xiao, A. (2014). Histone Variant H2A.X Deposition Pattern Serves as a Functional Epigenetic Mark for Distinguishing the Developmental Potentials of iPSCs. *Cell Stem Cell* **15**, this issue, 281–294.
- 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.
- Zhu, H., Shah, S., Shyh-Chang, N., Shinoda, G., Einhorn, W.S., Viswanathan, S.R., Takeuchi, A., Grasemann, C., Rinn, J.L., Lopez, M.F., et al. (2010). Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. *Nat. Genet.* **42**, 626–630.
- Zhu, H., Shyh-Chang, N., Segrè, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G., et al.; DIAGRAM Consortium; MAGIC Investigators (2011). The Lin28/let-7 axis regulates glucose metabolism. *Cell* **147**, 81–94.

# Prolonged Fasting Reduces IGF-1/PKA to Promote Hematopoietic-Stem-Cell-Based Regeneration and Reverse Immunosuppression

Chia-Wei Cheng,<sup>1</sup> Gregor B. Adams,<sup>2</sup> Laura Perin,<sup>3</sup> Min Wei,<sup>1</sup> Xiaoying Zhou,<sup>2</sup> Ben S. Lam,<sup>2</sup> Stefano Da Sacco,<sup>3</sup> Mario Mirisola,<sup>4</sup> David I. Quinn,<sup>5</sup> Tanya B. Dorff,<sup>5</sup> John J. Kopchick,<sup>6</sup> and Valter D. Longo<sup>1,2,7,\*</sup>

<sup>1</sup>Longevity Institute, School of Gerontology, Department of Biological Sciences, University of Southern California, 3715 McClintock Avenue, Los Angeles, CA 90089, USA

<sup>2</sup>Eli and Edythe Broad Center for Regenerative Medicine and Stem Cell Research at USC, Keck School of Medicine, University of Southern California, 1425 San Pablo Street, Los Angeles, CA 90033, USA

<sup>3</sup>Saban Research Institute, Division of Urology, Children's Hospital Los Angeles, University of Southern California, 4650 Sunset Boulevard, Los Angeles, CA 90027, USA

<sup>4</sup>Department of Medical Biotechnology and Forensics, University of Palermo, via Divisi 83, 90133 Palermo, Italy

<sup>5</sup>Translational Oncology Program, Kenneth J. Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 1441 Eastlake Avenue, Los Angeles, CA 90033, USA

<sup>6</sup>Department of Biomedical Sciences, Heritage College of Osteopathic Medicine, Ohio University, 228 Irvine Hall, Athens, OH 45701, USA

<sup>7</sup>IFOM, FIRC Institute of Molecular Oncology, Via Adamello, 16, 20139 Milano, Italy

\*Correspondence: vlongo@usc.edu

<http://dx.doi.org/10.1016/j.stem.2014.04.014>

## SUMMARY

Immune system defects are at the center of aging and a range of diseases. Here, we show that prolonged fasting reduces circulating IGF-1 levels and PKA activity in various cell populations, leading to signal transduction changes in long-term hematopoietic stem cells (LT-HSCs) and niche cells that promote stress resistance, self-renewal, and lineage-balanced regeneration. Multiple cycles of fasting abated the immunosuppression and mortality caused by chemotherapy and reversed age-dependent myeloid-bias in mice, in agreement with preliminary data on the protection of lymphocytes from chemotoxicity in fasting patients. The proregenerative effects of fasting on stem cells were recapitulated by deficiencies in either IGF-1 or PKA and blunted by exogenous IGF-1. These findings link the reduced levels of IGF-1 caused by fasting to PKA signaling and establish their crucial role in regulating hematopoietic stem cell protection, self-renewal, and regeneration.

## INTRODUCTION

Prolonged fasting (PF) lasting 48–120 hr reduces progrowth signaling and activates pathways that enhance cellular resistance to toxins in mice and humans (Fontana et al., 2010; Guevara-Aguirre et al., 2011; Holzenberger et al., 2003; Lee and Longo, 2011; Longo et al., 1997). The physiological changes caused by PF are much more pronounced than those caused by calorie restriction or fasting lasting 24 hours or less in part because of the requirement to fully switch to a fat- and ketone

bodies-based catabolism after glycogen reserves are depleted during PF (Longo and Mattson, 2014). Studies in mice indicate that PF can protect them from chemotoxicity by reducing circulating insulin-like growth factor-1 (IGF-1) (Lee et al., 2010; Raffaghello et al., 2008). A preliminary case series study also indicates that PF has the potential to ameliorate several side effects caused by chemotherapy in humans (Safdie et al., 2009). One of the side effects, myelosuppression, is often dose limiting in chemotherapy treatment, in part because damage to adult stem/progenitor cells impairs tissue repair and regeneration (Kofman et al., 2012; Mackall et al., 1994; van Tilburg et al., 2011; Williams et al., 2004). Despite the rising interest in nutrient-dependent changes in stem cell populations, little is known about how acute or periodic dietary interventions affect the hematopoietic system.

Hematopoietic stem and progenitor cells (HSPCs) residing in the adult bone marrow (BM) are part of the Lin<sup>-</sup>Sca-1<sup>+</sup>c-Kit<sup>+</sup> (LSK) population of cells, which include the self-renewing long-term and short-term hematopoietic stem cells (LSK-CD48<sup>-</sup>CD150<sup>+</sup>, LT-HSC, and LSK-CD48<sup>-</sup>CD150<sup>-</sup>, ST-HSC) and the multipotent progenitors (LSKCD48<sup>+</sup>, multipotent progenitor [MPP]) (Figure S1 available online) (Challen et al., 2009; Rathinam et al., 2011). Together, these cells are responsible for adult hematopoietic regeneration. In the heterogeneous hematopoietic stem cells (HSCs), several subtypes are identified as lymphoid- (Ly-HSCs), balanced HSC (Bala-HSC), and myeloid-HSCs (My-HSCs) according to their distinct mature blood cell outputs (Figure S1) (Benz et al., 2012; Challen et al., 2010; Muller-Sieburg et al., 2004). In both mice and humans, these HSC subtypes modulate hematopoietic lineage potential and play an important role in lineage-homeostasis during aging (Beerman et al., 2010; Challen et al., 2010; Cho et al., 2008; Pang et al., 2011). Here, we studied the role of multiple PF cycles on chemotherapy-induced and age-dependent immunosuppression and investigated how PF affects HSC self-renewal, the Ly-, My-, and Bala-HSC subtypes as well as their hematopoietic reconstitution outcomes.

## RESULTS

### Cycles of Prolonged Fasting Reduce Damage in Bone Marrow Stem and Progenitor Cells and Protect Mice against Chemotoxicity

Chemotherapy drugs cause immunosuppression by inducing DNA damage and cell death in both peripheral blood (PB) and bone marrow (BM), which often results in the long-term impairment of hematopoiesis (Bedford et al., 1984; Yahata et al., 2011). To test whether PF may protect the hematopoietic system against immunosuppressive toxicity, mice were fasted or fed an ad libitum diet (AL) and then challenged with cyclophosphamide (CP) for multiple cycles (Figure 1A) (Adams et al., 2007). In agreement with our previous results with etoposide and doxorubicin, we observed a major protective effect of cycles of 48 hr PF against CP-induced mortality (Figures 1B and S1A) (Raffaghello et al., 2008). The PF cycles also led to a decrease in the DNA damage caused by CP in leukocytes and BM cells (Figures 1C and S1B).

To determine whether PF can protect from chemotherapy-induced toxicity, we collected BM cells at the end of six cycles of CP or PF + CP treatments and measured apoptosis. Given that the HSPCs represent a minor fraction of the total BM, we further examined apoptosis in the subpopulations of these cells (i.e., LT-HSCs, ST-HSCs, and MPPs) by performing TUNEL assay. The results indicate that without affecting BM cellularity, PF diminished CP-induced apoptosis in HSPCs ( $p < 0.05$ , *t* test), particularly in ST-HSCs and MPPs (Figures 1D, S1C, and S1D). The PF-induced protection against CP-induced apoptosis was also confirmed by Annexin V binding assay for HSPCs (Figure S1E).

### Prolonged Fasting Cycles Promote Lineage-Balanced Hematopoietic Regeneration

To assess whether the protection of HSPCs improved the hematopoietic recovery, we compared the hematological profiles of CP and PF + CP mice at baseline (before CP treatments, after PF), at nadir (2–4 days after CP) and during the recovery phase (8–10 days after CP) for each cycle of chemotherapy. Multicycle CP treatments resulted in a major decline in white blood cell (WBC) counts (Figure 1E). In the control group, WBC suppression, especially the number of lymphocytes, persisted for more than 70 days (six cycles) (Figure 1E). PF reduced WBC counts independently of chemotherapy and did not prevent the CP-induced decrease in the number of WBCs (Figure 1E, time 0). However, the beneficial effect of PF was evident starting on cycle 4 (day 39) followed by the return of lymphocytes to normal levels after the fifth cycle (day 56) (Figure 1E). At the end of six cycles of treatment, mice in the PF group also showed normal or close to normal levels of lymphoid cells and normal ratios of lymphoid and myeloid cells (L/M) (Figure 1E, right panel). This recovery was observed at similar time points in three independent experiments ( $n = 20$ ).

To begin to determine whether PF cycles can potentially promote a similar effect in humans, we also analyzed the hematological profiles of cancer patients from a phase I clinical trial for the feasibility and safety of a 24–72 hr PF period in combination with chemotherapy. Although three different platinum-based drug combinations were used (Table S1), the results from a phase I clinical trial indicate that 72 but not 24 hr of PF in combination with chemotherapy were associated with normal lymphocyte counts and maintenance of a normal lineage balance in

WBCs (Figure 1F). These encouraging preliminary results will need to be expanded and confirmed in the ongoing phase II randomized phase of the clinical trial.

In agreement with the effect of PF on the recovery in WBC numbers and improvement in lymphoid/myeloid ratio, results of fluorescence-activated cell sorting (FACS) analyses for stem cell populations indicated an improved preservation of LT-HSCs and ST-HSCs and the enhanced resistance to the myeloid bias in the PF group after six cycles of CP treatment in mice (Figures 1G and 1H).

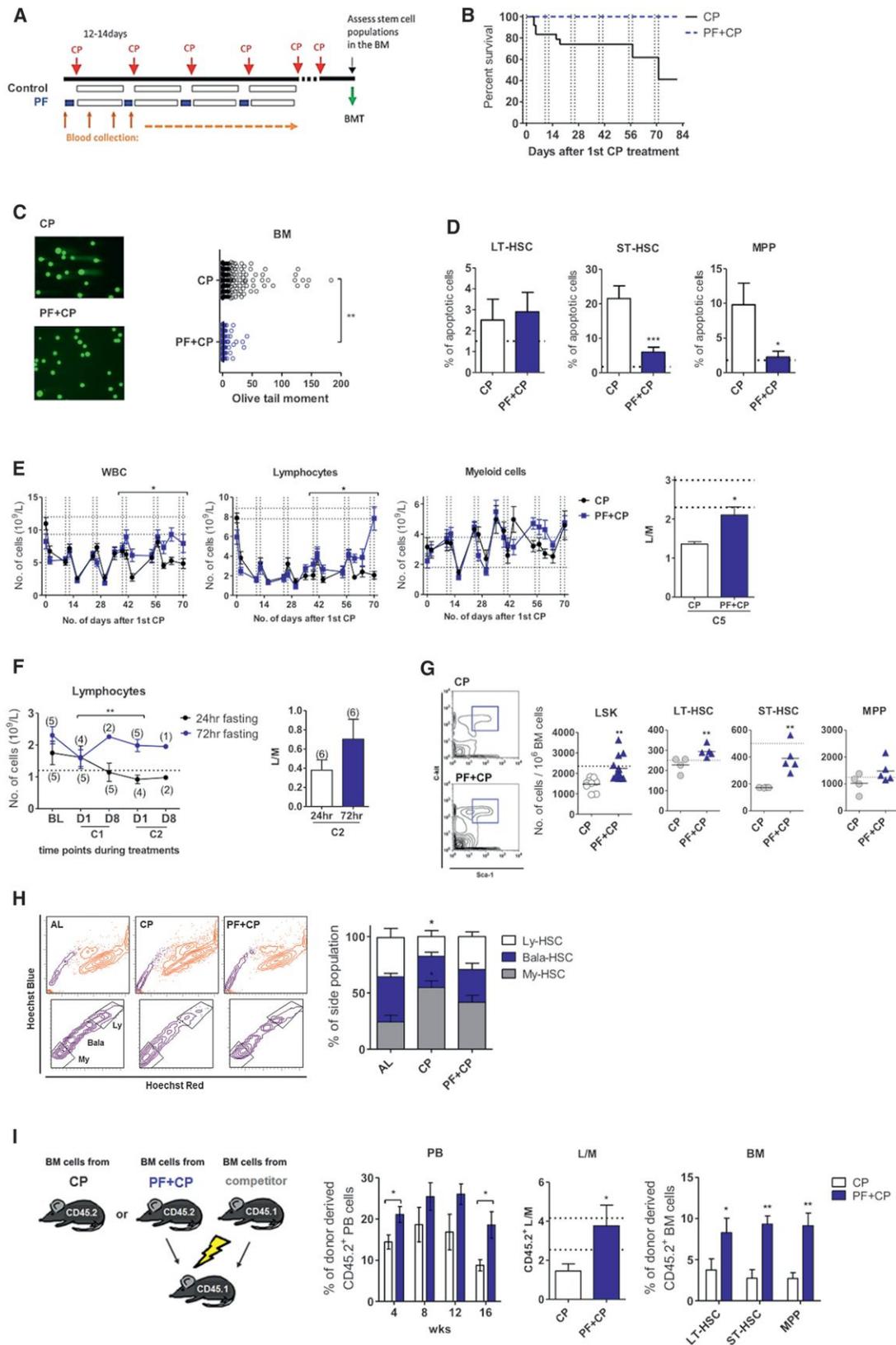
To assess whether the increased HSCs in BM from PF + CP mice can enhance hematopoietic regeneration, we collected BM cells from the CP- or PF + CP-treated mice and transplanted the same number of cells into the immunocompromised (irradiated) recipient mice. Results of this competitive repopulation assay indicate that, compared to the control group fed ad libitum, the BM cells from mice exposed to six cycles of CP treatment preceded by PF have higher regeneration capacity leading to efficient blood reconstitution with normal lymphoid/myeloid ratio (L/M), as evident from the improved engraftment in the blood and BM (Figures 1I, S1G, and S1H).

### Prolonged Fasting Cycles Regulate Stem Cells Independently of Chemotherapy and Help Reverse Immunosenescence

We tested whether the cycles of PF alone could also stimulate HSC self-renewal. Results using bromodeoxyuridine (BrdU) incorporation assays indicated an approximately 6-fold increase of newly generated (BrdU<sup>+</sup>) HSPCs (i.e., LT-HSC, ST-HSC, and MPP) in PF mice, which represents 93.7% of the total increase in HSPCs after PF cycles (Figure 2A). We found that the increase in LSK cell number is due mainly to an increase in LT-HSCs and ST-HSCs (Figure 2B). By contrast, the number of total BM cells and that of progenitors (i.e., MPP, multipotent progenitors; CLP, common lymphoid progenitors; CMP, common myeloid progenitor) was not increased by PF, and, in fact, the number of CMP was slightly decreased during PF (Figures 2C and S2A).

Results from cell-cycle analyses indicate that PF alone induced a major increase of S/G2/M phase LT-HSCs, ST-HSCs, and MPPs (Figure 2D). The significant induction in cell-cycle entry could explain at least part of the PF-induced increase in HSCs. In addition to the Ki67/Hoechst 33342 staining for cell-cycle analysis, the PF-induced self-renewal proliferation was confirmed by analysis using Pyronin Y/Hoechst 33342 staining (Figure S2B). On the other hand, results from the TUNEL assay indicate that apoptosis was barely detectable in any subpopulation of HSPCs from either AL-fed or PF mice when no chemotherapy treatment was applied. Apoptosis analysis using Annexin V and 7AAD indicate similar results (Figure S2C). Although PF alone reduces the apoptosis rate in ST-HSCs significantly, the small reduction (from 1.57% to 0.72%) in apoptosis/cellular death could only contribute to a very small portion of the PF-induced increase in HSCs and MPP (Figure 2E). However, because studies of HSCs have shown that induction of proliferation may sometimes be accompanied by an increase of apoptosis (Nakada et al., 2010; Tothova et al., 2007), it is important to note that this was not observed in PF-induced self-renewal proliferation.

Besides the increase in the number of HSCs and MPP, we also observed a PF-dependent alteration of lymphoid-,



**Figure 1. Prolonged Fasting Cycles Protect the Hematopoietic System and Reverse Chemotherapy-Induced Hematopoietic Suppression**  
 (A) Diagrammatic representation of the experimental procedure to analyze the effects of prolonged fasting (PF, 48 hr) during six cycles of cyclophosphamide chemotherapy (CP, 200 mg/kg, i.p.).

(legend continued on next page)

myeloid-biased, and balanced-HSCs ratio (Figures 2F, S2D, and S2E). Whereas most HSCs from young mice are balanced in lymphopoiesis and myelopoiesis, the majority of HSCs from elderly mice are myeloid biased (Beerman et al., 2010; Challen et al., 2010; Cho et al., 2008; Dykstra et al., 2007; Morita et al., 2010; Muller-Sieburg et al., 2004; Pang et al., 2011). We therefore investigated if PF cycles can correct this bias in aged mice. Results from 18-month-old mice indicate that eight cycles of PF could reverse the age-dependent myeloid bias in HSC subtypes and reverse the effect of aging on WBC number in whole blood (Figures 2F and 2G), similar to the changes observed in mice and possibly patients undergoing PF in combination with chemotherapy (Figures 1E, 1F, and 1H). Taken together, these results suggest that PF cycles can also stimulate the HSCs in a chemotherapy-independent manner, which leads to a lineage-balanced hematopoietic regeneration.

### Mimicking the Effects of Prolonged Fasting by Deficiency in GHR/IGF-1 Signaling Promotes Hematopoietic Recovery

We previously showed that PF reduces circulating IGF-1 levels and that IGF-1 deficiency is sufficient to protect mice against chemotherapy toxicity (Lee et al., 2010). To determine whether the improved hematopoietic regeneration caused by PF in mice can be replicated by IGF-1 deficiency, we studied the hematopoietic system in growth hormone receptor knockout (GHRKO) mice, which have very low circulating and BM IGF-1 levels (Al-Regaiey et al., 2005) (Figures 3A and S3A; Table S2). We found that CP-induced DNA damage measured by the comet assay in PB and BM cells of GHRKO mice was significantly reduced compared to that in cells from wild-type littermates (Figure 3B). Similar to what was observed in mice undergoing pre-chemo PF cycles, ST-HSCs of the GHRKO mice were protected from CP-induced apoptosis (Figure 3C). Also, the number of HSCs (i.e., LT-HSCs and ST-HSCs) preserved in the BM of GHRKO mice was higher than that in the wild-type littermates (Figure 3D). An improvement in hematopoietic recovery analogous to that caused by PF was also observed in GHRKO mice (Figure 3E).

We found that IGF-1 deficiency also caused the protective effects and the regenerative effects independently of chemotoxicity. Unlike PF mice, GHRKO mice did not have higher levels of

total HSPCs (Figure S3B). However, similarly to what we observed after PF cycles, the levels of HSCs (i.e., LT-HSCs plus ST-HSCs) was significantly higher in GHRKO mice compared to those in age- and sex-matched littermates, with increased cell-cycle entry but no detectable differences in apoptosis (Figures 3F, 3G, and S3C–S3E). Also, the age-dependent myeloid bias was not observed in the GHRKO mice (Figure 3H). These data suggest that the periodically reduced IGF-1 signaling caused by PF cycles may play a crucial role in the hematopoietic regeneration observed in mice.

### Prolonged Fasting Promotes Hematopoietic Regeneration in a IGF-1/PKA-Dependent Manner

To understand the molecular mechanism by which PF and GHR/IGF-1 deficiency promote hematopoietic recovery/regeneration, we reanalyzed two of our previously published microarray data sets and looked for genes whose expression significantly changed in response to PF with a focus on genes similarly affected by exposure of epithelial cells to IGF-1-deficient serum (Guevara-Aguirre et al., 2011; Kim and Volsky, 2005; Kirschner et al., 2009; Lee et al., 2012). In starved mice, the expression of the PKA catalytic subunit alpha (PKA $\alpha$ ) was significantly reduced in all tissues tested (Table S3). Similarly, IGF-1-deficient serum from growth hormone receptor-deficient (GHRD) human subjects caused changes in the expression of both positive and negative regulators of PKA consistent with an inhibition of its kinase activity (Table S4).

As PKA phosphorylates the cAMP response element-binding transcription factor (CREB) at Ser133, p-CREB is commonly used as an indicator of intracellular PKA activity (Gonzalez and Montminy, 1989). Using mouse embryonic fibroblasts devoid of the endogenous IGF-1 receptor (R<sup>-</sup> cells) and those overexpressing the human *IGF1R* (R<sup>+</sup> cell), we showed that CREB phosphorylation is positively regulated by IGF-1/IGF-1R in a PKA-dependent manner, confirming the link between IGF-1 and PKA/CREB signaling in mammalian cells (Figure 4A). IGF-1 receptor (IGF-1R) expression, which was higher in progenitor cells compared to LT-HSCs (Venkatraman et al., 2013), was not affected by PF (Figure S4A). Taken together, our *in vivo* results indicate that PF reduces PKA signaling in BM cells at least in part through reduced IGF-1 levels, but without affecting IGF-1R expression (Figure 4B).

(B) Survival curve with vertical dashed lines indicating the prechemo fasting period;  $p < 0.01$ , log-rank (Mantel-Cox) test;  $n = 20$  (ten male and ten female).

(C) DNA damage measurement (olive tail moment) in bone marrow (BM) cells (day 81, sixth recovery phase).

(D) Apoptosis measurement (TUNEL assay) in HSCs and MPP (day 81, sixth recovery phase).

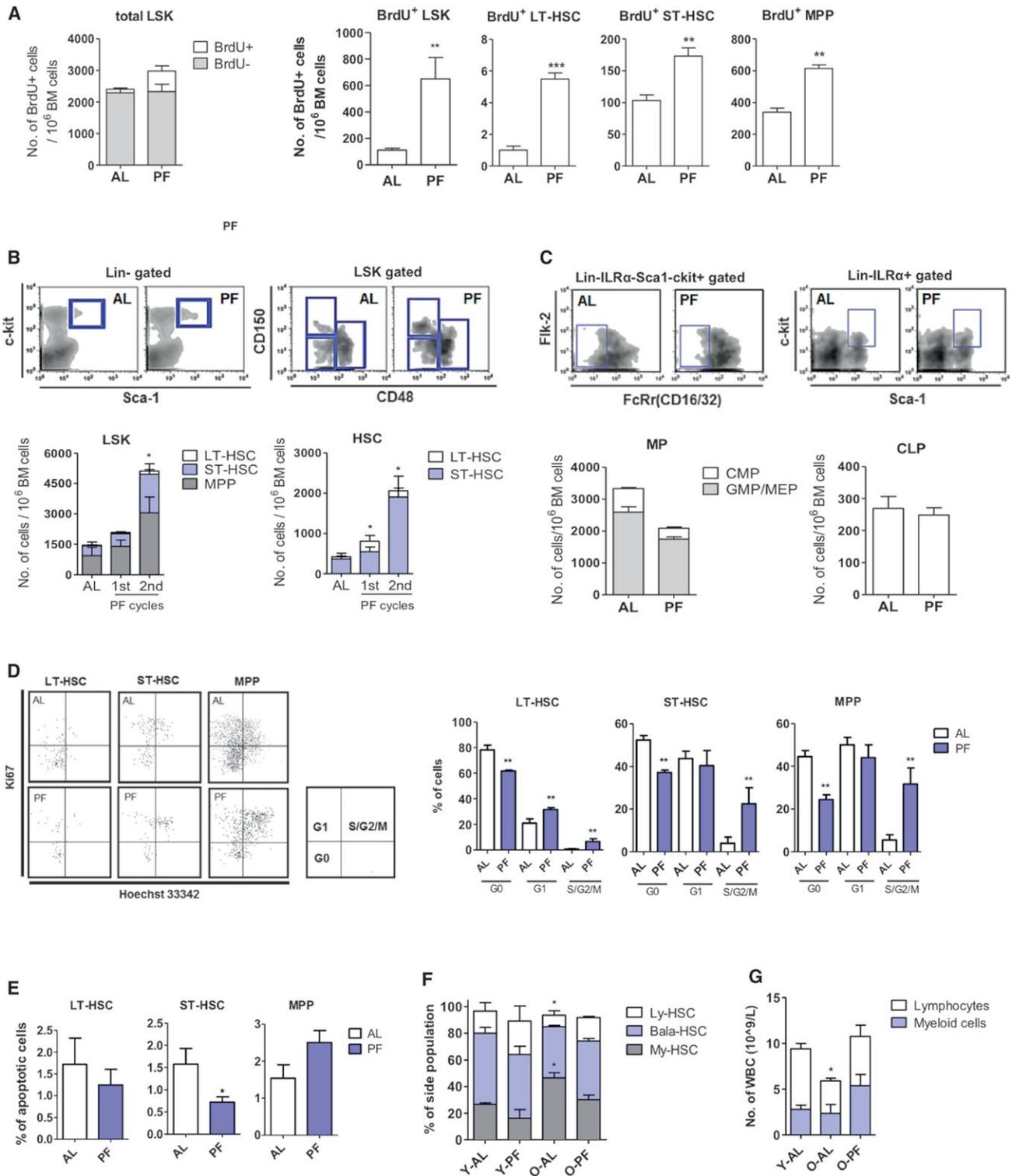
(E) Hematological profile of mice. Total white blood cell (WBC), lymphocyte counts, and lymphoid/myeloid ratio (L/M) in mice treated with six cycles of CP (200 mg/kg, *i.p.*). Each point represents the mean  $\pm$  SEM; horizontal dashed lines indicate the ranges of baseline values;  $*p < 0.05$ , two-way ANOVA, comparing CP versus PF + CP during the recovery phase,  $n = 12$  (six male and six female); L/M ratio of peripheral blood (PB) is defined as number of lymphocytes divided by number of myeloid cells (i.e., granulocytes and monocytes). See also Figures S1F and S1G.

(F) Hematological profile of human subjects. Lymphocyte counts and lymphoid/myeloid ratio (L/M) in patients undergoing two cycles (C1 and C2) of platinum-based doublet chemotherapy in combination with either 24 or 72 hr (48 before and 24 hr after chemo) prolonged fasting; D1 and D8 indicate the first day (before chemo) and eighth day after each chemotherapy cycle; each point represents the mean  $\pm$  SEM;  $**p < 0.01$ , two-way ANOVA; sample size is indicated in parentheses.

(G) FACS analysis of hematopoietic stem and progenitor cells (day 84, end of sixth cycle); horizontal dashed lines indicate the baseline value.

(H) Proportion of the lymphoid-biased (Ly-HSC), balanced (Bala-HSC), and of the myeloid-biased (My-HSC) hematopoietic stem cells. The markers used are lower side population of LSK (lower-SP<sup>LSK</sup>) for My-HSC, middle-SP<sup>LSK</sup> for Bala-HSC, and upper-SP<sup>LSK</sup> for Ly-HSC. The lower panels show a magnification of the SP population in the upper panels.  $*p < 0.05$ , one-way ANOVA comparing to AL.

(I) BM cells collected from mice treated with either CP or PF + CP were transplanted into the recipient mice. The chimerism of donor-derived cells in PB and that in BM was determined 16 weeks after primary BM transplantation. The ratio of lymphocytes to myeloid cells (L/M) in the reconstituted blood was also measured. For (G) and (I),  $n = 6$ –10 per group,  $*p < 0.05$ ,  $**p < 0.01$ , *t* test comparing the PF with the nonfasted control group both in combination with cyclophosphamide treatment. Error bars represent SEM.



**Figure 2. Prolonged Fasting Cycles Promote Chemotherapy-Independent Hematopoietic Regeneration**

Mice in the control group were fed ad libitum and those in the PF group were fasted for one or two cycles as indicated. n = 4–12 female mice per group. Error bars represent SEM.

(A) BrdU incorporation assay for LSK cells. Mice undergoing 24 + 48 hr prolonged fasting were injected (i.p.) with BrdU (0.1 mg/g, twice a day, for 2 days, starting after 24 hr of fasting.

(B) Number of long-term hematopoietic stem cells (LT-HSC), short-term hematopoietic stem cells (ST-HSC), and multipotent progenitors (MPP).

(C) Number of common lymphoid progenitors (CLP) and myeloid progenitors (MP)

(legend continued on next page)

To demonstrate that IGF-1 is a mediator of PF-dependent effects on HSCs, we tested whether exogenous IGF-1 can blunt the effect of PF on HSC number and PKA activity. Fasted mice were injected with IGF-1 (200  $\mu\text{g}/\text{kg}$ ) to reverse the reduction of IGF-1 during PF. Results indicate that IGF-1 administration significantly blunted PF-induced reduction of PKA/pCREB in the LSK population, particularly in HSCs (Figure 4C). It also blunted the PF-induced increase in HSCs but not in MPPs (Figure 4D; Table S4). We further investigated whether the induction in HSCs can lead to enhanced engraftment and whether this effect is IGF-1/PKA dependent. Results of competitive repopulation assays indicate the PF improved hematopoietic reconstitution in PB and in BM. This effect was blocked by exogenous IGF-1 (Figures 4E, 4F, S4B, and S4C). Results of secondary transplantation further confirmed the effects in long-term repopulation capacity (Figures 4G and 4H). Overall, these results strongly support a role for lower IGF-1 and the consequent reduced activity of PKA in PF-dependent stimulation of HSC self-renewal and the improvement in both short- and long-term hematopoietic repopulation capacities (Figures 4E–4H).

Because IGF-1R signaling and IGF-1 expression were both reduced in the BM stromal niche cells ( $\text{Lin}^- \text{CD45}^-$ ) from fasted mice (Figure 5A), we investigated whether the stromal niche could play a role in promoting PF-induced HSC self-renewal by reducing IGF-1 levels in the microenvironment (as previously shown in Figure 3A). To test this, LT-HSCs were purified ( $\text{CD45}^+ \text{LSK} \text{CD150}^+ \text{CD48}^-$ ) from mice on either PF or the control diet and then cross-exposed to the stromal niche cells ( $\text{CD45}^- \text{Lin}^-$  fraction) from mice on either PF or the ad lib diet using coculture systems (Figure 5B). Notably, LT-HSCs are unable to survive in the absence of niche cells, so the isolated LT-HSCs were not studied alone. Results indicate that the effect of PF on LT-HSC is sufficient to promote the self-renewal of LT-HSC and its capacity to generate ST-HSC and non-LSK progenitors ( $\text{Lin}^- \text{non-LSK}$ ) (Figure 5C, comparing A to B, C to D). Also, the PF-treated niche cells could increase the generation of ST-HSCs from ad lib diet LT-HSCs (comparing A to C) and increases further the ST-HSC number generated by PF-treated LT-HSCs (comparing B to D). These results confirm the role of LT-HSCs in mediating PF-dependent hematopoietic regeneration but also indicate that niche cells exposed to PF can contribute to the ST-HSC component of this regeneration in vitro.

### Reduction of IGF-1 or PKA Signaling Promotes HSC Self-Renewal

PKA has conserved proaging roles in yeast and mammals (Fabrizio et al., 2001; Rinaldi et al., 2010). In yeast, integration of an extra copy of the regulatory and inhibitory subunit of PKA, *BCY1* (*BCY1<sup>oe</sup>*) enhanced, whereas mutations in *BCY1* that activate PKA decreased, cellular resistance to  $\text{H}_2\text{O}_2$ -induced oxidative stress (Figure 6A) in agreement with our previous results with

*RAS2*- and adenylate cyclase-deficient mutants (Fabrizio et al., 2003; Fabrizio et al., 2001). In mammalian cells, it was confirmed by us and others that disruption of PKA signaling protects against stress (Figures S5A–S5C) (Yan et al., 2007).

The role of PKA in hematopoietic regeneration, however, is poorly understood. It is known that PKA negatively regulates *Foxo1* and positively regulates CREB and G9a (Chen et al., 2008; Gonzalez and Montminy, 1989; Lee et al., 2011; Yamamizu et al., 2012a). FoxOs maintain hematopoietic stress-resistance, self-renewal and lineage homeostasis (Tothova et al., 2007), whereas CREB and G9a promote hematopoietic lineage commitment and differentiation (Chen et al., 2012; Yamamizu et al., 2012b). We found that in PF mice, the reduction of IGF-1/pAkt and PKA/pCREB signaling was associated with an induction of *Foxo1* expression and a reduction of G9a (Figures 6B and S5E), but it did not affect the expressions of *Foxo3a* and *Foxo4* (Figures S5F–S5H). Also, the results indicated that the numbers of ST-HSC and MPP were significantly increased after treatment with PKA small interfering RNA (siRNA) as well as after treatment with IGF-1 siRNA (Figures 6C and S5D; Table S5), in agreement with the finding that inhibition of G9a increases primitive HSCs (Chen et al., 2012).

Given that inhibition of mTOR, another key effector of nutrient signaling, is known to enhance HSC self-renewal and maintenance autonomously and nonautonomously, we examined the crosstalk between mTOR and PKA in HSCs and MPPs (Chen et al., 2009; Huang et al., 2012). Ex vivo rapamycin (an mTOR inhibitor) treatment alone did not cause an induction in the number of HSCs as expected based upon previous studies with in vivo treatments (Figure 6C) (Nakada et al., 2010; Yilmaz et al., 2012). This could be due to the need for a longer period of mTOR inhibition to achieve HSC induction (Nakada et al., 2010). In fact, when cotreated with PKA siRNA, rapamycin caused an additional induction in ST-HSC and MPP, compared to that caused by PKA knockdown alone suggesting that diminished PKA signaling promotes the induction of HSCs, which can be further potentiated by mTOR inhibition in certain stem and progenitor cell subpopulations (Figure 6C). Notably, the double inhibition of PKA and mTOR resulted in the synergistic induction in ST-HSC and MPP but blunted the induction in LT-HSC caused by PKA knockdown alone, which is similar to what was caused by IGF-R knockdown (Figure 6C), and in agreement with the potential role for IGF-1 in the regulation of both PKA and mTOR in HSCs (Fontana et al., 2010; Longo and Fabrizio, 2002).

The BM cells treated with IGF-1R siRNA or PKA siRNA ex vivo (exBM) were further transplanted into the irradiated recipient mice to assess their hematopoietic reconstitution capacity. In agreement with the effects observed in PF mice, the PKA or IGF-1R-deficient BM cells caused a significant improvement in engraftment in PB and in BM compared to untreated BM cells

(D) Cell-cycle analysis for BM cells using Ki67 and Hoechst 33342.

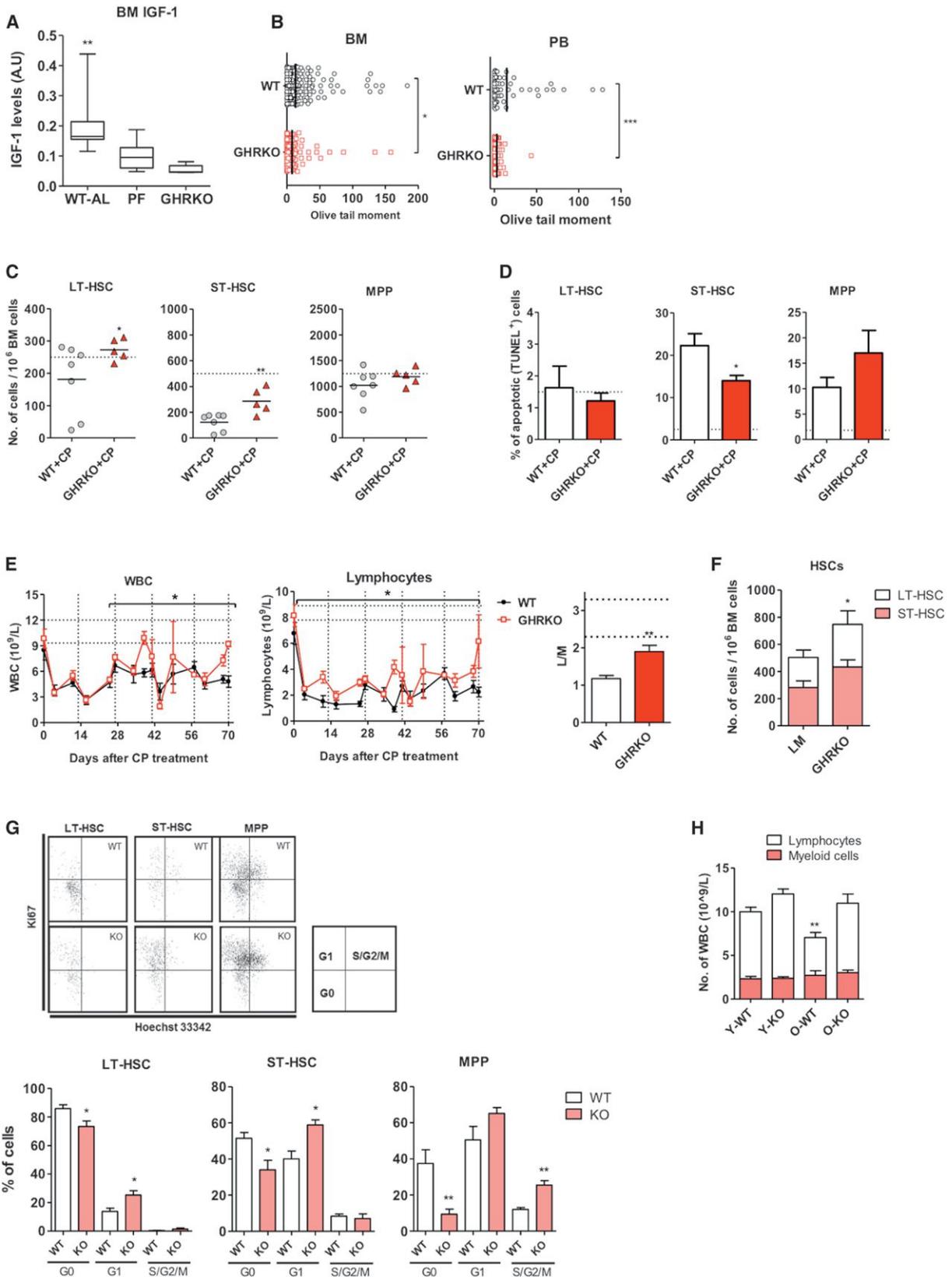
(E) Apoptosis analysis for BM cells using TUNEL assay.

For (A)–(E), \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.005$ , t test comparing the AL-fed controls.

(F) Proportion of the lymphoid-biased (Ly-HSC), balanced (Bala-HSC), and the myeloid-biased (My-HSC) hematopoietic stem cells. The markers used are lower side population of LSK (lower- $\text{SP}^{\text{LSK}}$ ) for My-HSC, middle- $\text{SP}^{\text{LSK}}$  for Bala-HSC, and upper- $\text{SP}^{\text{LSK}}$  for Ly-HSC.

(G) Number of lymphocytes and myeloid cells in young (6 months, 48 hr fasting) and old (18 months, eight cycles of fasting) mice.

For (F) and (G), \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.005$ , one-way ANOVA.



(legend on next page)

(Std) (Figure 6D). The long-term repopulation capacity was also confirmed by secondary transplantation (Figure S5I).

## DISCUSSION

When considering changes in gene expression and metabolism, as well as the levels of various hormones, PF promotes coordinated effects that would be difficult to achieve with any pharmacological or other dietary intervention. In yeast, the key changes responsible for the protective effects of starvation are the downregulation of the glucose-sensing Ras/adenylate cyclase/PKA and of the amino acid-sensing Tor/Sch9 (S6K) pathways (Figure 7A) (Fontana et al., 2010). When mutations in both pathways are combined, cells are extremely resistant to a wide variety of toxins and can live up to 5-fold longer than normal (Fabrizio et al., 2001; Kaeberlein et al., 2005; Kenyon, 2001; Longo and Finch, 2003; Wei et al., 2009). In mammals, mutations that cause deficiency in the GHR-IGF-1 axis promote a range of phenotypes that overlap with those in the highly protected yeast with deficiencies in nutrient signaling pathways including dwarfism, stress resistance, and longevity extension (Figure 7A) (Lee and Longo, 2011). In fact, cells from GHR/IGF-1-deficient mice are protected from multiple forms of stress (Brown-Borg et al., 2009; Salmon et al., 2005) and the IGF-1-deficient (LID) mice, with an over 70% reduction in circulating IGF-1, are resistant to several chemotherapy drugs (Lee et al., 2010). Here, we connect the GHR-IGF-1 and the PKA proaging pathways by showing that PKA functions downstream of IGF-1 to sensitize BM cells in agreement with results in yeast and with the previously established connection between IGF-1 and PKA in mammalian neuronal cells (Subramaniam et al., 2005).

However, the studies of growth-deficient yeast and mice could not have predicted the remarkable effect of PF cycles in promoting stem cell-based regeneration of the hematopoietic system. Calorie intake was previously shown to affect the balance of stem cell self-renewal and differentiation, which is important for somatic maintenance and long-term survival (Bondolfi et al., 2004; Chen et al., 2003; Ertl et al., 2008; Jasper and Jones, 2010; Rafalski and Brunet, 2011; Rando and Chang, 2012). In mice, chronic calorie restriction (CR) promotes the self-renewal of intestinal stem cells, muscle stem cell engraftment and neural regeneration, preserves the long-term regenerative capacity of HSCs, and prevents the decline of HSC number during aging in certain mouse strains (Lee et al., 2002; Bondolfi et al., 2004; Cerletti et al., 2012; Chen et al., 2003; Ertl et al., 2008; Rafalski

and Brunet, 2011; Yilmaz et al., 2012). Reduction of mTOR signaling has been implicated as one of the major molecular mechanisms responsible for the effects of CR on enhanced stem cell function (Huang et al., 2012; Rafalski and Brunet, 2011; Yilmaz et al., 2012). However, neither CR nor other dietary intervention had previously been shown to promote a coordinated effect leading to the regeneration and/or rejuvenation of a major portion of a system or organ.

Because during PF mammalian organisms minimize energy expenditure in part by rapidly reducing the size of a wide range of tissues, organs, and cellular populations including blood cells, the reversal of this effect during refeeding represents one of the most potent strategies to regenerate the hematopoietic and possibly other systems and organs in a coordinated manner. Here, we show that PF causes a major reduction in WBC number, followed, during refeeding, by a coordinated process able to regenerate this immune system deficiency by changes beginning during the fasting period, which include a major increase in LT-HSC and ST-HSC and redirection of the frequency of Ly-HSC/Bala-HSC/My-HSC leading to a lineage-balanced mode. In fact, we show that PF alone causes a 28% decrease WBC number, which is fully reversed after refeeding (Figures 7B and S2F). Even after WBCs are severely suppressed or damaged as a consequence of chemotherapy or aging, cycles of PF are able to restore the normal WBC number and lineage balance, suggesting that the organism may be able to exploit its ability to regenerate the hematopoietic system after periods of starvation, independently of the cause of the deficiency (Figure 7B).

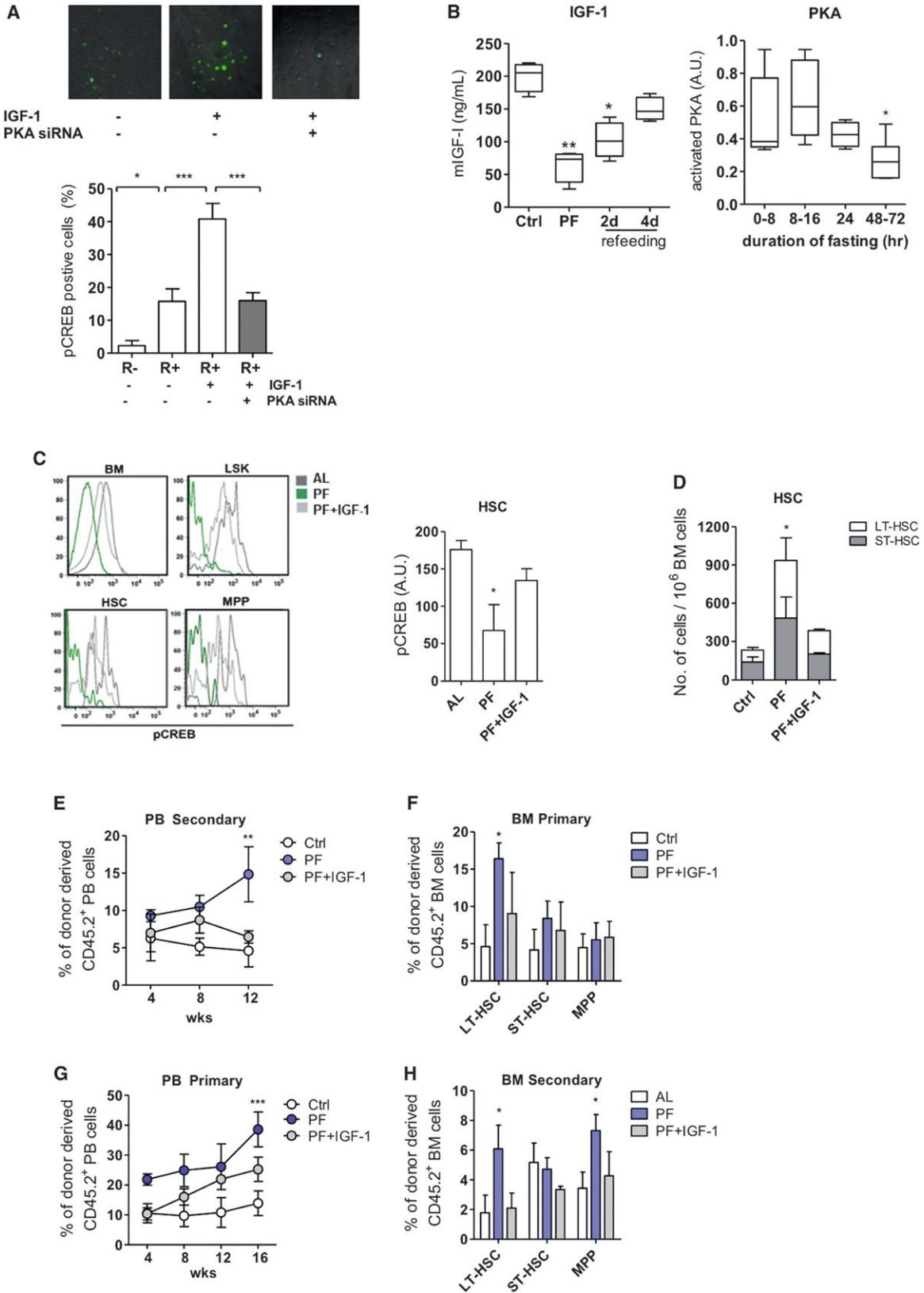
In agreement with our results, starvation protects germline stem cells (GSCs) and extends reproductive longevity in *C. elegans* through an adaptive energy shift toward the less committed cells (Angelo and Van Gilst, 2009). In contrast, short-term fasting ( $\leq 24$  hr) in *Drosophila* promotes the differentiation of hematopoietic progenitors to mature blood cells (Shim et al., 2012). It will be important to determine whether the coordinated regenerative changes observed during PF and refeeding may resemble at least in part the sophisticated program responsible for the generation of the hematopoietic system during development.

Recent studies revealed that HSCs rely heavily on the metabolic programs that prevent aerobic metabolism to maintain their quiescent state and self-renewal capacity (Ito et al., 2012; Takubo et al., 2013; Yu et al., 2013). In the case of PF, the energy metabolism is switched progressively from a carbohydrate-based to a fat- and ketone body-based catabolism, which could

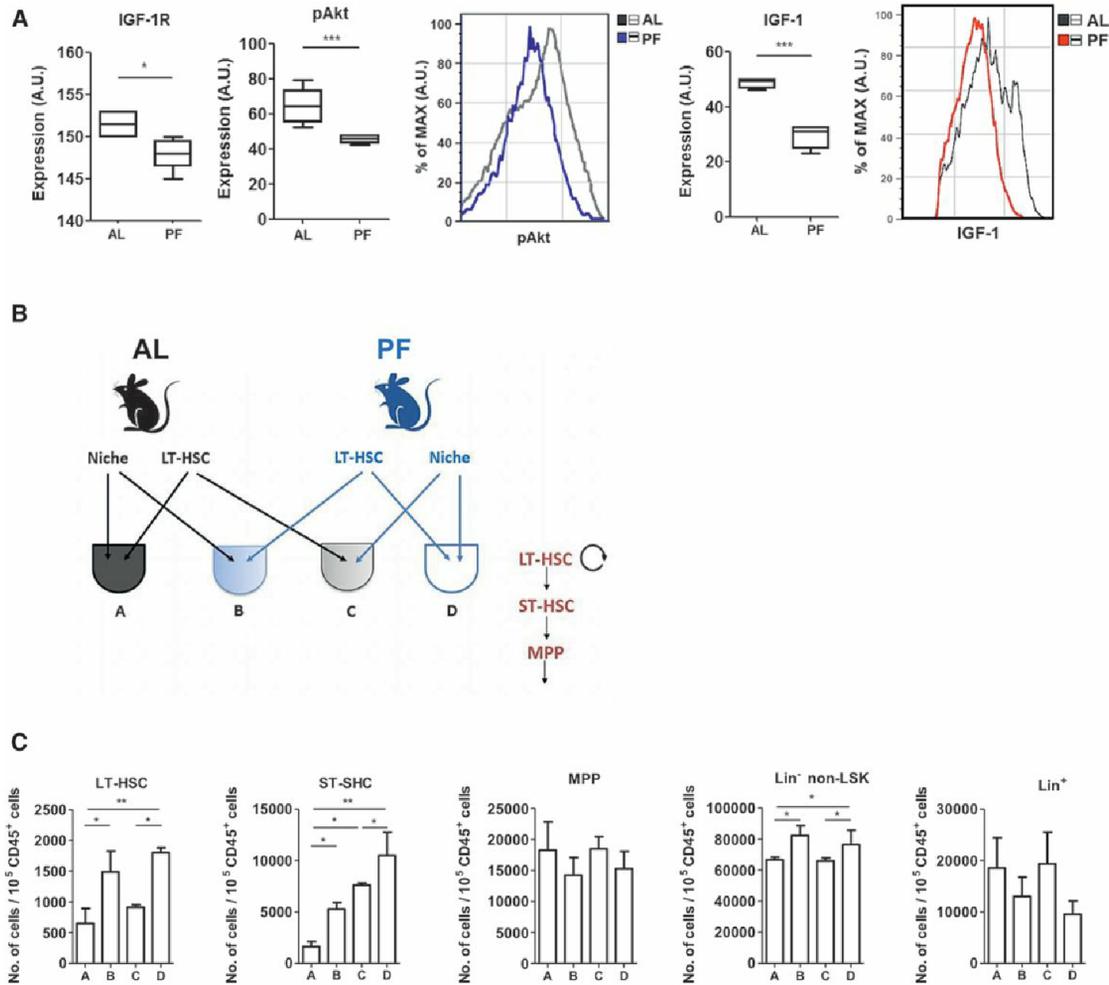
### Figure 3. Deficiency in GHR-IGF-1 Signaling Promotes Hematopoietic Regeneration in Both Chemo-Treated and Untreated Mice

Measurements were performed in GHRKO and their age-matched littermates, with or without treatment with six cycles of CP (200 mg/kg, i.p.). n = 4–8 female mice per group. Error bars represent SEM.

- (A) BM IGF-1 level in GHRKO mice and PF mice compared to wild-type mice fed ad libitum (WT-AL), \*p < 0.05, \*\*p < 0.01, one-way ANOVA.  
 (B) DNA damage measurement (olive tail moment) in BM cells and mononuclear peripheral blood cells (PB) from GHRKO and their littermates (WT) (day 81, sixth recovery phase).  
 (C) Apoptosis measurement (TUNEL assay) in hematopoietic stem and progenitor cells (day 81, sixth recovery phase).  
 (D) Number of hematopoietic stem and progenitor cells (day 84, end of sixth cycle); horizontal dashed lines indicate the chemo-free baseline value.  
 (E) Total white blood cell (WBC) and lymphocyte counts in PB of GHRKO mice and their littermates (WT); each point represents the mean  $\pm$  SEM; vertical dashed lines indicate CP treatments; horizontal dashed lines indicate baseline value; \*p < 0.05, two-way ANOVA for recovery phases; lymphoid/myeloid ratio (L/M) after six cycles of CP treatments. PB L/M ratio is defined as the number of lymphocytes divided by the number of myeloid cells (i.e., granulocytes and monocytes).  
 (F) Number of long-term hematopoietic stem cells (LT-HSC) and short-term hematopoietic stem cells (ST-HSC).  
 (G) Cell-cycle analysis using Ki67 and Hoechst 33342.  
 (H) Number of lymphocytes and myeloid cells in young (aged 6 months) and old (aged 18 months) mice.  
 For (B)–(D) and (F)–(H), \*p < 0.05, \*\*p < 0.01; t test comparing to the wild-type control.



(legend on next page)



**Figure 5. The Role of Stromal Niche in PF-Induced HSC Self-Renewal**

(A) Levels of the indicated proteins in BM stromal niche cells (Lin-CD45<sup>-</sup>).

(B) Diagrammatic representation of the coculture experiment.

(C) Number of CD45<sup>+</sup> progenies generated by the purified LT-HSCs exposed to the indicated niche cells. Cells were grown in the contacting coculture system for 3 days and then analyzed by FACS. Error bars represent SEM.

\*p < 0.05, \*\*p < 0.01, and \*\*\*p < 0.005, t test for (A) and one-way ANOVA for (C).

contribute to HSC self-renewal, in agreement with findings that fatty acid oxidation promotes HSC asymmetric self-renewal over the symmetric commitment (Ito et al., 2012).

PKA is known to promote lineage specification of HSC through CREB and G9a (Chen et al., 2012; Yamamizu et al., 2012b). As inhibition of G9a has been a key strategy to promote reprogramming (Huangfu et al., 2008; Shi et al., 2008), the PF-induced downregulation of G9a shown here may redirect cell fate through a similar process causing the induction in HSCs, analogously to that caused by G9a inhibition (Figure 5B)(Chen

et al., 2012). Recent studies also indicate that PKA can directly phosphorylate and negatively regulate FoxO1 (Chen et al., 2008; Lee et al., 2011), which has a profound role in stem cell stress resistance, self-renewal and pluripotency maintenance (Tothova et al., 2007; Zhang et al., 2011). Whereas PKA is implicated in stem cell differentiation, our study suggests that cycles of PF downregulate IGF-1 and PKA to promote stem cell self-renewal.

A therapeutic challenge of hematopoietic regeneration is to stimulate stem cell production for immediate tissue repair while

**Figure 4. Prolonged Fasting Promotes IGF-1/PKA-Dependent Hematopoietic Regeneration**

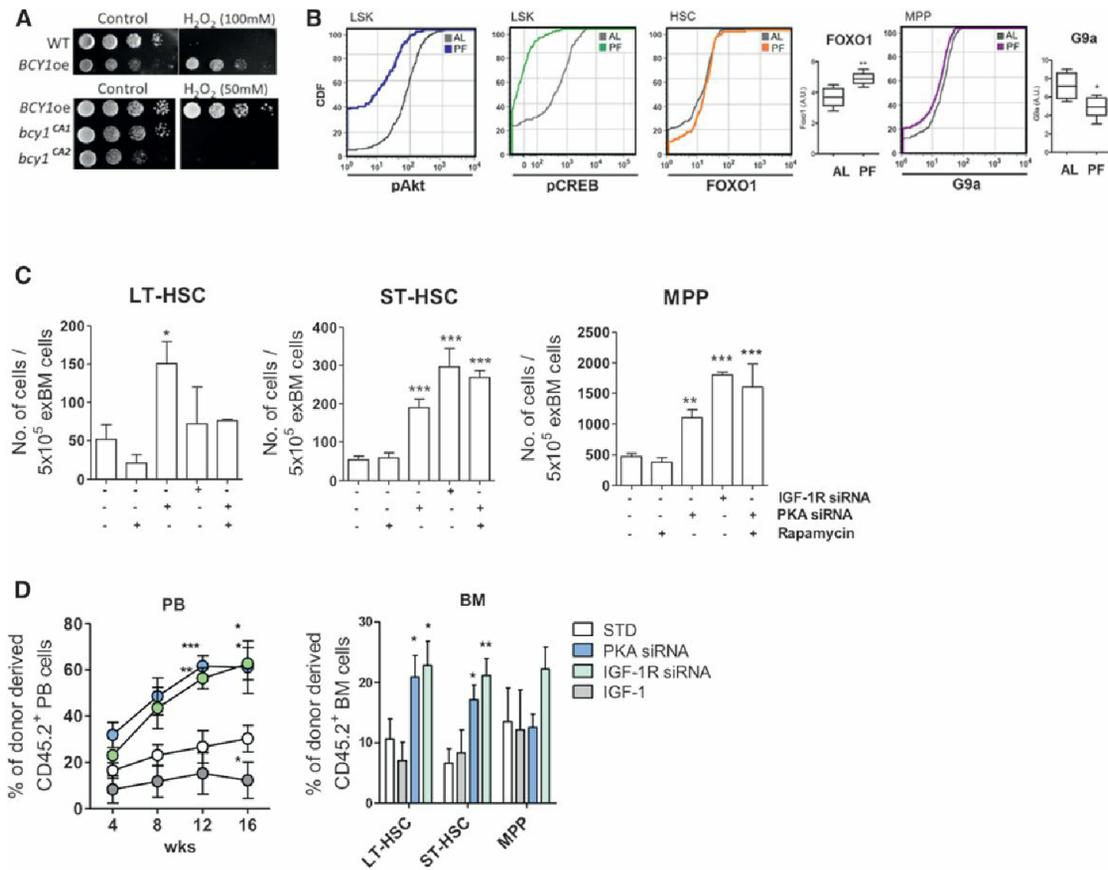
(A) PKA-dependent phosphorylation of CREB visualized by ICC in mouse embryonic fibroblast (MEFs) devoid of endogenous IGF-1R (R<sup>-</sup> cells) or overexpressing human IGF1R (R<sup>+</sup> cells). R<sup>+</sup> cells were treated with IGF-1 and compared to cells transfected with PKAC $\alpha$  siRNA.

(B) Prolonged fasting (PF) reduces both circulating IGF-1 levels and PKA activity in BM cells in mice.

(C and D) IGF-1 injection blunted the PF-induced (C) reduction of PKA/pCREB (D) increase in hematopoietic stem cells.

(E-H) The chimerism of donor-derived cells in PB and that in the BM was determined 16 weeks after primary and secondary BM transplantation.

n = 4-8 female mice per group, \*p < 0.05, \*\*p < 0.01, and \*\*\*p < 0.005, one-way ANOVA. Error bars represent SEM.



**Figure 6. Reduction of IGF-1-PKA Signaling Promotes Hematopoietic Stem Cell Self-Renewal**

(A) Yeast cells (DBY746 background) overexpressing BCY1 (BCY1oe), which reduces PKA activity, or cells carrying mutations that activate PKA activity (bcy1CA1 and bcy1CA2) were grown in SDC for 3 days and treated with H<sub>2</sub>O<sub>2</sub> (50 or 100 mM) for 30 min at 30°C. Cells were serially diluted and plated onto YPD plates. (B) PKA-regulated self-renewal pathways in PF mice. The levels of phosphorylation or expression of intracellular proteins in the indicated cellular populations and expression of indicated genes in total BM cells. BM cells were collected from mice with or without 48 hr starvation (AL and PF). n = 4 female mice per group, \*\*p < 0.01, \*p < 0.05, t test.

(C) Number of hematopoietic stem cells (per 5 × 10<sup>5</sup> total BM) and progenitor cells (LT-HSC, ST-HSC, and MPP) under the indicated treatments. \*p < 0.05, \*\*p < 0.01, and \*\*\*p < 0.005, one-way ANOVA. See also Figures S7F and S7G.

BM cells treated with PKA siRNA, IGF-1R siRNA or IGF-1 (versus nontreated cells) were transplanted into immunocompromised-recipient mice.

(D) The engraftment in PB was measured at indicated time point after primary transplantation and the engraftment in BM was measured at the end of the 16 weeks after primary transplantation. n = 4–8 female mice per group, \*p < 0.05, \*\*p < 0.01, and \*\*\*p < 0.005, one-way ANOVA. Error bars represent SEM.

avoiding stem cell depletion under stress (Pang et al., 2011). Our results indicate that cycles of an extreme dietary intervention represent a powerful mean to modulate key regulators of cellular protection and tissue regeneration but also provide a potential therapy to reverse or alleviate the immunosuppression or immunosenescence caused by chemotherapy treatment and aging, respectively, and possibly by a variety of diseases affecting the hematopoietic and immune systems and other systems and organs. The clinical data shown here provide preliminary results supporting the possibility that these effects can also be translated into effective clinical applications.

## EXPERIMENTAL PROCEDURES

### Mice

C57BL/6J mice (Jackson Laboratory) were used in this study. Mice are either fasted for 48 hr or fed ad libitum before chemotherapy treatment. Cyclophos-

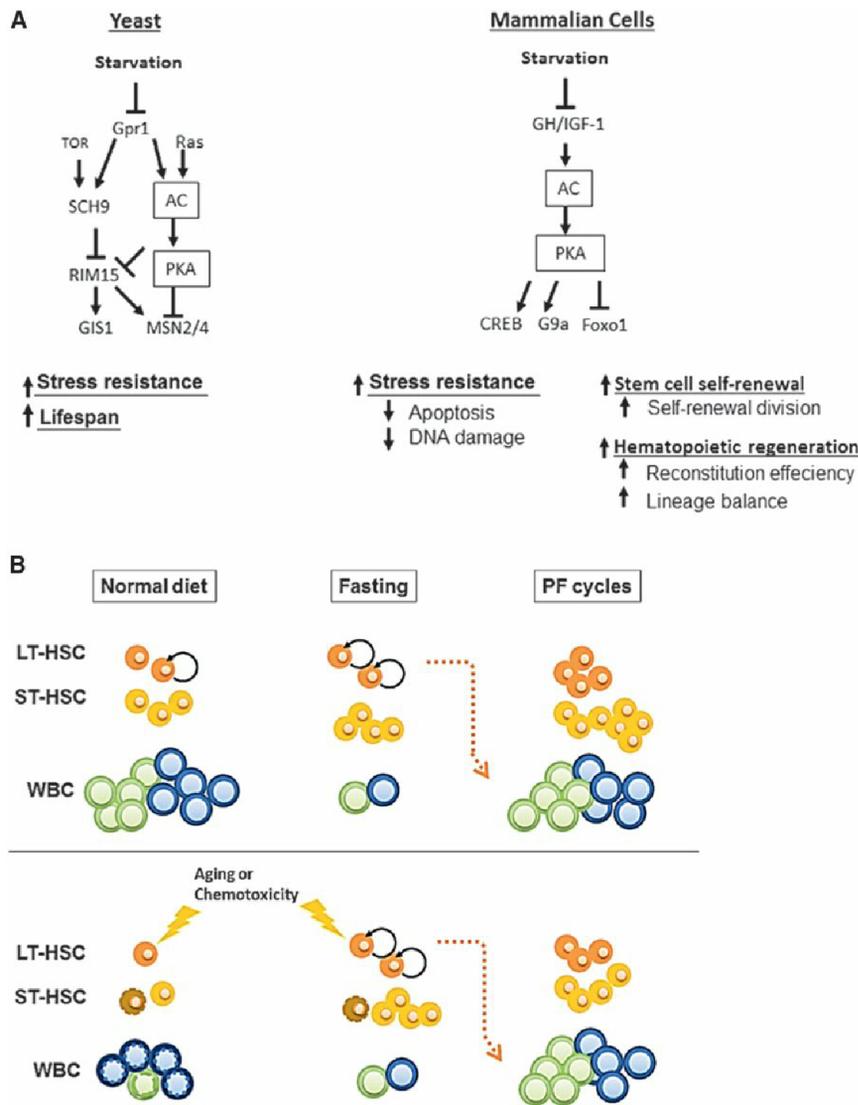
phamide (CP) was administered intraperitoneally (i.p.) at the dose of 200 mg/kg every 12–14 days (six cycles total). IGF-1 was injected (i.p.) at the dose of 100 μg/kg, twice a day. Six- to 8-week-old B6.SJL mice (Taconic) were used as recipient mice in the competitive repopulation assay. Genotyping for GHRKO mice was performed as shown in Figure S3A. All animal experiments were done in accordance with the USC Institutional Animal Care and Use Committee and NIH guidelines.

### Comet Assay

DNA damage (including single-stranded DNA and double-stranded DNA breaks) in freshly collected blood and bone marrow (BM) cells was assessed by CometAssay (Trevigen) with a Nikon Eclipse TE300 fluorescent microscope and analyzed with the Comet Score (TriTek, v.1.5). One hundred to two hundred cells were scored per experimental sample.

### Complete Blood Count

Peripheral blood (PB) was collected via tail bleeds into heparinized microhematocrit capillary tubes (Fisher Scientific) and analyzed using BC-2800 Auto Hematology Analyzer (Mindray). CBC profiles from clinical



**Figure 7. PF Reduces IGF-1/PKA to Promote Lineage-Balanced Hematopoietic Regeneration**

(A) A simplified model for a partially conserved nutrient signaling PKA pathway in yeast and mammalian cells. Arrows show activating actions, and horizontal bars indicate inhibitory actions. GH, growth hormone; AC, adenylate cyclase; PKA, protein kinase A; CREB, cAMP response element-binding protein; Foxo1, Forkhead box protein O1; G9a, H3 Lys-9 methyltransferase.

(B) A simplified model for PF-induced effects on WBC and HSCs. Fasting causes a major reduction in WBCs followed by their replenishment after re-feeding, based on effects on HSCs self-renewal resulting in increased progenitor and immune cells. These effects of PF can result in reversal of chemotherapy-based immunosuppression but also in the rejuvenation of the immune cell profile in old mice.

was collected from tail vein. Fifty to one hundred microliters of blood was diluted 1:1 with PBS and incubated with anti-CD45.1, anti-CD45.2 antibodies, and anti-CD11b (BD Biosciences). Analysis was performed with BD FACS diva on LSR II.

**BrdU Incorporation**

For detecting cell genesis, mice were injected (i.p.) with the filter sterilized *BrdU* 2.0% solution (Sigma) at 0.1 mg/g body weight in PBS, twice a day, for 2 days, starting after 24 hr of prolonged fasting (PF mice). BM cells were collected and stained with anti-BrdU combining with the plasma membrane marker antibodies as mentioned above and analyzed on BD FACS diva on an LSR II, according to the manufacturer's protocol (BD Biosciences).

**Oxidative Stress Assay for Yeast**

Day 3 cells were diluted to an OD<sub>600</sub> of 1 in K-phosphate buffer (pH 6) and treated with 50 or 100 mM hydrogen peroxide for 30 min. Serial dilutions of

untreated and treated cells were spotted onto YPD plates and incubated at 30°C for 2–3 days.

untreated and treated cells were spotted onto YPD plates and incubated at 30°C for 2–3 days.

**Cell Culture and Treatments**

Cell lines and primary cells used in this study were cultured at 37°C and 5% CO<sub>2</sub>. Mouse embryonic fibroblast with overexpressed human IGF1R (R<sup>+</sup> cells) were derived from IGF1R knockout mice (obtained from Dr. Baserga) and cultured in DMEM/F12 supplemented with 10% fetal bovine serum (FBS). Cells were seeded at 80% (R<sup>+</sup> cells) or 50% (exBM or hAFSCs) confluence for IGF-1R and PKAC $\alpha$  siRNA transfection (100 nM, with 1% X-tremeGENE transfection reagents, Roche) and/or rapamycin treatment (5 nM), and the inhibition efficiencies of the target proteins are shown in Table S5. The IGF-1 induction (10 nM, 15 min) was performed at 24 hr after standard incubation. CREB phosphorylation was measured by immunocytochemistry (ICC) with the pCREB-AF488 antibody (cell signaling, 1:200, overnight at 4°C). Explanted BM cells, isolated HSCs and BM stromal cells were incubated with alpha-MEM + 10% FBS. Cell contents were analysis by FACS as described above.

**Statistical Analysis**

The significance of the differences in mouse survival curves was determined by Log-rank (Mantel-Cox). Unless otherwise indicated in figure legends, data are presented as means  $\pm$  SEM. Student's t tests for two groups and

**Competitive Repopulation Assay**

BM collection and transplantation were performed as previously described (Adams et al., 2007). Briefly, BM cells were collected from mice (C57B/6J) treated with six cycle CP. BM cells (2.5  $\times$  10<sup>5</sup>) from CP-treated mice were mixed with an equal number of those from a wild-type competitor mouse (B6.SJL) and injected into recipient B6.SJL, lethally irradiated 24 hr previously with 10 Gy of radiation. The relative contribution of engraftment from the different cell sources was assessed by flow cytometry of the PB with CD45.2 (C57B/6) and CD45.1 (B6.SJL) antigens.

**FACS Analysis**

FACS analyses for LT-HSCs (LSK-CD48<sup>+</sup>CD150<sup>+</sup>), ST-HSCs (LSK-CD48<sup>-</sup>CD150<sup>-</sup>), and MPPs (LSK CD48<sup>+</sup>CD150<sup>-</sup>) in BM were performed as previously described (Figure S1) (Adams et al., 2007; Challen et al., 2010). Freshly harvested BM cells were stained with lineage, stem, and progenitor markers, followed by Annexin-V/7-AAD staining and TUNEL assay for apoptosis analysis or stained with PY/Hoechst 33342 or Ki67/Hoechst 33342 for cell-cycle analysis. For competitive repopulation analysis, PB

ANOVA for multiple groups were used to assess statistical significance (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.005$ ).

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and five tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2014.04.014>.

#### ACKNOWLEDGMENTS

We thank Dr. Lora Barsky (USC Flow Cytometry Core Facility) for assistance in the FACS analysis, and Dr. R. Baserga (Thomas Jefferson University) for providing R<sup>+</sup> and R<sup>-</sup> cells. This study was funded in part by NIH/NIA grants AG20642, AG025135, and P01 AG034906. V.D.L. has equity interest in L-Nutra, a company that develops medical food.

Received: October 9, 2012

Revised: December 5, 2013

Accepted: April 16, 2014

Published: June 5, 2014

#### REFERENCES

- Adams, G.B., Martin, R.P., Alley, I.R., Chabner, K.T., Cohen, K.S., Calvi, L.M., Kronenberg, H.M., and Scadden, D.T. (2007). Therapeutic targeting of a stem cell niche. *Nat. Biotechnol.* **25**, 238–243.
- Al-Regaiey, K.A., Masternak, M.M., Bonkowski, M., Sun, L., and Bartke, A. (2005). Long-lived growth hormone receptor knockout mice: interaction of reduced insulin-like growth factor *i*/insulin signaling and caloric restriction. *Endocrinology* **146**, 851–860.
- Angelo, G., and Van Gilst, M.R. (2009). Starvation protects germline stem cells and extends reproductive longevity in *C. elegans*. *Science* **326**, 954–958.
- Bedford, P., Berger, M.R., Eisenbrand, G., and Schmähl, D. (1984). The level of DNA interstrand crosslinking in bone marrow parallels the extent of myelosuppression in mice treated with four chloroethylnitrosoureas. *J. Cancer Res. Clin. Oncol.* **108**, 141–147.
- Beerman, I., Bhattacharya, D., Zandi, S., Sigvardsson, M., Weissman, I.L., Bryder, D., and Rossi, D.J. (2010). Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc. Natl. Acad. Sci. USA* **107**, 5465–5470.
- Benz, C., Copley, M.R., Kent, D.G., Wohrer, S., Cortes, A., Aghaepour, N., Ma, E., Mader, H., Rowe, K., Day, C., et al. (2012). Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs. *Cell Stem Cell* **10**, 273–283.
- Bondolfi, L., Ermini, F., Long, J.M., Ingram, D.K., and Jucker, M. (2004). Impact of age and caloric restriction on neurogenesis in the dentate gyrus of C57BL/6 mice. *Neurobiol. Aging* **25**, 333–340.
- Brown-Borg, H.M., Rakoczy, S.G., Sharma, S., and Bartke, A. (2009). Long-lived growth hormone receptor knockout mice: potential mechanisms of altered stress resistance. *Exp. Gerontol.* **44**, 10–19.
- Cerletti, M., Jang, Y.C., Finley, L.W., Haigis, M.C., and Wagers, A.J. (2012). Short-term calorie restriction enhances skeletal muscle stem cell function. *Cell Stem Cell* **10**, 515–519.
- Challen, G.A., Boles, N., Lin, K.K.Y., and Goodell, M.A. (2009). Mouse hematopoietic stem cell identification and analysis. *Cytometry A* **75**, 14–24.
- Challen, G.A., Boles, N.C., Chambers, S.M., and Goodell, M.A. (2010). Distinct hematopoietic stem cell subtypes are differentially regulated by TGF- $\beta$ 1. *Cell Stem Cell* **6**, 265–278.
- Chen, J., Astle, C.M., and Harrison, D.E. (2003). Hematopoietic senescence is postponed and hematopoietic stem cell function is enhanced by dietary restriction. *Exp. Hematol.* **31**, 1097–1103.
- Chen, H., Lin, A.S., Li, Y., Reiter, C.E., Ver, M.R., and Quon, M.J. (2008). Dehydroepiandrosterone stimulates phosphorylation of FoxO1 in vascular endothelial cells via phosphatidylinositol 3-kinase- and protein kinase A-dependent signaling pathways to regulate ET-1 synthesis and secretion. *J. Biol. Chem.* **283**, 29228–29238.
- Chen, C., Liu, Y., Liu, Y., and Zheng, P. (2009). mTOR regulation and therapeutic rejuvenation of aging hematopoietic stem cells. *Sci. Signal.* **2**, ra75.
- Chen, X., Skutt-Kakaria, K., Davison, J., Ou, Y.L., Choi, E., Malik, P., Loeb, K., Wood, B., Georges, G., Torok-Storb, B., and Paddison, P.J. (2012). G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment. *Genes Dev.* **26**, 2499–2511.
- Cho, R.H., Sieburg, H.B., and Muller-Sieburg, C.E. (2008). A new mechanism for the aging of hematopoietic stem cells: aging changes the clonal composition of the stem cell compartment but not individual stem cells. *Blood* **111**, 5553–5561.
- Dykstra, B., Kent, D., Bowie, M., McCaffrey, L., Hamilton, M., Lyons, K., Lee, S.J., Brinkman, R., and Eaves, C. (2007). Long-term propagation of distinct hematopoietic differentiation programs in vivo. *Cell Stem Cell* **1**, 218–229.
- Ertl, R.P., Chen, J., Astle, C.M., Duffy, T.M., and Harrison, D.E. (2008). Effects of dietary restriction on hematopoietic stem-cell aging are genetically regulated. *Blood* **111**, 1709–1716.
- Fabrizio, P., Pozza, F., Pletcher, S.D., Gendron, C.M., and Longo, V.D. (2001). Regulation of longevity and stress resistance by Sch9 in yeast. *Science* **292**, 288–290.
- Fabrizio, P., Liou, L.L., Moy, V.N., Diaspro, A., Valentine, J.S., Gralla, E.B., and Longo, V.D. (2003). SOD2 functions downstream of Sch9 to extend longevity in yeast. *Genetics* **163**, 35–46.
- Fontana, L., Partridge, L., and Longo, V.D. (2010). Extending healthy life span—from yeast to humans. *Science* **328**, 321–326.
- Gonzalez, G.A., and Montminy, M.R. (1989). Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* **59**, 675–680.
- Guevara-Aguirre, J., Balasubramanian, P., Guevara-Aguirre, M., Wei, M., Madia, F., Cheng, C.W., Hwang, D., Martin-Montalvo, A., Saavedra, J., Ingles, S., et al. (2011). Growth hormone receptor deficiency is associated with a major reduction in pro-aging signaling, cancer, and diabetes in humans. *Sci. Transl. Med.* **3**, 70ra13.
- Holzenberger, M., Dupont, J., Ducos, B., Leneuve, P., Gélouën, A., Even, P.C., Cervera, P., and Le Bouc, Y. (2003). IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* **421**, 182–187.
- Huang, J., Nguyen-McCarty, M., Hexner, E.O., Danet-Desnoyers, G., and Klein, P.S. (2012). Maintenance of hematopoietic stem cells through regulation of Wnt and mTOR pathways. *Nat. Med.* **18**, 1778–1785.
- Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* **26**, 795–797.
- Ito, K., Carracedo, A., Weiss, D., Arai, F., Ala, U., Avigan, D.E., Schafer, Z.T., Evans, R.M., Suda, T., Lee, C.H., and Pandolfi, P.P. (2012). A PML–PPAR- $\delta$  pathway for fatty acid oxidation regulates hematopoietic stem cell maintenance. *Nat. Med.* **18**, 1350–1358.
- Jasper, H., and Jones, D.L. (2010). Metabolic regulation of stem cell behavior and implications for aging. *Cell Metab.* **12**, 561–565.
- Kaerberlein, M., Powers, R.W., 3rd, Steffen, K.K., Westman, E.A., Hu, D., Dang, N., Kerr, E.O., Kirkland, K.T., Fields, S., and Kennedy, B.K. (2005). Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. *Science* **310**, 1193–1196.
- Kenyon, C. (2001). A conserved regulatory system for aging. *Cell* **105**, 165–168.
- Kim, S.Y., and Volsky, D.J. (2005). PAGE: parametric analysis of gene set enrichment. *BMC Bioinformatics* **6**, 144.
- Kirschner, L.S., Yin, Z., Jones, G.N., and Mahoney, E. (2009). Mouse models of altered protein kinase A signaling. *Endocr. Relat. Cancer* **16**, 773–793.
- Kofman, A.E., McGraw, M.R., and Payne, C.J. (2012). Rapamycin increases oxidative stress response gene expression in adult stem cells. *Aging (Albany, N.Y. Online)* **4**, 279–289.

- Lee, J., Seroogy, K.B., and Mattson, M.P. (2002). Dietary restriction enhances neurotrophin expression and neurogenesis in the hippocampus of adult mice. *J. Neurochem* 80, 539–547.
- Lee, C., and Longo, V.D. (2011). Fasting vs dietary restriction in cellular protection and cancer treatment: from model organisms to patients. *Oncogene* 30, 3305–3316.
- Lee, C., Safdie, F.M., Raffaghello, L., Wei, M., Madia, F., Parrella, E., Hwang, D., Cohen, P., Bianchi, G., and Longo, V.D. (2010). Reduced levels of IGF-1 mediate differential protection of normal and cancer cells in response to fasting and improve chemotherapeutic index. *Cancer Res.* 70, 1564–1572.
- Lee, J.W., Chen, H., Pullikotil, P., and Quon, M.J. (2011). Protein kinase A- $\alpha$  directly phosphorylates FoxO1 in vascular endothelial cells to regulate expression of vascular cellular adhesion molecule-1 mRNA. *J. Biol. Chem.* 286, 6423–6432.
- Lee, C., Raffaghello, L., Brandhorst, S., Safdie, F.M., Bianchi, G., Martin-Montalvo, A., Pistoia, V., Wei, M., Hwang, S., Merlino, A., et al. (2012). Fasting cycles retard growth of tumors and sensitize a range of cancer cell types to chemotherapy. *Sci. Transl. Med.* 4, 124ra127.
- Longo, V.D., and Fabrizio, P. (2002). Regulation of longevity and stress resistance: a molecular strategy conserved from yeast to humans? *Cell. Mol. Life Sci.* 59, 903–908.
- Longo, V.D., and Finch, C.E. (2003). Evolutionary medicine: from dwarf model systems to healthy centenarians? *Science* 299, 1342–1346.
- Longo, V.D., and Mattson, M.P. (2014). Fasting: molecular mechanisms and clinical applications. *Cell Metab.* 19, 181–192.
- Longo, V.D., Ellerby, L.M., Bredesen, D.E., Valentine, J.S., and Gralla, E.B. (1997). Human Bcl-2 reverses survival defects in yeast lacking superoxide dismutase and delays death of wild-type yeast. *J. Cell Biol.* 137, 1581–1588.
- Mackall, C.L., Fleisher, T.A., Brown, M.R., Magrath, I.T., Shad, A.T., Horowitz, M.E., Wexler, L.H., Adde, M.A., McClure, L.L., and Gress, R.E. (1994). Lymphocyte depletion during treatment with intensive chemotherapy for cancer. *Blood* 84, 2221–2228.
- Morita, Y., Ema, H., and Nakauchi, H. (2010). Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J. Exp. Med.* 207, 1173–1182.
- Muller-Sieburg, C.E., Cho, R.H., Karlsson, L., Huang, J.F., and Sieburg, H.B. (2004). Myeloid-biased hematopoietic stem cells have extensive self-renewal capacity but generate diminished lymphoid progeny with impaired IL-7 responsiveness. *Blood* 103, 4111–4118.
- Nakada, D., Saunders, T.L., and Morrison, S.J. (2010). Lkb1 regulates cell cycle and energy metabolism in hematopoietic stem cells. *Nature* 468, 653–658.
- Pang, W.W., Price, E.A., Sahoo, D., Beerman, I., Maloney, W.J., Rossi, D.J., Schrier, S.L., and Weissman, I.L. (2011). Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc. Natl. Acad. Sci. USA* 108, 20012–20017.
- Rafalski, V.A., and Brunet, A. (2011). Energy metabolism in adult neural stem cell fate. *Prog. Neurobiol.* 93, 182–203.
- Raffaghello, L., Lee, C., Safdie, F.M., Wei, M., Madia, F., Bianchi, G., and Longo, V.D. (2008). Starvation-dependent differential stress resistance protects normal but not cancer cells against high-dose chemotherapy. *Proc. Natl. Acad. Sci. USA* 105, 8215–8220.
- Rando, T.A., and Chang, H.Y. (2012). Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock. *Cell* 148, 46–57.
- Rathinam, C., Matesic, L.E., and Flavell, R.A. (2011). The E3 ligase Itch is a negative regulator of the homeostasis and function of hematopoietic stem cells. *Nat. Immunol.* 12, 399–407.
- Rinaldi, J., Wu, J., Yang, J., Ralston, C.Y., Sankaran, B., Moreno, S., and Taylor, S.S. (2010). Structure of yeast regulatory subunit: a glimpse into the evolution of PKA signaling. *Structure* 18, 1471–1482.
- Safdie, F.M., Dorff, T., Quinn, D., Fontana, L., Wei, M., Lee, C., Cohen, P., and Longo, V.D. (2009). Fasting and cancer treatment in humans: A case series report. *Aging (Albany, N.Y. Online)* 1, 988–1007.
- Salmon, A.B., Murakami, S., Bartke, A., Kopchick, J., Yasumura, K., and Miller, R.A. (2005). Fibroblast cell lines from young adult mice of long-lived mutant strains are resistant to multiple forms of stress. *Am. J. Physiol. Endocrinol. Metab.* 289, E23–E29.
- Shi, Y., Do, J.T., Despons, C., Hahm, H.S., Schöler, H.R., and Ding, S. (2008). A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2, 525–528.
- Shim, J., Mukherjee, T., and Banerjee, U. (2012). Direct sensing of systemic and nutritional signals by haematopoietic progenitors in *Drosophila*. *Nat. Cell Biol.* 14, 394–400.
- Subramaniam, S., Shahani, N., Strelau, J., Laliberté, C., Brandt, R., Kaplan, D., and Unsicker, K. (2005). Insulin-like growth factor 1 inhibits extracellular signal-regulated kinase to promote neuronal survival via the phosphatidylinositol 3-kinase/protein kinase A/c-Raf pathway. *J. Neurosci.* 25, 2838–2852.
- Takubo, K., Nagamatsu, G., Kobayashi, C.I., Nakamura-Ishizu, A., Kobayashi, H., Ikeda, E., Goda, N., Rahimi, Y., Johnson, R.S., Soga, T., et al. (2013). Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. *Cell Stem Cell* 12, 49–61.
- Tothova, Z., Kollipara, R., Huntly, B.J., Lee, B.H., Castrillon, D.H., Cullen, D.E., McDowell, E.P., Lazo-Kallanian, S., Williams, I.R., Sears, C., et al. (2007). FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* 128, 325–339.
- van Tilburg, C.M., van Gent, R., Bierings, M.B., Otto, S.A., Sanders, E.A.M., Nibbelke, E.E., Gaiser, J.F., Janssens-Korpela, P.L., Wolfs, T.F.W., Bloem, A.C., et al. (2011). Immune reconstitution in children following chemotherapy for haematological malignancies: a long-term follow-up. *Br. J. Haematol.* 152, 201–210.
- Venkatraman, A., He, X.C., Thorvaldsen, J.L., Sugimura, R., Perry, J.M., Tao, F., Zhao, M., Christenson, M.K., Sanchez, R., Yu, J.Y., et al. (2013). Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature* 500, 345–349.
- Wei, M., Fabrizio, P., Madia, F., Hu, J., Ge, H., Li, L.M., and Longo, V.D. (2009). Tor1/Sch9-regulated carbon source substitution is as effective as calorie restriction in life span extension. *PLoS Genet.* 5, e1000467.
- Williams, M.D., Braun, L.A., Cooper, L.M., Johnston, J., Weiss, R.V., Qualy, R.L., and Linde-Zwirble, W. (2004). Hospitalized cancer patients with severe sepsis: analysis of incidence, mortality, and associated costs of care. *Crit. Care* 8, R291–R298.
- Yahata, T., Takanashi, 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.
- Yamamizu, K., Fujihara, M., Tachibana, M., Katayama, S., Takahashi, A., Hara, E., Imai, H., Shinkai, Y., and Yamashita, J.K. (2012a). Protein kinase A determines timing of early differentiation through epigenetic regulation with G9a. *Cell Stem Cell* 10, 759–770.
- Yamamizu, K., Matsunaga, T., Katayama, S., Kataoka, H., Takayama, N., Eto, K., Nishikawa, S., and Yamashita, J.K. (2012b). PKA/CREB signaling triggers initiation of endothelial and hematopoietic cell differentiation via Etv2 induction. *Stem Cells* 30, 687–696.
- Yan, L., Vatner, D.E., O'Connor, J.P., Ivessa, A., Ge, H., Chen, W., Hirotani, S., Ishikawa, Y., Sadoshima, J., and Vatner, S.F. (2007). Type 5 adenylyl cyclase disruption increases longevity and protects against stress. *Cell* 130, 247–258.
- Yilmaz, O.H., Katajisto, P., Lamming, D.W., Gültekin, Y., Bauer-Rowe, K.E., Sengupta, S., Birsoy, K., Dursun, A., Yilmaz, V.O., Selig, M., et al. (2012). mTORC1 in the Paneth cell niche couples intestinal stem-cell function to calorie intake. *Nature* 486, 490–495.
- Yu, W.M., Liu, X., Shen, J., Jovanovic, O., Pohl, E.E., Gerson, S.L., Finkel, T., Broxmeyer, H.E., and Qu, C.K. (2013). Metabolic regulation by the mitochondrial phosphatase PTPMT1 is required for hematopoietic stem cell differentiation. *Cell Stem Cell* 12, 62–74.
- Zhang, X., Yalcin, S., Lee, D.F., Yeh, T.Y., Lee, S.M., Su, J., Mungamuri, S.K., Rimmelé, P., Kennedy, M., Sellers, R., et al. (2011). FOXO1 is an essential regulator of pluripotency in human embryonic stem cells. *Nat. Cell Biol.* 13, 1092–1099.



# LOST IN PAPERS?

## Find your way with Cell Press Reviews

Whether you're looking for answers to fundamental questions, brushing up on a new subject or looking for a synthesis on a hot topic, turn to Cell Press Reviews. We publish reviews in a vast array of disciplines spanning the life sciences.

Propel your research forward, faster with Cell Press Reviews. Our insightful, authoritative reviews—published in our primary research and Trends journals—go beyond synthesis and offer a point of view.

[www.cell.com/reviews](http://www.cell.com/reviews)

**CellPress**  
Your work is our life

# An iCRISPR Platform for Rapid, Multiplexable, and Inducible Genome Editing in Human Pluripotent Stem Cells

Federico González,<sup>1,4</sup> Zengrong Zhu,<sup>1,4</sup> Zhong-Dong Shi,<sup>1,4</sup> Katherine Lelli,<sup>1</sup> Nipun Verma,<sup>1,2</sup> Qing V. Li,<sup>1,3</sup> and Danwei Huangfu<sup>1,\*</sup>

<sup>1</sup>Developmental Biology Program, Sloan-Kettering Institute, 1275 York Avenue, New York, NY 10065, USA

<sup>2</sup>Weill Graduate School of Medical Sciences at Cornell University/The Rockefeller University/Sloan-Kettering Institute Tri-Institutional M.D.-Ph.D. Program, 1300 York Avenue, New York, NY 10065, USA

<sup>3</sup>Louis V. Gerstner Jr. Graduate School of Biomedical Sciences, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA

<sup>4</sup>Co-first author

\*Correspondence: [huangfud@mskcc.org](mailto:huangfud@mskcc.org)

<http://dx.doi.org/10.1016/j.stem.2014.05.018>

## SUMMARY

Human pluripotent stem cells (hPSCs) offer a unique platform for elucidating the genes and molecular pathways that underlie complex traits and diseases. To realize this promise, methods for rapid and controllable genetic manipulations are urgently needed. By combining two newly developed gene-editing tools, the TALEN and CRISPR/Cas systems, we have developed a genome-engineering platform in hPSCs, which we named iCRISPR. iCRISPR enabled rapid and highly efficient generation of biallelic knockout hPSCs for loss-of-function studies, as well as homozygous knockin hPSCs with specific nucleotide alterations for precise modeling of disease conditions. We further demonstrate efficient one-step generation of double- and triple-gene knockout hPSC lines, as well as stage-specific inducible gene knockout during hPSC differentiation. Thus the iCRISPR platform is uniquely suited for dissection of complex genetic interactions and pleiotropic gene functions in human disease studies and has the potential to support high-throughput genetic analysis in hPSCs.

## INTRODUCTION

The identification and functional validation of sequence variants affecting diverse human traits, including disease susceptibility, is key to understanding human biology and disease mechanisms. Advances in next-generation sequencing and genome-wide association studies have led to the rapid discovery of numerous disease-associated sequence variants in recent years. To functionally validate this increasing number of disease-associated mutations, an ideal platform should not only closely recapitulate their genomic, cellular, and human-specific contexts (Cooper and Shendure, 2011), but also offer superior speed and capacity to meet the growing demand.

Human pluripotent stem cells (hPSCs), including human embryonic stem cells (hESCs) and the closely related human induced pluripotent stem cells (hiPSCs), offer a promising solution to meet these challenges (Zhu and Huangfu, 2013). First, hPSCs have unlimited self-renewal capacity, providing a renewable source of experimental cells suitable for rapid, large-scale analyses. Second, they have the potential to generate all adult cell types, including rare or inaccessible human cell populations, providing a unique platform to recapitulate the cellular and human-specific contexts required for disease studies. Yet, to fulfill this potential, it is of paramount importance to develop methods for rapid, efficient, and controllable genetic manipulation in hPSCs. Unfortunately, while classic gene-targeting technology via homologous recombination in mouse ESCs (mESCs) has proven a powerful tool to dissect gene function (Capecchi, 2005; Thomas and Capecchi, 1986; Thomas et al., 1986), this approach has been extremely inefficient when applied to hPSCs (Hockemeyer and Jaenisch, 2010).

Recently, with the advent of programmable site-specific nucleases, genome engineering has become a much easier task in a wide range of organisms and cultured cell types including hPSCs (Joung and Sander, 2013; Ran et al., 2013b; Urnov et al., 2010). Acting as “DNA scissors,” they induce double strand breaks (DSBs) at desired genomic loci, triggering the endogenous DNA repair machinery. Processing of DSBs by the error-prone nonhomologous end-joining (NHEJ) pathway leads to small insertions and deletions (Indels) useful for generating loss-of-function mutations, whereas error-free homology directed repair (HDR) enables targeted integration of exogenously provided DNA sequences for introducing precise nucleotide (nt) alterations or knockin reporters.

Transcription activator-like effector nucleases (TALENs) and the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems have emerged as powerful and versatile site-specific nucleases for genome modification in a variety of model systems. TALENs are typically designed as pairs to bind the genomic sequences flanking the target site. Each TALEN arm consists of a programmable, sequence-specific TALE DNA-binding domain linked to a nonspecific DNA cleavage domain derived from the bacterial restriction endonuclease FokI (Cermak et al., 2011; Miller

et al., 2011). Recent studies have also successfully adapted the prokaryotic type II CRISPR/Cas system for genome editing in eukaryotic systems (Cong et al., 2013; DiCarlo et al., 2013; Gratz et al., 2013; Hwang et al., 2013; Mali et al., 2013b). The type II CRISPR/Cas system requires two components: the DNA endonuclease Cas9 protein for DNA cleavage and a variable CRISPR RNA (crRNA) and *trans*-activating crRNA (tracrRNA) duplex for DNA target recognition (Jinek et al., 2012). Binding of crRNA/tracrRNA to the target sequence via Watson-Crick base pairing directs Cas9 to any genomic locus of interest for site-specific DNA cleavage. The CRISPR/Cas system has now been further improved for use in mammalian systems through Cas9 codon optimization and replacement of the crRNA/tracrRNA duplex with a single chimeric guide RNA (gRNA) (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013b).

While programmable site-specific nucleases have significantly improved our capacity to genetically modify hPSCs (Ding et al., 2013a, 2013b; Hockemeyer et al., 2009, 2011; Hsu et al., 2013; Mali et al., 2013b; Soldner et al., 2011), none of the available methods has achieved multiplexed gene targeting or inducible gene knockout in hPSCs. These two features are crucial for interrogating complex genetic interactions and pleiotropic gene functions, which are often difficult to study using animal models due to the involvement of multiple alleles and random segregation of these alleles through breeding.

To achieve rapid, multiplexable, and inducible genome editing in hPSCs, we have developed a genome-engineering platform. Through TALEN-mediated gene targeting, we created a panel of hPSC lines for robust, doxycycline-inducible expression of Cas9 (referred to as iCas9 hPSCs). By transfecting iCas9 hPSCs with gRNAs targeting different genes, we have generated biallelic knockout hPSC lines for six individual genes with high efficiency (~20% to ~60%). This highly efficient platform enabled us to generate double- and triple-gene knockout hPSC lines in a single step with up to ~10% efficiency. Moreover, cotransfection of gRNAs with a single-stranded DNA (ssDNA) HDR template yielded homozygous knockin clones at a rate of up to ~10%, allowing efficient and scarless introduction of defined nucleotide modifications in different hPSC lines for disease modeling. Finally, we achieved stage-specific inducible gene knockout during hPSC differentiation. This versatile new platform, which we called “iCRISPR,” allows rapid generation of mutant hPSCs for analysis of complex disease phenotypes in isogenic backgrounds and could be easily scalable for high-throughput genetic analysis.

## RESULTS

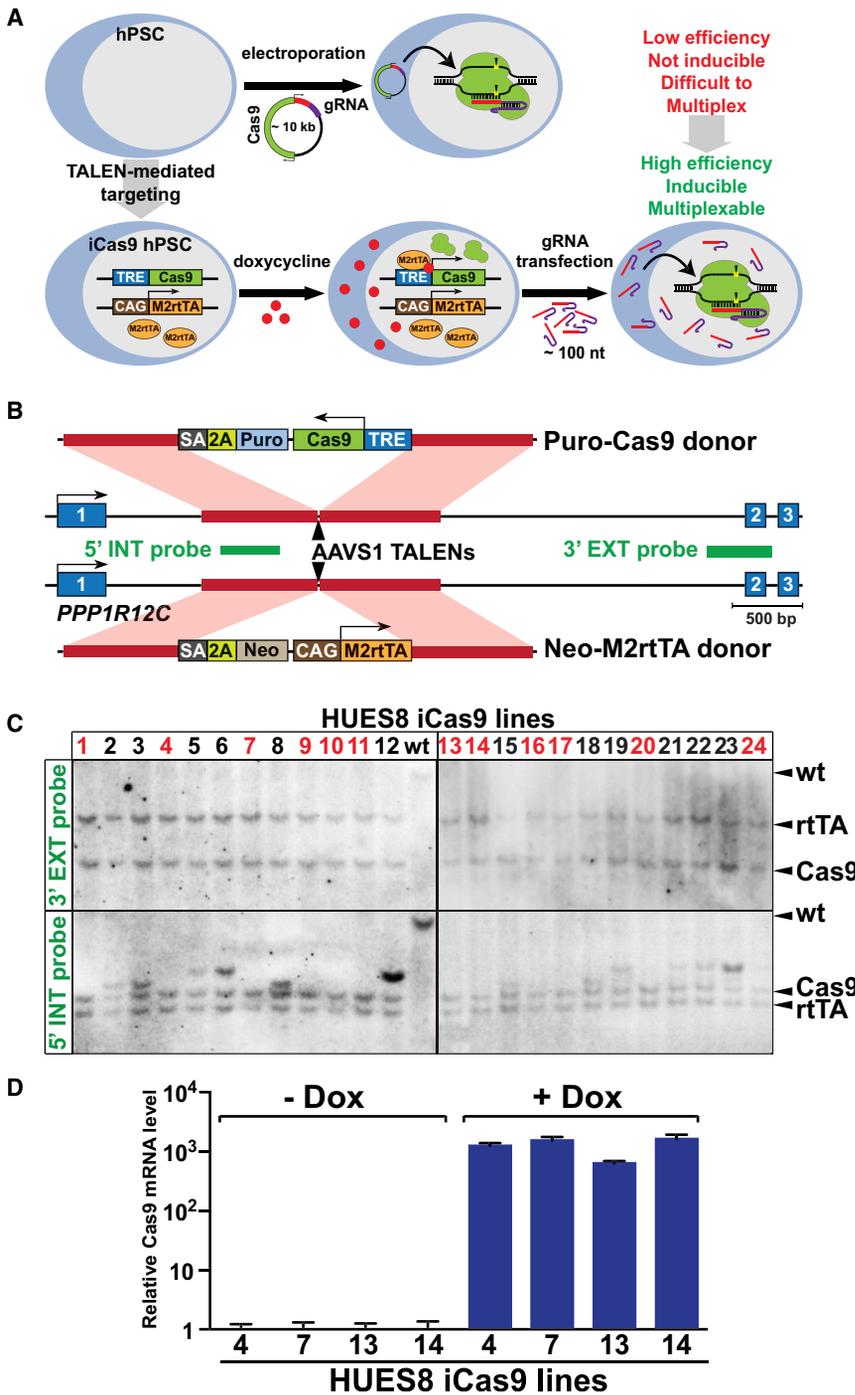
### An iCRISPR Platform for Rapid and Versatile Genome Editing

A recent study by Jaenisch and colleagues successfully applied CRISPR/Cas for highly efficient genome modifications in mESCs (Wang et al., 2013), which suggests similar approaches may work in hPSCs. Encouraged by these findings, we investigated the efficiency of CRISPR/Cas-mediated genome editing in hESCs through plasmid electroporation to transiently express Cas9 and a specific gRNA, a method that is currently used in hESCs (Figure 1A, upper panel). Although we successfully

generated heterozygous *GATA6* mutant lines at 2%–6% efficiencies, no homozygous mutants were identified out of 384 total hESC clones analyzed (Table 1). This finding is consistent with the variable and generally low gene editing efficiencies observed by others in hPSCs (Ding et al., 2013b; Hsu et al., 2013; Mali et al., 2013b) and highlights the need for a more efficient method to model human traits caused by recessive or multiple-gene mutations.

We reasoned that one could develop a more efficient and versatile genome-editing platform by first generating hPSCs that express Cas9, the invariable component of the CRISPR/Cas system. We anticipated that Cas9-expressing hPSCs would be easily transfected with gRNAs due to their small size (~100 nt), which in turn could lead to reproducible and highly efficient genome editing (Figure 1A, lower panel). We first determined the efficiency of lipid-mediated transfection of small RNAs. Using a control fluorescence-labeled, double-stranded RNA (dsRNA) probe, we estimated that ~60% hESCs would be transfected with gRNAs (Figure S1A available online). In contrast, only ~10% GFP<sup>+</sup> hESCs were detected by flow cytometry after electroporation of a GFP-expressing plasmid of comparable size with the Cas9/gRNA vectors (Figures S1A–S1D). Importantly and in contrast to plasmid electroporation, lipid-mediated transfection is associated with very low cytotoxicity. This opens up the possibility of repeated transfections and cotransfection of multiple gRNAs for multiplexed genome editing, as well as gRNA transfection during a specific stage of hPSC differentiation for inducible gene knockout.

We next engineered iCas9 hPSC lines for doxycycline-inducible expression of Cas9 through TALEN-mediated gene targeting. We chose to target the transgenes into the *AAVS1* (also known as *PPP1R12C*) locus because it has been shown to support robust and sustained transgene expression in a manner similar to that of the *Rosa26* locus in mice (Smith et al., 2008). Based on a gene trap approach used by Jaenisch and colleagues (Hockemeyer et al., 2011), we coelectroporated the *AAVS1* TALEN constructs with two donor plasmids targeting the first intron of the *PPP1R12C* gene in three hESC lines (HUES8, HUES9, and MEL-1) and one hiPSC line (BJ iPSC) (Cowan et al., 2004; Huangfu et al., 2008). One donor plasmid contains a doxycycline-inducible Cas9 expression cassette (Puro-Cas9 donor), and the other carries a constitutive reverse tetracycline transactivator (M2rtTA) expression cassette (Neo-M2rtTA donor) (Figures 1B, S1E, and S1F). Southern blot analysis revealed a high biallelic targeting efficiency: >50% of the clonal lines have both transgenes correctly inserted without additional random integrations (Figures 1C and S1G). qRT-PCR analysis confirmed the induction of Cas9 expression upon doxycycline treatment in all clonal iCas9 lines examined (Figures 1D and S1H). Further analysis showed that iCas9 hPSCs display uniform expression of the pluripotency markers OCT4, NANOG, and SOX2 (Figure S1I) and maintain the capacity to differentiate into tissue derivatives of the three embryonic germ layers in teratoma assays (Figure S1J). Furthermore, the *AAVS1* targeting strategy does not introduce apparent chromosomal aberrations, as confirmed by karyotyping analysis (two of the iCas9 lines inherited a duplication present in the parental line before the targeting experiment) (Table S1).



**Figure 1. Engineering an iCRISPR Platform through Generating iCas9 hPSCs**

(A) Schematic comparison of the current genome editing approach in hPSCs by electroporation of a ~10 kilobase (kb) CRISPR/Cas9 vector (upper panel) with the iCRISPR platform (lower panel) for genome editing in hPSCs. Cas9 protein (green) binds a chimeric gRNA composed of a constant transactivating structural region (purple) and a variable DNA recognition site (red). The Cas9/gRNA complex binds to DNA and induces a DSB (yellow). TRE, tetracycline response element; CAG, constitutive synthetic promoter; M2rtTA, reverse tetracycline transactivator sequence and protein; doxycycline, red dots.

(B) Generation of iCas9 hPSCs through TALEN-mediated gene targeting at the *AAVS1* locus. Red lines indicate homology to *PPP1R12C* intron 1; SA, splice acceptor; 2A, self-cleaving 2A peptide; Puro, Puromycin resistance gene; Neo, Neomycin resistance gene.

(C) Southern blot analysis of HUES8 iCas9 lines using 3' external and 5' internal probes. Lines carrying desired targeted insertions of the Puro-Cas9 and Neo-M2rtTA donor sequences without random integrations are indicated in red.

(D) Quantitative real-time RT-PCR (qRT-PCR) analysis of Cas9 transcript levels with or without doxycycline (Dox) treatment in HUES8 iCas9 lines. See also Figure S1 and Table S1.

**One-Step Creation of Single-, Double- and Triple-Gene Knockout hPSCs**

To determine the capacity of iCas9 hPSCs for gene editing, we designed a panel of gRNAs targeting six genes (*NGN3*, *GATA4*, *GATA6*, *TET1*, *TET2*, and *TET3*) located on five different chromosomes (Figure S2A). We first selected gRNAs effectively targeting each gene by performing experiments in 293T cells and assessing Indel rates using T7 Endonuclease I (T7EI), which specifically cleaves heteroduplexes formed by the hybridization of wild-type and mutant DNA sequences (Mashal et al., 1995)

(Figure S2B). Next, iCas9 hPSCs were treated with doxycycline and transfected with selected gRNAs generated through in vitro transcription (Figure 2A). We observed ~30% Indel rates based on T7EI assays performed on genomic DNA extracted 2 or 3 days after gRNA transfection (Figure 2B). For *TET1* and *TET2*, we designed Restriction Fragment Length Polymorphism (RFLP) assays to directly assess the loss of a restriction site in proximity to the predicted Cas9 cleavage site (Figure 2C). RFLP assays revealed a higher mutation rate compared to T7EI (42% versus 26% for *TET1*-Cr2, 42% versus 31% for *TET2*-Cr4; Figure 2D), suggesting that T7EI assays may underestimate the rate of mutations as suggested for similar Indel assays (Guschin et al., 2010). Encouraged by these high Indel rates, we then performed multiplexed genome editing through cotransfection of gRNAs targeting two or three genomic loci. Multiplexed targeting of *GATA4* and *GATA6* induced 28% and 16% of Indels in the respective loci (Figure 2E). Similar efficiencies were observed for multiplexed targeting of *TET1*, *TET2*, and *TET3* (to 17%, 11%, and 18%, respectively) through optimizing the transfection conditions (Figure 2E and Figures S3A–S3D). These results suggest that our system supports efficient genome editing including multiplexing.

**Table 1. CRISPR/Cas-Mediated Single-Gene Targeting in hESCs**

Gene	CRISPR	Delivery Method	Mutant Alleles per Clone			
			1	2	compound heterozygotes	homozygotes
NGN3	Cr5	gRNA transfection	2/48 (4.2%)	2/48	10/48	12/48 (25%)
	Cr6	gRNA transfection	1/36 (2.8%)	1/36	5/36	6/36 (16.7%)
GATA4	Cr2	gRNA transfection	2/96 (2.1%)	19/96	5/96	24/96 (25%)
GATA6	Cr1	gRNA transfection	14/40 (35%)	7/40	0/40	7/40 (17.5%)
	Cr8	gRNA transfection	6/48 (12.5%)	13/48	1/48	14/48 (29.2%)
	Cr8	plasmid electroporation	2/96 (2.1%)	0/96	0/96	0/96 (0%)
	Cr8	plasmid electroporation*	17/288 (5.9%)	0/288	0/288	0/288 (0%)
TET1	Cr2	gRNA transfection	4/48 (8.3%)	10/48	3/48	13/48 (27.1%)
TET2	Cr4	gRNA transfection	6/48 (12.5%)	14/48	14/48	28/48 (58.3%)
TET3	Cr4	gRNA transfection	4/48 (8.3%)	24/48	8/48	32/48 (66.6%)

In vitro transcribed gRNAs targeting *NGN3*, *GATA4*, *GATA6*, *NGN3*, *TET1*, *TET2*, or *TET3* were transfected in doxycycline-induced iCas9 hESCs or electroporated as plasmids (*GATA6* Cr8) in wild-type HUES8 or HUES9 (\*) hESCs. The number of lines containing each specific number of mutated alleles (1 or 2) is shown in relation to the total number of lines screened in each experiment.

We next established clonal lines after gRNA transfection (Figure 3A). For single-gene targeting, we generated mutant hESC lines affecting six individual genes after typically sequencing 36 to 96 clones for each targeting experiment. Approximately 42% (ranging from 20% to 75%) clones carried mutations in at least one allele. Notably, the majority of mutant clones carried mutations in both alleles, including both homozygous and compound heterozygous mutations (Figures 3B and 3D and Table 1). Some gRNAs induced nonrandom deletions on target sequences in a significant number of clones, a phenomenon which has been reported by others in the murine system (Wang et al., 2013). For instance, 10 out of 12 biallelic mutant lines generated with an *NGN3*-targeting gRNA (*NGN3*-Cr6) contained an identical homozygous 7 base pair (bp) deletion, and 9 out of 14 biallelic mutant clones generated with a *TET2*-targeting gRNA (*TET2*-Cr2) carried the same 3 bp deletion (Figure S3E). Preferential generation of these alleles is likely caused by microhomology-mediated repair that utilizes short stretches of repeated sequences flanking the DSB site.

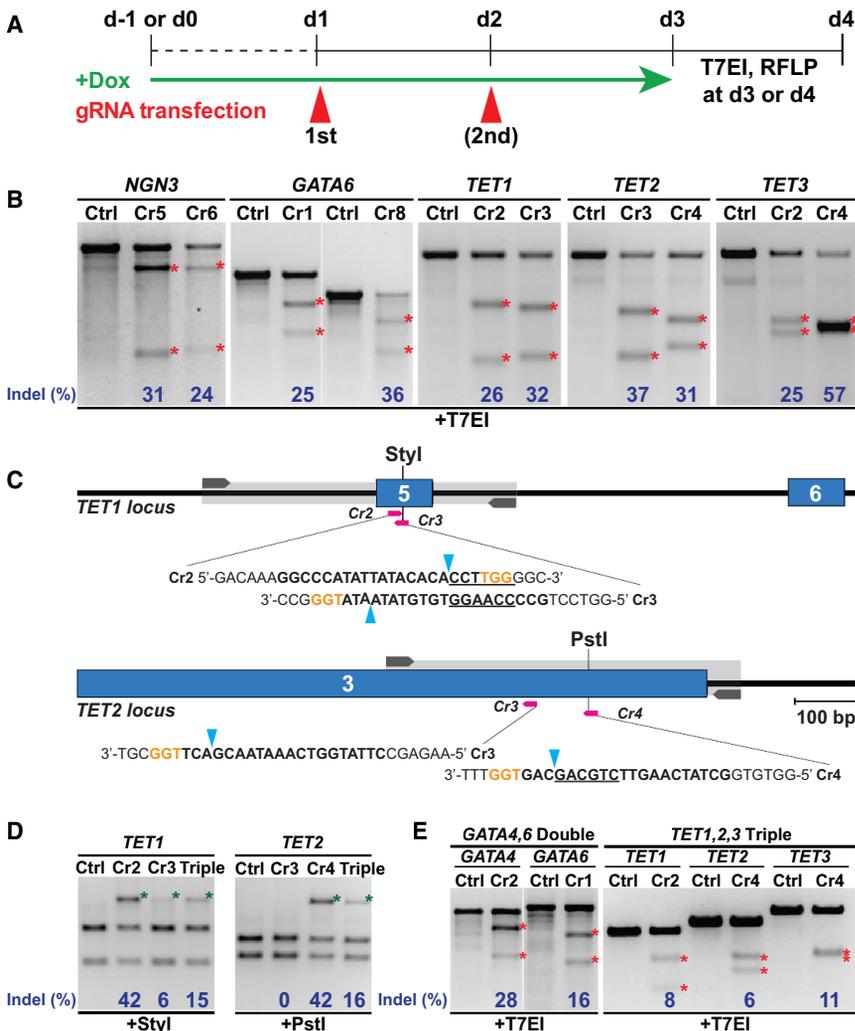
The use of iCas9 hPSCs also enabled efficient one-step multiplexed gene editing. We were able to identify mutant clones affecting two genes (i.e., all four alleles of *GATA4* and *GATA6*) and three genes (i.e., all six alleles of *TET1*, *TET2*, and *TET3*) with ~5% efficiencies (Figures 3C and 3D and Table S2). Further optimization of gRNA transfection conditions (Figures S3A–S3D) improved the efficiency of biallelic triple-gene targeting to close to 10% (Table S2). Since TET proteins are responsible for catalyzing the conversion of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) (Ito et al., 2010; Tahiliani et al., 2009), we analyzed the 5hmC levels in triple-targeted hESC lines. As expected from loss-of-function alleles, we detected a significant reduction of 5hmC levels in all triple-gene mutant lines compared with wild-type control hESCs (Figure 3E). These findings demonstrate rapid, single-step generation of multiple-gene knockout lines for loss-of-function studies. Importantly, despite the high gene-editing efficiencies, we did not detect off-target mutations in multiple single- and triple-gene mutant hESC lines analyzed (Table S3).

### Precise HDR-Mediated Genome Editing for Disease Modeling

To accurately dissect gene function in human development and disease, precise nucleotide alterations are required to either create disease-specific variants in wild-type cells or correct disease-associated mutations in patient cells. To test this approach, we cotransfected doxycycline-treated iCas9 hPSCs with a *GATA6*-targeting gRNA (*GATA6*-Cr8) and a 110 nt ssDNA repair template introducing a single nt mutation (Figures 4A and 4B). T7E1 analysis revealed ~50% and ~35% mutation rates in HUES8 and BJ iPSCs, respectively. Because the HDR ssDNA template creates a silent BsgI restriction site in *GATA6*, we next performed RFLP analysis to discriminate successful HDR-mediated gene editing from NHEJ-mediated Indels, and we determined that ~15% of HUES8 and ~8% of BJ iPSC genomic DNA sequences integrated the desired modification (Figures 4C and 4D). Next, we replated transfected cells to establish mutant clones (Figure 4B). RFLP analysis on 96 clones identified 33 (34%) clones with at least one BsgI site and 15 (16%) clones containing one BsgI site in each *GATA6* allele (Figure 4E). Sequence analysis verified that nine clones carried the desired homozygous mutations without additional sequence alterations (Figures 4F and 4G). Similar results were obtained with the *NGN3* locus as determined by RFLP analysis (Figures S4A and S4B).

Since our platform allows rapid generation of precise homozygous knockin mutations, we tested the feasibility of generating an allelic series for modeling disease susceptibility.

*APOE* is the most common risk locus associated with late-onset Alzheimer's disease (LOAD). There are three common allelic variants of *APOE*, known as  $\epsilon 2$ ,  $\epsilon 3$ , and  $\epsilon 4$ . The most common  $\epsilon 3$  variant is considered "neutral," whereas the  $\epsilon 4$  variant is the major known risk factor with a dose-dependent effect: an increased number of the  $\epsilon 4$  alleles (i.e., from 0 to 1 and to 2) is associated with increased risk and decreased onset age (Corder et al., 1993). Sequence analysis showed that both HUES8 and BJ iPSCs carry the common  $\epsilon 3/\epsilon 3$  genotype (Figure 5A). We therefore sought to create hPSCs carrying the  $\epsilon 3/\epsilon 4$  and  $\epsilon 4/\epsilon 4$  genotypes, which would be useful for modeling disease susceptibility.



**Figure 2. Single and Multiplexed gRNA Transfection Efficiently Induces Indels in iCas9 hPSCs**

(A) Schematic representation of the experimental procedure.

(B) T7EI assay for Cas9-mediated cleavage in HUES8 iCas9 cells using single gRNAs targeting *NGN3* (Cr5, 6), *GATA6* (Cr1, 8), *TET1* (Cr2, 3), *TET2* (Cr3, 4), and *TET3* (Cr2, 4). All gRNAs were transfected twice, except for *GATA6* (Cr1, 8), which were transfected once. In all figures in this study red asterisks indicate the expected T7EI-specific fragments used to quantify Indel frequency (blue).

(C) Schematic of Cas9/gRNA-targeting sites (pink arrows in all figures in this study) in *TET1* and *TET2* loci showing exon structure (blue boxes in all figures in this study), PCR amplicons (light gray boxes in all figures in this study), and Styl or PstI restriction sites used for RFLP analysis. gRNA-targeting sequences are in bold; protospacer-adjacent motif (PAM) sequences are in orange; Cas9 cleavage sites are indicated by blue arrow heads; and restriction sites are underlined, here and in all figures in this study.

(D) RFLP analysis upon *TET1* (Cr2, 3), *TET2* (Cr3, 4), or multiplexed *TET1*, 2, 3 (Triple) gRNA transfection. In all figures in this study green asterisks indicate the uncut PCR fragment used to quantify Indel frequency by RFLP (blue).

(E) T7EI assay in HUES8 iCas9 cells transfected with multiplexed *GATA4*, 6 (Double) or *TET1*, 2, 3 (Triple) gRNAs.

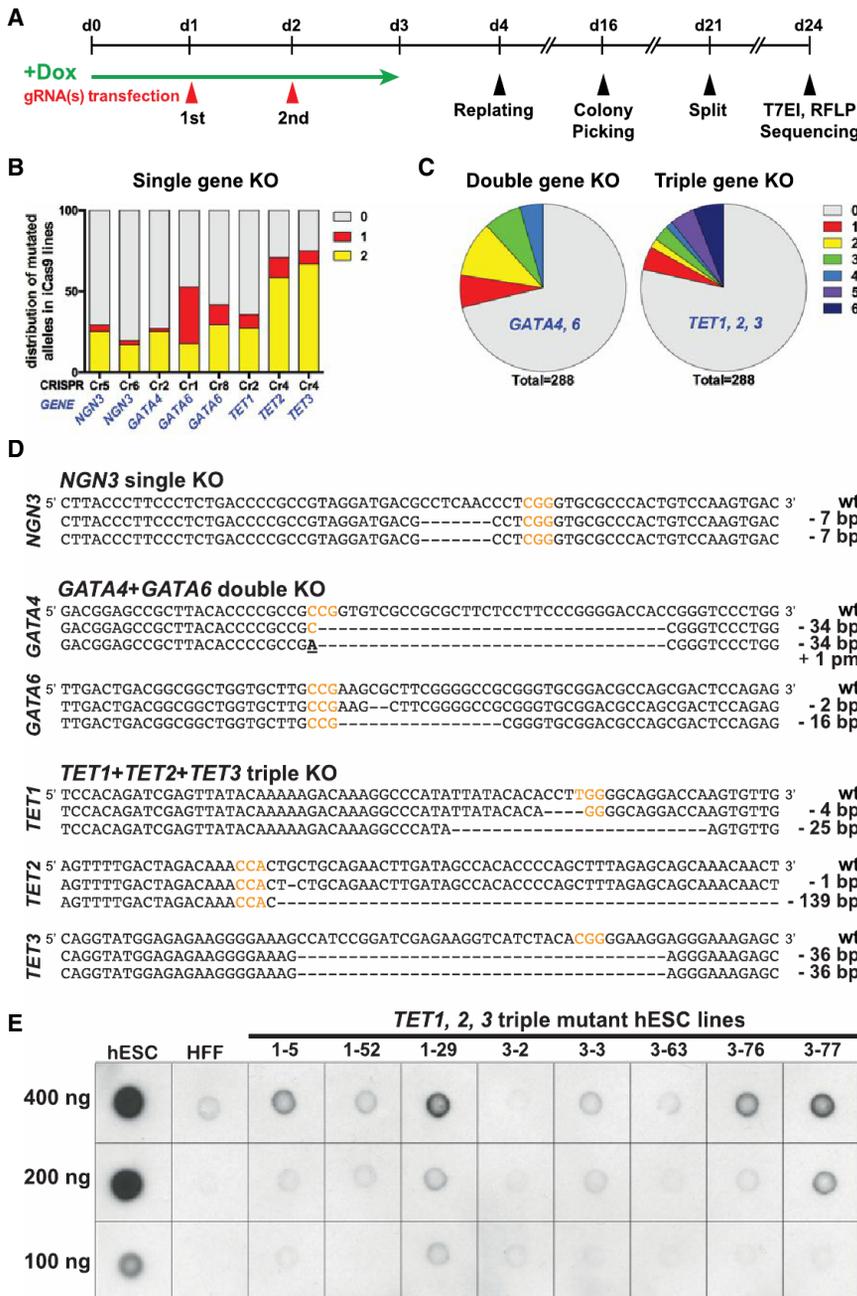
See also Figure S2.

### Inducible Gene Knockout

Inactivating genes in a temporal or tissue-specific manner has greatly facilitated the study of genes with pleiotropic effects. Since the iCRISPR platform allows inducible Cas9 expression and tightly regulated delivery of gRNA with minimal toxicity, we investigated the feasibility of conducting inducible gene knockout. We were able to differentiate iCas9 hESCs first into definitive endoderm (DE) and subsequently into pancreatic progenitor (PP) and insulin-expressing  $\beta$ -like cells based on an established protocol (Kroon et al., 2008) (Figures S5A and S5B). We proceeded to determine the efficiencies of inducible knockout by inducing Cas9 expression and performing gRNA transfection at the PP stage (Figure 6A). T7EI analysis revealed an  $\sim$ 20% Indel rate in PP cells transfected with an *NGN3*-targeting gRNA (*NGN3*-Cr6) (Figure 6B). To further investigate the feasibility of using iCas9 cells for knocking out genes in a tissue-specific manner, we enriched CXCR4-expressing DE cells through fluorescence-activated cell sorting (FACS) (D'Amour et al., 2005) and differentiated the cells further to the PP stage for gRNA transfection (Figure 6A). T7EI analysis showed that PP cells transfected with single gRNAs targeting different genes carried on average  $\sim$ 30% Indels in optimal transfection conditions (Figure 6C). As observed in other experiments, RFLP assays detected higher Indel rates, around 60% with gRNAs targeting *TET1* and *TET2* (Figure 6D). Sequencing analysis of the PCR amplicon clones, a more direct measurement

We targeted the single-nucleotide polymorphism (SNP) rs429358 in *APOE* exon 4 to introduce a T-to-C transition for conversion from  $\epsilon$ 3 to  $\epsilon$ 4 (Figure 5B). We designed gRNAs targeting *APOE* (Figure S4C) and selected the most effective one (*APOE*-Cr3) through the T7EI assay (Figure S4D). Next, we co-transfected doxycycline-treated iCas9 hPSCs with *APOE*-Cr3 gRNA and a 120 nt ssDNA repair template carrying the  $\epsilon$ 4 polymorphism and a *Ban*I silent restriction site (Figure 5B). T7EI analysis revealed  $\sim$ 21% and  $\sim$ 13% mutation rates in the HUES8 and BJ iPSC *APOE* loci, respectively. As observed previously, RFLP assay using *Not*I revealed a higher mutation rate (49% in HUES8 and 23% in BJ iPSCs; Figures 5C and 5D). Using *Ban*I RFLP we further discriminated HDR-mediated gene editing from NHEJ-mediated Indels and estimated that  $\sim$ 2%–3% of the target sequence had the desired modification (Figures 5C and 5D). Indeed, sequence analysis confirmed successful generation of clonal lines carrying an *APOE* allelic series in both HUES8 and BJ iPSCs: the original  $\epsilon$ 3/ $\epsilon$ 3 and the modified  $\epsilon$ 3/ $\epsilon$ 4 and  $\epsilon$ 4/ $\epsilon$ 4 genotypes (Figure 5A). This strategy could be easily modified to convert the risk-associated  $\epsilon$ 4 variant to the standard  $\epsilon$ 3 variant to create isogenic control cells for studies of hPSCs carrying the  $\epsilon$ 4 variant or for therapeutic purposes.

able Cas9 expression and tightly regulated delivery of gRNA with minimal toxicity, we investigated the feasibility of conducting inducible gene knockout. We were able to differentiate iCas9 hESCs first into definitive endoderm (DE) and subsequently into pancreatic progenitor (PP) and insulin-expressing  $\beta$ -like cells based on an established protocol (Kroon et al., 2008) (Figures S5A and S5B). We proceeded to determine the efficiencies of inducible knockout by inducing Cas9 expression and performing gRNA transfection at the PP stage (Figure 6A). T7EI analysis revealed an  $\sim$ 20% Indel rate in PP cells transfected with an *NGN3*-targeting gRNA (*NGN3*-Cr6) (Figure 6B). To further investigate the feasibility of using iCas9 cells for knocking out genes in a tissue-specific manner, we enriched CXCR4-expressing DE cells through fluorescence-activated cell sorting (FACS) (D'Amour et al., 2005) and differentiated the cells further to the PP stage for gRNA transfection (Figure 6A). T7EI analysis showed that PP cells transfected with single gRNAs targeting different genes carried on average  $\sim$ 30% Indels in optimal transfection conditions (Figure 6C). As observed in other experiments, RFLP assays detected higher Indel rates, around 60% with gRNAs targeting *TET1* and *TET2* (Figure 6D). Sequencing analysis of the PCR amplicon clones, a more direct measurement



**Figure 3. Single and Multiplexed Gene Targeting**

(A) Strategy of gene targeting. (B and C) Allelic sequence distribution in HUES8 clonal lines generated with single gRNAs targeting *NGN3* (Cr5, 6), *GATA4* (Cr2), *GATA6* (Cr1, 8), *TET1* (Cr2), *TET2* (Cr4), or *TET3* (Cr4) (B), or multiplexed gRNAs targeting *GATA4* and *GATA6* (Cr2 and Cr1, respectively) or *TET1*, *TET2*, and *TET3* (Cr2, Cr4, and Cr4, respectively) (C). (D) Representative sequences of various knockout (KO) mutant clones with PAM sequences labeled in orange. pm, point mutation. (E) Analysis of 5hmC levels in DNA isolated from *TET1*, 2, 3 triple-targeted hESC clones by dot blot assay using an anti-5hmC antibody. Wild-type HUES8 iCas9 cells and human foreskin fibroblasts (HFFs) were used as controls. See also Figure S3 and Tables S2 and S3.

AAVS1-TALEN constructs allows the generation of hPSC lines for convenient, inducible knockout studies: all cells would express the gRNA of interest, which, upon doxycycline treatment, would target the induced Cas9 protein to the desired genomic locus (Figure 6F). Following this strategy, we generated hESC lines for inducible knockout of *NGN3* and *TET2*, named iCr*NGN3* and iCr*TET2*, respectively (Figure S5D). Doxycycline treatment of differentiated iCr*NGN3* and iCr*TET2* hESCs at the PP stage resulted in consistent induction of Indels as revealed by T7E1 and RFLP assays (Figures 5G and 5H). Sequencing analysis of the PCR amplicons revealed ~55% and ~75% mutation rates in induced iCr*NGN3* and iCr*TET2* hESCs (Figure 6I). The ratio of frameshift versus nonframeshift mutations differed significantly between induced iCr*NGN3* and iCr*TET2* hESCs, suggesting that the knockout efficiencies also depend on the types of mutations generated. Importantly, to determine the tightness of our

inducible system, we analyzed PCR amplicons after prolonged in vitro culture of iCr*NGN3* hESCs (15 passages after establishment of the line) in the absence of doxycycline treatment and identified no mutations associated with potential “leaky” targeting at the *NGN3* locus (Figure S5E). These findings demonstrate the versatile use of the iCRISPR platform as an approach for inducible gene knockout in hPSCs and their differentiated progeny.

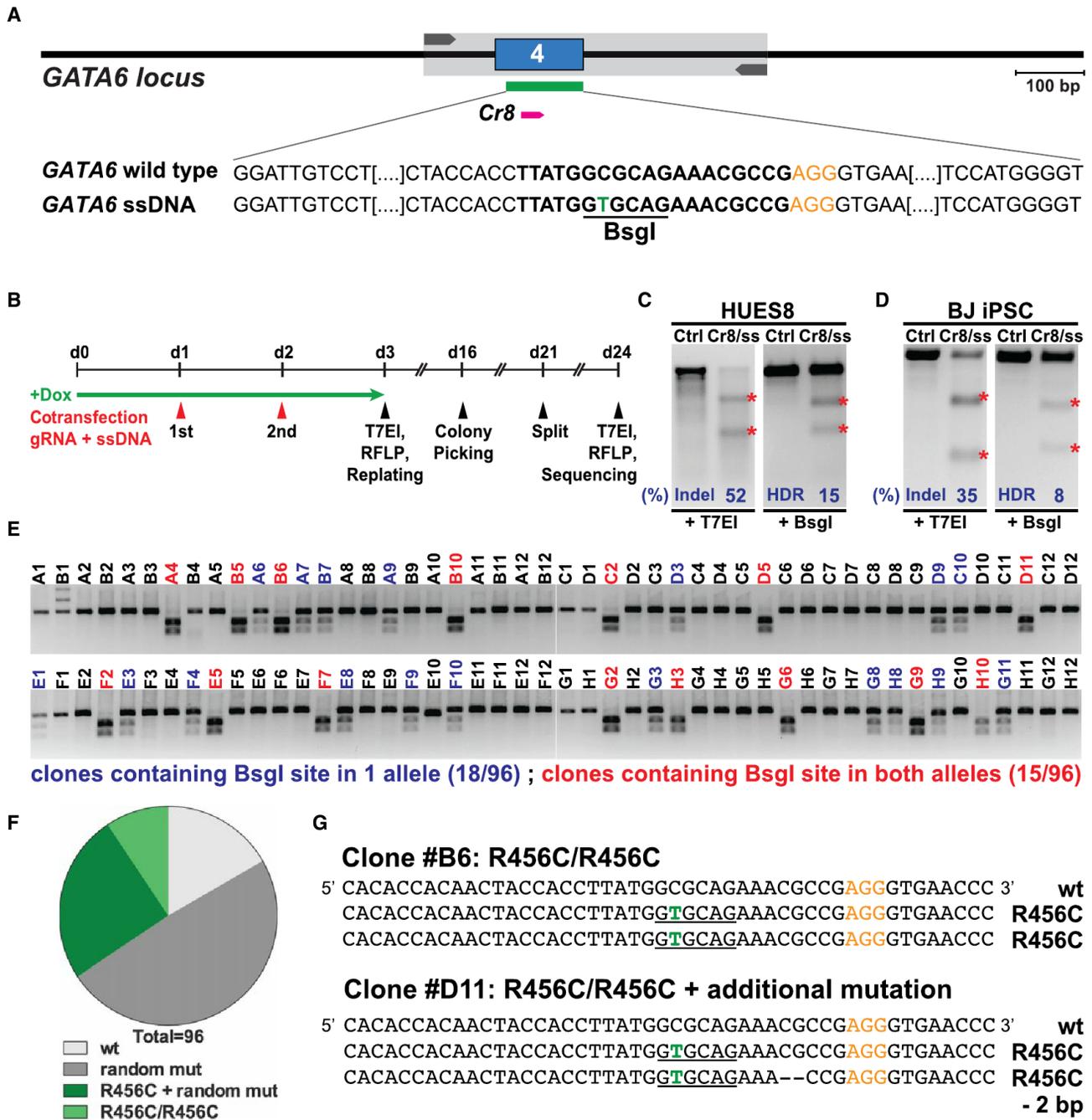
than T7E1 or RFLP, revealed an average of ~78% mutation rate for all targeting experiments, reaching up to ~97% with *TET3* (Figure 6E). Notably, the majority of mutations were frameshift, supporting the usage of this inducible gene-editing approach for examination of loss-of-function phenotypes in a temporal-specific and cell-type-specific manner.

The gRNA transfection method described above may interfere with some differentiation procedures. To overcome this limitation, we generated a Puro-Cr donor by modifying the Puro-Cas9 AAVS1 donor vector to include a constitutive gRNA expression module in addition to the doxycycline-inducible Cas9 expression cassette (Figure S5C). Coelectroporation of this Puro-Cr donor with the Neo-M2rtTA donor and the

inducible system, we analyzed PCR amplicons after prolonged in vitro culture of iCr*NGN3* hESCs (15 passages after establishment of the line) in the absence of doxycycline treatment and identified no mutations associated with potential “leaky” targeting at the *NGN3* locus (Figure S5E). These findings demonstrate the versatile use of the iCRISPR platform as an approach for inducible gene knockout in hPSCs and their differentiated progeny.

**DISCUSSION**

Efficient genetic engineering in hPSCs is crucial for elucidating the genes and molecular pathways that underlie complex human

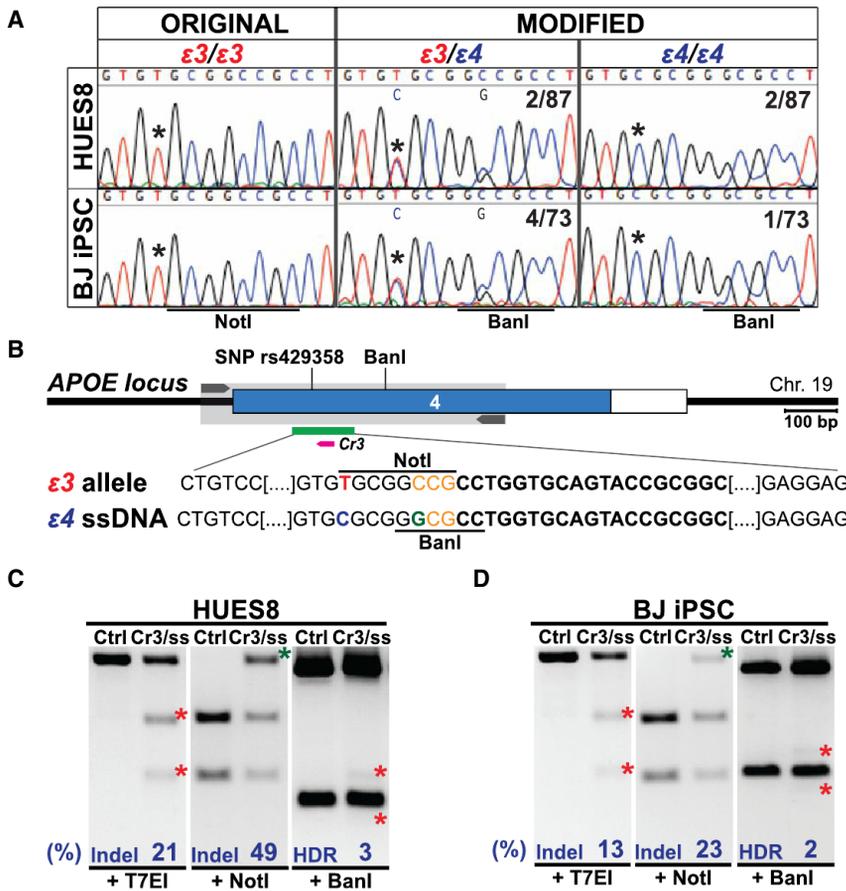


**Figure 4. HDR-Mediated Genome Editing**

(A) Schematic of Cas9/gRNA and ssDNA oligo targeting sites at the *GATA6* locus. A C > T substitution (green) was introduced in the ssDNA HDR template, generating a new *BsgI* restriction site (underlined), resulting in an R456C amino acid substitution in *GATA6*.  
 (B) Strategy of HDR-mediated genome editing.  
 (C and D) T7E1 and RFLP assay in HUES8 (C) and BJ iPSCs (D) cotransfected with *GATA6* gRNA (Cr8) and ssDNA HDR repair template.  
 (E and F) RFLP analysis (E) and allelic sequence distribution (F) in clones generated with *GATA6* gRNA/ssDNA. wt, wild-type; mut, mutation. "R456C + random mut" includes clones with undesired mutations in addition to the R456C modification in one or both alleles.  
 (G) Representative sequences of one homozygous (R456C/R456C) and one compound heterozygous (R456C/R456C + additional mutations) *GATA6* mutant clone.

traits. While recent approaches based on TALEN and CRISPR/Cas systems have led to encouraging results, a more efficient and universal platform would be highly desirable for large-scale

analysis of gene function in the postgenomic era. The iCRISPR platform offers a rapid and efficient approach to introduce mutations in any gene of interest. Based on the high targeting



**Figure 5. Generation of an Allelic Series at the APOE Locus**

(A) Sequence analysis of SNP rs429358 in parental HUES8 hESCs and BJ iPSCs lines ( $\epsilon 3/\epsilon 3$ ) and derived HDR-mediated edited clones ( $\epsilon 3/\epsilon 4$  or  $\epsilon 4/\epsilon 4$ ). Ratios indicate the number of colonies with the specified genotype out of the total number of colonies analyzed.

(B) Schematic of Cas9/gRNA and ssDNA oligo targeting sites at the APOE locus. A T (red) > C (blue) substitution was introduced in the  $\epsilon 4$  ssDNA to convert the  $\epsilon 3$  allele into  $\epsilon 4$ , and in addition, a C > G substitution (green) was introduced, generating a novel BanI and disrupting the endogenous NotI restriction site (underlined). SNP rs429358 and BanI sites are indicated in APOE exon 4.

(C and D) T7E1 and RFLP assay in HUES8 hESCs (C) and BJ iPSCs (D) cotransfected with APOE gRNA (Cr3) and  $\epsilon 4$  ssDNA HDR repair template. Ratios indicate the number of colonies with the specified genotype out of the total number of colonies analyzed.

See also Figure S4.

16 additional passages following line establishment, and karyotypically normal mutant lines have been established using the iCRISPR system (Table S1); second, even though iCrNGN3 hESCs constitutively express the gRNA targeting the NGN3 locus, sequencing analysis revealed no mutations in the targeted site even after 15 additional passages in the absence of

efficiencies observed in our study, the analysis of ~24 colonies should be sufficient to establish multiple monoallelic and biallelic mutant lines for loss-of-function studies of a single gene. We expect that a trained individual could readily perform gRNA transfection and analyze ~300 colonies at a time, enabling the generation of mutant lines affecting 12 genes in just 1 month, or 144 genes in a year. Further optimization would enable the use of the iCRISPR platform for high-throughput genetic analysis of disease phenotypes in hPSCs. For instance, one may use the iCRISPR platform to screen an array of human disease-associated genes for their functional relevance.

Using hESC lines in the current NIH hESC registry (HUES8, HUES9, and MEL-1) as well as BJ iPSCs, we have generated multiple iCas9 lines which will be available to the academic community upon request. Establishing the iCRISPR platform in different hPSC backgrounds requires minimal efforts due to the highly efficient TALEN-mediated AAVS1 targeting approach: only a handful of clonal lines need to be analyzed and the entire procedure takes around 1 month. Once the system is established, it can be used repeatedly and reliably for rapid and versatile genetic studies in hPSCs, which is highly desirable for laboratories interested in systematically interrogating biological and disease mechanisms in isogenic backgrounds. Despite the stable integration of Cas9 in hPSCs, Cas9 activity is tightly regulated by doxycycline treatment, and no adverse effects were observed. Two observations support this conclusion: first, the karyotype of wild-type iCas9 lines remains normal even after

doxycycline treatment (Figure S5E). Future studies can also include the option to remove the Cas9 gene when desired by flanking the Cas9 gene with loxP sites, for example.

It is worth noting that iCRISPR may create complete loss-of-function (null), partial loss-of-function (hypomorphic), or, less frequently, gain-of-function (e.g., dominant-negative) alleles. Studying such allelic series may provide valuable insights into the molecular basis of distinct phenotypes caused by different mutations in a single gene, though careful analysis is needed to determine the exact nature of individual mutant alleles. Recent studies suggest that the CRISPR/Cas system may have off-target effects (Cho et al., 2014; Fu et al., 2013; Hsu et al., 2013; Mali et al., 2013a; Pattanayak et al., 2013), potentially confounding genetic studies in hPSCs. Although our analysis so far has not identified off-target mutations, CRISPR design using recently developed algorithms can reduce potential off-target effects (Hsu et al., 2013; Mali et al., 2013b). Characterizing the full extent of off-target activities in future studies would likely allow more rational design of gRNAs with higher specificity. To increase target specificity, a double nicking approach has been developed recently using a nickase version of Cas9 (nCas9) that requires the cooperation between two gRNAs to create a DNA DSB (Mali et al., 2013a; Ran et al., 2013a). This strategy significantly ameliorates, but does not completely eliminate, off-target activities. We propose two practical solutions to overcome this limitation for genetic studies in hPSCs. In a first approach, one may generate independent mutant lines using

gRNAs targeting different sequences in the same gene: observing the same phenotype from multiple independent mutant lines would strongly suggest that the phenotype is caused by the disruption of the gene of interest. In a complementary approach, one may perform rescue experiments using the iCRISPR platform, which supports highly efficient HDR-mediated scarless correction of mutant sequences.

The iCRISPR platform is also highly versatile. Our platform allows the generation of hPSCs carrying biallelic mutations in multiple genes in just 1 month, which will greatly accelerate the study of complex multigenic interactions that are often challenging to model *in vivo* due to the random segregation of alleles. This is highly valuable not only for studying multigenic disorders, but also for genetic epistasis analysis involving multiple genes. Additionally, iCRISPR can also be used to introduce specific nucleotide modifications with high efficiency. This is critical for dissecting protein functional domains, precisely modeling human diseases, and potentially correcting disease-associated mutations for therapeutic intervention. A challenge for modeling complex diseases lies in the large number of susceptibility loci that are each associated with multiple sequence variants. For instance, *APOE* is associated with three common polymorphisms, and it is among the close to 20 LOAD-susceptibility loci identified so far (Lambert et al., 2013). The iCRISPR platform offers an ideal solution: it enables the rapid generation of allelic series, as demonstrated for the *APOE* locus, and it is expected to greatly facilitate the perturbation of multiple disease-associated loci either individually or in combination in isogenic backgrounds.

Finally, we demonstrate that the CRISPR/Cas system can be used for inducible gene knockout, which to our knowledge has not yet been shown in any *in vivo* or cell culture systems. While the ubiquitous rTA expression system used in the present study only allowed temporal regulation of Cas9 expression, we were able to achieve tissue-specific regulation by enriching defined cell populations through FACS. Future studies may also achieve tissue specificity by expressing rTA from tissue-specific promoters. This approach may be further extended to *in vivo* systems as a simpler alternative to conventional conditional knockout strategies, which generally require complex genetic configurations involving tissue-specific expression of a site-specific recombinase (e.g., Cre) combined with a conditional allele (e.g., a “floxed” allele).

In addition to generating mutant alleles with small Indels or precise nucleotide alterations through NHEJ- or HDR-mediated repair mechanisms as shown in the present study, the iCRISPR platform is likely to also greatly facilitate other types of genome engineering in hPSCs. For instance, iCRISPR may be used to create larger deletions for study of noncoding RNAs or gene regulatory regions such as promoters and enhancers. It may also facilitate the generation of reporter alleles through HDR-mediated gene targeting using long donor DNA templates or defined chromosomal rearrangements such as translocations. Also, by swapping Cas9 with other newly developed Cas9 variants (e.g., nCas9, dCas9, or dCas9-KRAB) (Gilbert et al., 2013; Mali et al., 2013a; Qi et al., 2013; Ran et al., 2013b), one may repurpose iCRISPR for additional tasks such as gene regulation. Because of the highly efficient TALEN-mediated gene targeting at the *AAVS1* locus and the superior speed, one can easily apply

iCRISPR to any human cell line of interest, including patient-specific hiPSCs or cancer cell lines. This versatility allows the study of gene function in practically any hPSC line or genetic background of choice.

## EXPERIMENTAL PROCEDURES

### Construction of TALENs, *AAVS1* Targeting Vectors, and Cas9/gRNA-Expressing Vectors

A pair of TALENs (*AAVS1*-TALEN-L targeting CCCCTCCACCCCACAGT and *AAVS1*-TALEN-R targeting TTTCTGTACCAATCCT) was generated to target the first intron of the constitutively expressed gene *PPP1R12C* at the *AAVS1* locus (Hockemeyer et al., 2011). The TALEN constructs were constructed following a published protocol (Sanjana et al., 2012). In brief, a library of TALE monomers with complementary overhangs was built by PCR using vector templates from Addgene (32180, 32181, 32182, or 32183). Monomers were joined into hexamers according to the target DNA sequence. Next, the hexamers were linked together and incorporated into the full-length TALEN expression backbone (Addgene 32190).

Neo-M2rTA donor (Figure S1F) was kindly provided by D. Hockemeyer. Puro-Cas9 donor (Figure S1E) was constructed by replacing EGFP in the TRE-TIGHT-EGFP-BW plasmid (Addgene plasmid 22077) (Hockemeyer et al., 2011) with the human codon-optimized *Streptococcus pyogenes* Cas9 cDNA amplified by PCR from pX260 (Addgene plasmid 42229) (Cong et al., 2013). Puro-Cr donor was generated by introducing a chimeric gRNA expression cassette, PCR amplified from pX330 (Addgene plasmid 42230) (Cong et al., 2013), 3' of Cas9 in Puro-Cas9 donor (Figure S5C). piCRct Entry (Figure S1C), a human codon-optimized *S. pyogenes* Cas9 expression entry vector carrying a bicistronic crRNA/tracrRNA duplex expression cassette, was generated by modifying pX260. piCRg Entry (Figure S1D) was built by deleting the bicistronic crRNA/tracrRNA duplex expression cassette from piCRct Entry and replacing it with the chimeric gRNA expression cassette PCR amplified from pX330. To generate CRISPR/Cas9 expression vectors targeting specific genomic loci, 30 or 20 bp of sequence located 5' of the PAM sequence was cloned in piCRct Entry or piCRg Entry, respectively, following an established protocol (Cong et al., 2013). Briefly, piCRct Entry or piCRg Entry were digested with BbsI, dephosphorylated, and gel purified. A pair of oligos including either 30 or 20 bp homology (Table S4) was annealed and phosphorylated, generating BbsI overhangs that could be cloned into the BbsI-digested, dephosphorylated vector. Vectors described in this manuscript will be available to academic researchers through Addgene.

### Production of gRNAs through In Vitro Transcription

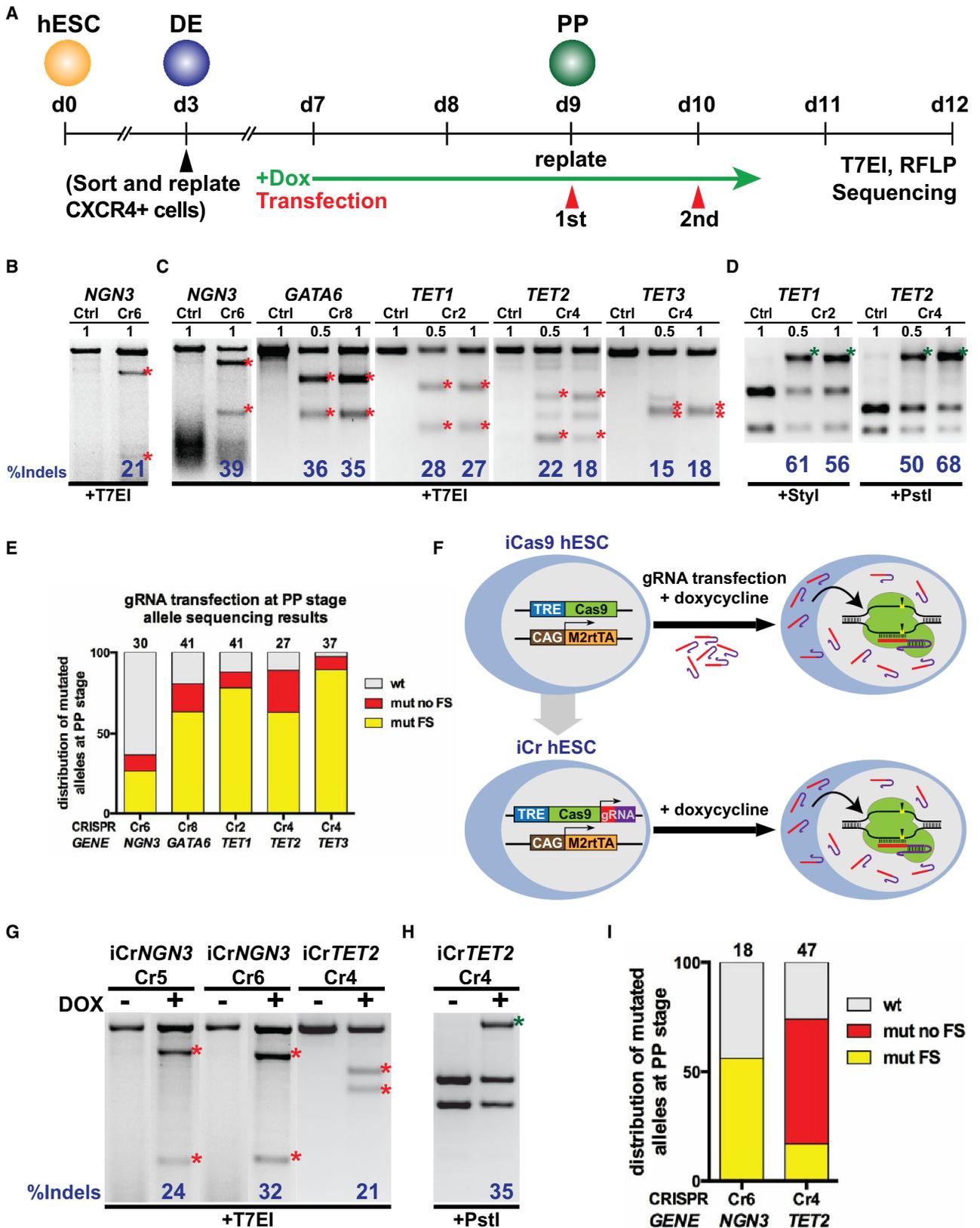
For production of gRNA, we first generated a T7-gRNA *in vitro* transcription (IVT) template by adding the T7 promoter to the gRNA sequence in the piCRg Entry vector through PCR amplification using CRISPR-specific forward primers and a universal reverse primer (Table S4).

Alternatively, for *APOE* gRNAs we designed a 120 nt oligo including the T7 promoter and the full-length gRNA sequence. This oligo was used as a template for PCR amplification using T7 and gRNA universal primers (Table S4).

T7-gRNA PCR products were used as templates for IVT using the MEGAscript T7 kit (Life Technologies). The resulting gRNAs were purified using the MEGAclear kit (Life Technologies), eluted in RNase-free water, and stored at  $-80^{\circ}\text{C}$  until use.

### gRNA or gRNA + ssDNA Transfection

iCas9 hPSCs were treated with doxycycline (2  $\mu\text{g}/\text{ml}$ ) for 1 or 2 days before and during transfection. For transfection, cells were dissociated using Accutase (Stem Cell Technologies) or TrypLE (Life Technologies), replated onto iMEF-coated plates, and transfected in suspension with gRNAs or gRNA + ssDNA using Lipofectamine RNAiMAX (Life Technologies) following the manufacturer's instructions. Briefly, gRNA and ssDNA were added at a 10 nM and 20 nM final concentration, respectively, unless otherwise indicated. gRNAs (or gRNA + ssDNA) and Lipofectamine RNAiMAX were diluted separately in Opti-MEM (Life Technologies), mixed together, incubated for 5 min at RT, and added dropwise to cultured hPSCs. A second transfection was performed 24 hr later in some experiments.



(legend on next page)

### T7E1 and RFLP Analysis for Assessment of Genome Modification and Off-Target Analysis

Genomic DNA was extracted 2 or 3 days after the last gRNA transfection. Genomic regions flanking the CRISPR target sites were PCR amplified (Table S4). For T7E1 assays, 12  $\mu$ l of PCR products were denatured and reannealed in NEB Buffer 2 (New England Biolabs) in a total volume of 25  $\mu$ l using the following protocol: 95°C, 5 min; 95°C–85°C at  $-2^\circ\text{C/s}$ ; 85°C–25°C at  $-0.1^\circ\text{C/s}$ ; hold at 4°C. Then, 12.5  $\mu$ l of hybridized PCR products were treated with 5 U of T7E1 at 37°C for 15 min in 13  $\mu$ l final reaction volume. Products were then analyzed on 2.5% agarose gels and imaged with a Gel Doc gel imaging system (Bio-Rad). Quantification was based on relative band intensities using ImageJ. Indel percentage was determined by the formula  $100 \times (1 - (b + c) / (a + b + c))^{1/2}$ , where *a* is the integrated intensity of the undigested PCR product, and *b* and *c* are the integrated intensities of each cleavage product (Hsu et al., 2013). For RFLP analysis, 10  $\mu$ l PCR products were digested with enzymes and analyzed on 2.5% agarose gel. Indel percentage was determined by the formula  $100 \times a / (a + b + c)$  or  $100 \times (b + c) / (a + b + c)$ .

For off-target analysis, ectopic gRNA targets were identified using the rules outlined in a previous study (Mali et al., 2013b). The most likely off targets falling in gene coding sequences (four or five sites per gRNA-mediated targeting experiment) were analyzed through sequencing. Primers for PCR amplification and sequencing of each off-target site are summarized in Table S4.

### Establishment of Knockout or Knockin Lines through NHEJ- or HDR-Mediated Repair

Two days after the last gRNA or gRNA + ssDNA (Table S4) transfection, hPSCs were dissociated into single cells and replated at  $\sim 2,000$  cells per 10 cm dish. Cells were allowed to grow until colonies from single cells became visible ( $\sim 10$  days). Single colonies were randomly picked based on hPSC morphology, mechanically disaggregated, and replated into individual wells of 96-well plates. Colonies were amplified, replated as described above, and analyzed by Sanger sequencing to enable identification of mutant clones. Clonal cell lines carrying desired mutations were amplified and frozen down. Alternatively, for *APOE* experiments, colonies were picked and directly analyzed by Sanger sequencing to enable the identification of mutant clones.

### SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes five figures, four tables, and Supplemental Experimental Procedures and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2014.05.018>.

### AUTHOR CONTRIBUTIONS

D.H., F.G., Z.Z., Z.S., and K.L. conceived the experiments; F.G., Z.Z., Z.S., K.L., N.V., and Q.V.L. performed the experiments; and D.H., F.G., and Z.Z. wrote the manuscript.

### ACKNOWLEDGMENTS

We thank Feng Zhang and Rudolf Jaenisch for providing vectors through Addgene; Dirk Hockemeyer for the Neo-M2rtTA donor and the *AAVS1* 3' external probe vector; Qiurong Ding from Chad Cowan's laboratory for

providing technical advice on picking hPSC colonies for identification and establishment of mutant clones; Drs. Ed Stanley and Andrew Elefanty for kindly providing MEL-1 cells; Gouri Nanjangud at the MSKCC Molecular Cytogenetics Core for performing karyotype analysis; Daniela Georgieva for technical assistance; and Lorenz Studer, Wenjun Guo, and members of the Huangfu laboratory for insightful discussions and critical reading of the manuscript. This study was funded in part by the NIH (R01DK096239), NYSTEM (C029156), the Basil O'Connor Starter Scholar Award from March of Dimes Birth Defects Foundation, the Tri-Institutional Stem Cell Initiative, a Louis V. Gerstner Jr. Young Investigators Award, and the MSKCC Society Special Projects Committee. F.G. and Z.Z. were supported by the New York State Stem Cell Science fellowship from the Center for Stem Cell Biology of the Sloan-Kettering Institute. N.V. was supported by a Howard Hughes Medical Institute (HHMI) Medical Research Fellowship.

Received: January 27, 2014

Revised: May 6, 2014

Accepted: May 28, 2014

Published: June 12, 2014

### REFERENCES

- Capecci, M.R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.* 6, 507–512.
- Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 39, e82.
- Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* 31, 230–232.
- Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132–141.
- 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.
- Cooper, G.M., and Shendure, J. (2011). Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat. Rev. Genet.* 12, 628–640.
- Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L., and Pericak-Vance, M.A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261, 921–923.
- Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D., and Melton, D.A. (2004). Derivation of embryonic stem-cell lines from human blastocysts. *N. Engl. J. Med.* 350, 1353–1356.
- D'Amour, K.A., Agulnick, A.D., Eliazar, S., Kelly, O.G., Kroon, E., and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat. Biotechnol.* 23, 1534–1541.

### Figure 6. Inducible Gene Knockout

- (A) Strategy of inducible gene knockout. DE, definitive endoderm stage; PP, pancreatic progenitor stage.
- (B and C) T7E1 assay in HUES8 iCas9 cells transfected with gRNAs targeting *NGN3* (Cr6) at the PP stage (B), or sorted CXCR4<sup>+</sup> DE cells differentiated to the PP stage and transfected with gRNAs targeting *NGN3* (Cr6), *GATA6* (Cr8), *TET1* (Cr2), *TET2* (Cr4), and *TET3* (Cr4) (C).
- (D) RFLP analysis of *TET1* Cr2 and *TET2* Cr4 transfected samples from (C).
- (E) Allelic sequence distribution in differentiated HUES8 iCas9 cells transfected with different gRNAs at PP stage. wt, wild-type; mut, mutation; FS, frameshift.
- (F) Schematic illustrating the TALEN-mediated establishment of iCr hESCs for inducible gene knockout. A U6 Pol III driving constitutive expression of a specific gRNA is included 3' of the inducible Cas9 expression cassette at the *AAVS1* locus, allowing gene knockout in all doxycycline-treated iCr hESCs.
- (G) T7E1 assay in differentiated iCr hESC lines expressing *NGN3*-Cr5 (#7), Cr6 (#12), or *TET2*-Cr4 (#2) gRNAs, treated with doxycycline at the PP stage.
- (H) RFLP analysis of iCr *TET2* (Cr4, #2) hESCs from (G).
- (I) Allelic sequence distribution in differentiated iCr hESCs treated with doxycycline at the PP stage. The number of independent clones analyzed is indicated above each column in (E) and (I).
- See also Figure S5.

- DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* *41*, 4336–4343.
- Ding, Q., Lee, Y.K., Schaefer, E.A., Peters, D.T., Veres, A., Kim, K., Kuperwasser, N., Motola, D.L., Meissner, T.B., Hendriks, W.T., et al. (2013a). A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* *12*, 238–251.
- Ding, Q., Regan, S.N., Xia, Y., Oostrom, L.A., Cowan, C.A., and Musunuru, K. (2013b). Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* *12*, 393–394.
- Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* *31*, 822–826.
- 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.
- Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., and O'Connor-Giles, K.M. (2013). Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* *194*, 1029–1035.
- Guschin, D.Y., Waite, A.J., Katibah, G.E., Miller, J.C., Holmes, M.C., and Rebar, E.J. (2010). A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.* *649*, 247–256.
- Hockemeyer, D., and Jaenisch, R. (2010). Gene targeting in human pluripotent cells. *Cold Spring Harb. Symp. Quant. Biol.* *75*, 201–209.
- Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* *27*, 851–857.
- 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.
- Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* *31*, 827–832.
- Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat. Biotechnol.* *26*, 1269–1275.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* *31*, 227–229.
- Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* *466*, 1129–1133.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816–821.
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. *eLife* *2*, e00471.
- Joung, J.K., and Sander, J.D. (2013). TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* *14*, 49–55.
- Kroon, E., Martinson, L.A., Kadoya, K., Bang, A.G., Kelly, O.G., Eliazar, S., Young, H., Richardson, M., Smart, N.G., Cunningham, J., et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat. Biotechnol.* *26*, 443–452.
- Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Granier-Boley, B., et al.; European Alzheimer's Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* *45*, 1452–1458.
- 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.
- Mashal, R.D., Koontz, J., and Sklar, J. (1995). Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. *Nat. Genet.* *9*, 177–183.
- Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., et al. (2011). A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* *29*, 143–148.
- Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* *31*, 839–843.
- 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.
- Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* *154*, 1380–1389.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013b). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* *8*, 2281–2308.
- Sanjana, N.E., Cong, L., Zhou, Y., Cunniff, M.M., Feng, G., and Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. *Nat. Protoc.* *7*, 171–192.
- Smith, J.R., Maguire, S., Davis, L.A., Alexander, M., Yang, F., Chandran, S., French-Constant, C., and Pedersen, R.A. (2008). Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* *26*, 496–504.
- Soldner, F., Laganière, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.I., Myers, R.H., Lindquist, S., et al. (2011). Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* *146*, 318–331.
- Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., and Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* *324*, 930–935.
- Thomas, K.R., and Capecchi, M.R. (1986). Targeting of genes to specific sites in the mammalian genome. *Cold Spring Harb. Symp. Quant. Biol.* *51*, 1101–1113.
- Thomas, K.R., Folger, K.R., and Capecchi, M.R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. *Cell* *44*, 419–428.
- Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* *11*, 636–646.
- Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* *153*, 910–918.
- Zhu, Z., and Huangfu, D. (2013). Human pluripotent stem cells: an emerging model in developmental biology. *Development* *140*, 705–717.



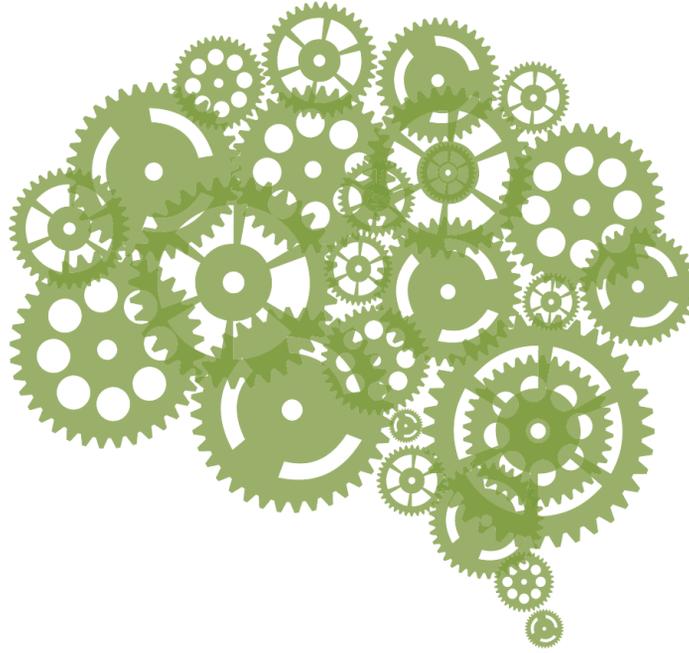
Dharmacon™  
part of GE Healthcare

# CRISPR-Cas9 gene editing

Simplify your workflow with synthetic RNA. The Dharmacon™ Edit-R™ CRISPR-Cas9 platform eliminates time-consuming cloning tasks and sequencing steps – saving you time and money. With transfection-ready, endotoxin-free DNA and synthetic RNA components, our platform allows you to rapidly target several sites per gene, for multiple genes.

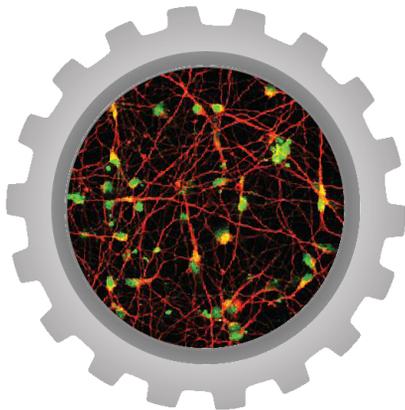
**Fast, scalable, and efficient gene editing**

[gelifesciences.com/CRISPR](http://gelifesciences.com/CRISPR)



# Model the **Human** Brain

Complete your human ES/iPS cell workflow with the new STEMdiff™ kits for directed differentiation to neurons, dopaminergic neurons or astrocytes



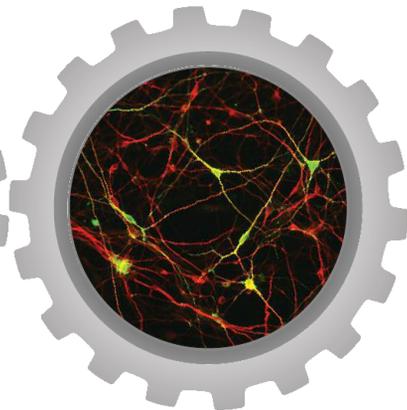
## Neurons

STEMdiff™ Neuron  
Differentiation and Maturation Kits



## Astrocytes

STEMdiff™ Astrocyte  
Differentiation and Maturation Kits



## Dopaminergic Neurons

STEMdiff™ Dopaminergic Neuron  
Differentiation and Maturation Kits

Learn more: [www.stemcell.com/STEMdiffNeural](http://www.stemcell.com/STEMdiffNeural)



Scientists Helping Scientists™ | [www.STEMCELL.COM](http://www.STEMCELL.COM)

FOR RESEARCH USE ONLY. NOT INTENDED FOR HUMAN OR ANIMAL DIAGNOSTIC OR THERAPEUTIC USES.  
STEMCELL TECHNOLOGIES INC.'S QUALITY MANAGEMENT SYSTEM IS CERTIFIED TO ISO 13485 MEDICAL DEVICE STANDARDS.

Copyright © 2015 by STEMCELL Technologies Inc. All rights reserved including graphics and images. STEMCELL Technologies & Design, STEMCELL Shield Design, Scientists Helping Scientists, STEMdiff are trademarks of STEMCELL Technologies Inc.