R&D SYSTEMS
a bio-techne brand

Get Your **Free Trial Vial** Today!

Consistency Is

N-2 MAX and N21-MAX Media Supplements for Neural and Stem Cell Cultures

Get your free trial vial | rndsystems.com/mediafreetrial

R&D SYSTEMS NOVUS BIOLOGICALS TOCRIS proteinsimple

bio-techne®
Global info@bio-techne.com bio-techne.com/find-us/distributors TEL +1 612 379 2956
North America TEL 800 343 7475 **Europe | Middle East | Africa** TEL +44 (0)1235 529449
China info.cn@bio-techne.com TEL +86 (21) 52380373
bio-techne.com

Cell Press Selections

iPSCs

Insights on the latest advances in stem cell reprogramming

Reprint supplement

CellPress
www.cell.com
CELLS10YRSIPSCS

biotechne®

Culturing stem cells can be tedious.

Let the good times roll with ProDots® Proteins.

ProDots®
PROTEINS

Rolls into cell culture media
Dissolves instantly
Eliminates aliquoting

Learn more | rndsystems.com/ProDots

R&D SYSTEMS®
a bio-techne® brand

R&D SYSTEMS
a bio-techne brand

Sample-Size Antibodies Now Available

Benefits of the Sample-Size:

> Test an antibody before purchasing a larger quantity.
> Economical option for small-scale experiments that require minimal antibody.
> Perfect for screening multiple antibodies to identify the best one for your target.

Learn more | rndsystems.com/SampleSizeAntibodies

R&D SYSTEMS
a bio-techne brand

Use Our Reagent Quality to Your Advantage

We've been the standard in quality research reagents for over 30 years. By choosing to use R&D Systems® reagents you've put your research at a distinct advantage over those who don't.

Our Quality Lets You...

Use reagents that work.

Not just reagents that work, reagents that are optimized and developed with industry-defining quality tests and standards. Be confident that your product isn't built with risky third party reagents.

Plan experiments with confidence.

Save yourself time and money by knowing that our products will perform the same today as they will 10 years from now.

Spend time answering real research questions.

Start focusing on discovery and stop worrying about your reagents. Our quality standards remove inconsistencies in reagent performance as a research variable.

Make an informed decision.

Don't gamble on the bioactivity of a protein or specificity of an antibody. Our performance data will tell you exactly what to expect from our products.

Know that we have your back.

Whether you're using R&D Systems® reagents or not, we are excited to make your research a success. Just call or email us with any of your questions.

Learn more | rndsystems.com/quality

Foreword

In the 10 years since their initial discovery by Shinya Yamanaka and Kazutoshi Takahashi, iPSCs have had a resounding impact on our understanding of mammalian development and human disease. The reviews and research articles in this *Cell Press Selections* edition on iPSCs offer a snapshot of the latest advances in this rapidly evolving field.

Over the past decade, innovations in iPSC research have converged with advances in seemingly unrelated areas, such as epigenetics, genome editing, and organoid technologies, and these synergies have fueled tremendous growth in this field. Applications of iPSC technology continue to expand their reach, with the first iPSC-based therapies already entering clinical testing.

The stem cell field is intently focused on filling in the gaps of how iPSCs are made and expanding their utility in disease modeling and therapeutic development. In recent years reprograming technology has also re-inspired analogous strategies for direct induction of alternate cell fates. The articles reprinted in this collection showcase current progress in these areas, all of which are ultimately aimed at improving the prospects for translation of iPSC technology to the clinic.

These articles represent only a small portion of the exciting research Cell Press has published and will publish on iPSCs, and we hope you'll visit www.cell.com on a regular basis to keep up with the latest iPSC news.

Finally, we are grateful to Biotechne for providing the generous support that allows us to bring this reprint collection to you.

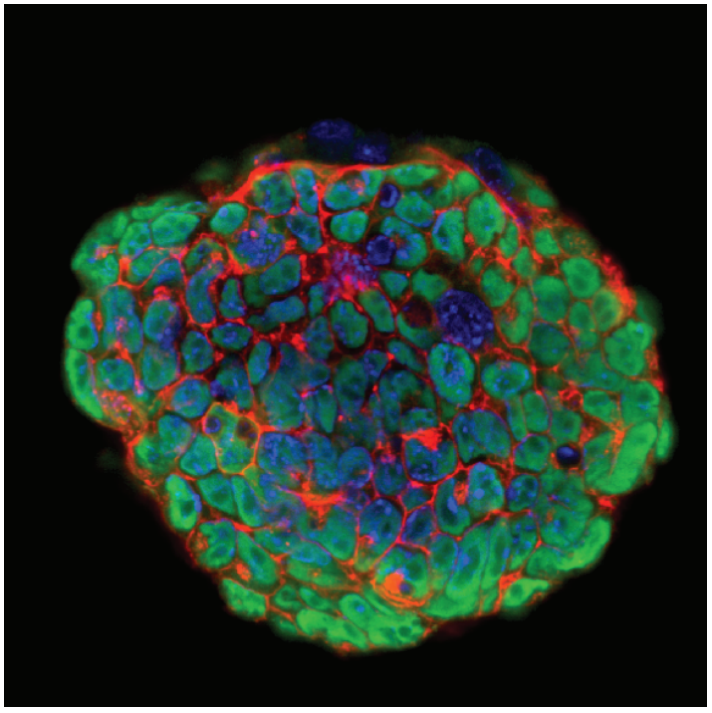
For more information about Cell Press Selections:
Gordon Sheffield
Program Director, Cell Press Selections
g.sheffield@cell.com
617-386-2189

R&D SYSTEMS
a bio-techne brand

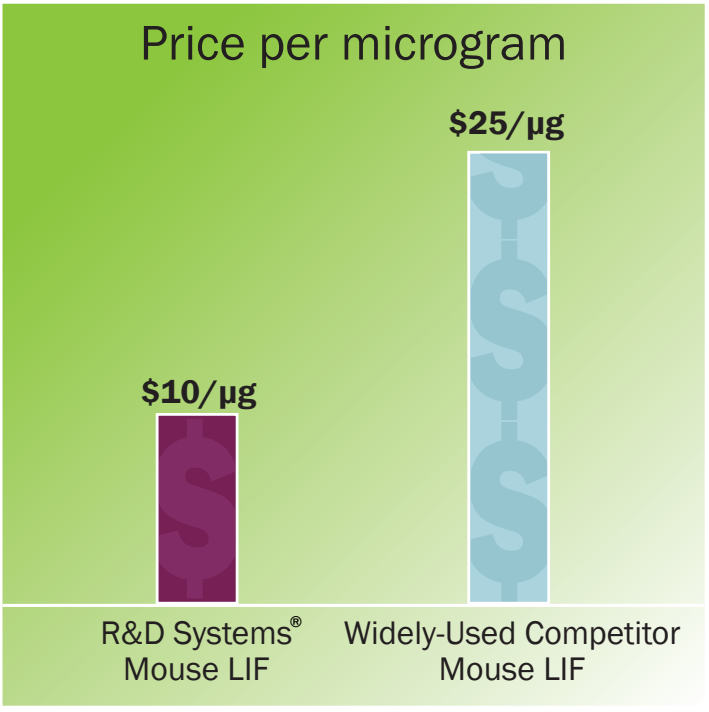
Recombinant Mouse LIF
Best Value on the Market

Leukemia Inhibitory Factor (LIF) is a staple media supplement for laboratories that culture pluripotent stem cells. With our newly-released Recombinant Mouse LIF, you get **more protein for less money**.

Excellent Bioactivity



Affordable Pricing



Learn more | rndsystems.com/mouseLIF

iPSCs

Insights into the next frontier in stem cell reprogramming

Forums and Reviews

From Dish to Bedside: Lessons Learned While Translating Findings from a Stem Cell Model of Disease to a Clinical Trial — *John McNeish, Jason P. Gardner, Brian J. Wainger, Clifford J. Woolf, and Kevin Eggan*

Programming and Reprogramming Cellular Age in the Era of Induced Pluripotency — *Lorenz Studer, Elsa Vera, and Daniela Cornacchia*

Complex Tissue and Disease Modeling using hiPSCs — *Robert Passier, Valeria Orlova, and Christine Mummery*

Transcriptional Control of Somatic Cell Reprogramming — *Yan Xu, Meng Zhang, Wenjuan Li, Xihua Zhu, Xichen Bao, Baoming Qin, Andrew P. Hutchins, and Miguel A. Esteban*

Articles

A XEN-like State Bridges Somatic Cells to Pluripotency during Chemical Reprogramming — *Yang Zhao, Ting Zhao, Jingyang Guan, Xu Zhang, Yao Fu, Junqing Ye, Jialiang Zhu, Gaofan Meng, Jian Ge, Susu Yang, Lin Cheng, Yaqin Du, Chaoran Zhao, Ting Wang, Linlin Su, Weifeng Yang, and Hongkui Deng*

Reprogrammed Stomach Tissue as a Renewable Source of Functional β Cells for Blood Glucose Regulation — *Chaiyaboot Ariyachet, Alessio Tovaglieri, Guanjue Xiang, Jiaqi Lu, Manasvi S. Shah, Camilla A. Richmond, Catia Verbeke, Douglas A. Melton, Ben Z. Stanger, David Mooney, Ramesh A. Shivdasani, Shaun Mahony, Qing Xia, David T. Breault, and Qiao Zhou*

A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells — *Courtney S. Young, Michael R. Hicks, Natalia V. Ermolova, Haruko Nakano, Majib Jan, Shahab Younesi, Saravanan Karumbayaram, Chino Kumagai-Cresse, Derek Wang, Jerome A. Zack, Donald B. Kohn, Atsushi Nakano, Stanley F. Nelson, M. Carrie Miceli, Melissa J. Spencer, and April D. Pyle*

Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency — *Davide Cacchiarelli, Cole Trapnell, Michael J. Ziller, Magali Soumillon, Marcella Cesana, Rahul Karnik, Julie Donaghey, Zachary D. Smith, Sutheera Ratanasirintrawoot, Xiaolan Zhang, Shannan J. Ho Sui, Zhaoting Wu, Veronika Akopian, Casey A. Gifford, John Doench, John L. Rinn, George Q. Daley, Alexander Meissner, Eric S. Lander, and Tarjei S. Mikkelsen*

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

SnapShot

Key Advances in hiPSC Disease Modeling — *Valeria Orlova and Christine Mummery*

CellPress

Cell Stem Cell
Forum

From Dish to Bedside: Lessons Learned While Translating Findings from a Stem Cell Model of Disease to a Clinical Trial

John McNeish,[1,*] Jason P. Gardner,[1] Brian J. Wainger,[2,3] Clifford J. Woolf,[3,4] and Kevin Eggan[3,5,6,*]
[1]Regenerative Medicine Discovery Performance Unit, GlaxoSmithKline, Cambridge, MA 02139, USA
[2]Departments of Neurology and Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, MA 02114, USA
[3]Harvard Stem Cell Institute, Cambridge, MA 02138, USA
[4]FM Kirby Neurobiology Center, Boston Children's Hospital and Harvard Medical School, Boston, MA 02115, USA
[5]Howard Hughes Medical Institute and Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
[6]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02139, USA
*Correspondence: john.d.mcneish@gsk.com (J.M.), eggan@mcb.harvard.edu (K.E.)
http://dx.doi.org/10.1016/j.stem.2015.06.013

While iPSCs have created unprecedented opportunities for drug discovery, there remains uncertainty concerning the path to the clinic for candidate therapeutics discovered with their use. Here we share lessons that we learned, and believe are generalizable to similar efforts, while taking a discovery made using iPSCs into a clinical trial.

Phenotypic assays using mature human cells derived from embryonic stem cells (ESCs) (McNeish, 2004) and induced pluripotent stem cells (iPSCs) (Grskovic et al., 2011) enable researchers to study disease-relevant phenotypes and provide a rich discovery tool for candidate therapeutics (Bellin et al., 2012). Furthermore, iPSCs have the potential for improving the identification of drug targets and candidate compounds as well as contributing to the optimized selection and stratification of trial participants. These applications could lead to more efficient clinical trials and reduced drug attrition during the development process.

While employing iPSCs is clearly attractive, the novelty of the approach has left its place in the drug discovery pipeline uncertain. On one hand, experiments with iPSCs can provide increased confidence in the relevance of a medicine to a patient's genetic makeup and human cellular physiology, suggesting that they may be more relevant to decision-making than existing animal models. On the other, the in vitro assays in which iPSCs are generally deployed leave reasonable questions concerning whether the mechanism of disease and candidate therapeutic identified have relevance in vivo. This natural and understandable tension leaves many engaged in iPSC research uncertain concerning the optimal path to the successful initiation of a clinical trial based on their discoveries. We reasoned it would therefore be useful to share our experiences and generalizable learning from translating a recent discovery made with iPSCs to an approved clinical trial for ALS patients (Wainger et al., 2014; Kiniskinis et al., 2014) (https://clinicaltrials.gov/, NCT# 02450552).

Recently, we showed that iPSC-derived motor neurons produced from ALS patients harboring *SOD1* mutations display reproducible, disease-relevant phenotypes (Kiskinis et al., 2014; Wainger et al., 2014.) These phenotypes included hyperexcitability with increased spontaneous action potentials and reduced survival (Wainger et al., 2014). More specifically, motor neurons from ALS patients displayed a reduced delayed-rectifier potassium channel activity. Further studies showed that this phenotype could be corrected by modulating the Kv7.2/3 class of potassium channels. Evidence that correcting motor neuron physiology was protective came through treatment with the approved antiepileptic and Kv7.2/3 potassium channel agonist ezogabine, which reduced neuronal excitability and improved cell survival (Wainger et al., 2014; Kiskinis et al., 2014). The first important lesson we learned from our desire to translate these studies was that the use of gene editing to correct the *SOD1* mutation, and with it the physiological changes we observed, was critical in building confidence in our findings.

To date, the only approved medicine available to ALS patients is Riluzole (Rilutek, Sanofi). The exact mechanisms by which Riluzole acts remain controversial, but have been proposed to include inhibition of Na^+ channels and glutamate activity (Bellingham, 2011). To date many additional mechanisms of drug action have been clinically evaluated in ALS, with seven compounds demonstrating positive phase 2 results. However, none of these seven have yet delivered positive findings in a pivotal phase 3 study. Thus, the identification of novel targets for ALS, like Kv7.2/3, worthy of being tested in the clinic remains sorely needed.

From the perspective of GlaxoSmith Kline (GSK), and likely other potential industry partners, a key concern with moving findings in ALS forward to the clinic has been the historically unreliable human translation of compounds evaluated in the *SOD1* mouse model (Perrin, 2014; Bellingham, 2011). In point of fact, Riluzole was brought to market prior to the development of this mouse model. Several hypotheses have been advanced concerning why discoveries from this mouse perform poorly in the clinic. One of the most reasonable is that given the genetic heterogeneity of ALS, it could be that features of disease in *SOD1* patients may not be central to disease progression in individuals harboring mutations in other genes. To determine how general

CrossMark

the relevance of our findings related to Kv7.2/3 were, we examined motor neuron physiology and response to ezogabine in a larger cohort of controls and patients harboring mutations in three distinct genes linked to ALS: *SOD1*, *C9orf72*, and *FUS* (Wainger et al., 2014). Our findings from these studies demonstrated similar physiological changes in the distinct patient classes, which were each rescued by ezogabine. Showing that the target phenotype and drug response were generalizable across patient forms was compelling enough to many key decision-makers to allow us to move forward without drug testing in the *SOD1* mouse model. Thus, while testing in animal models may be indispensable in some cases, we found that a higher premium was often placed on our data that supported the notion that the therapeutic approach proposed was valid to a broader portion of the patient population.

Another key factor enabling our clinical efforts was the relatively close alignment between the assays we carried out in vitro using iPSCs with emerging electrophysiological measures being made in the clinic. It has been shown using transcranial magnetic stimulation and threshold-tracking nerve conduction studies that ALS patients have a more excitable motor circuit and that the larger this change in physiology the worse a patient's prognosis (Bae et al., 2013). As a result, a clinical trial could be readily designed that employed clinical physiological measures to test the hypothesis that ezogabine might reduce motor circuit excitability in patients, much as it did in iPSC-derived motor neurons. In addition, we could propose to use physiological measures as pharmacodynamic biomarkers of ezogabine's impact on hyperexcitability, measuring the effect of two doses of ezogabine relative to placebo. Our experience suggests that carefully taking known in-patient biomarkers of disease into consideration when designing in vitro phenotypic assays with iPSCs can pay substantial dividends in later stages of translation. If we had not aimed our studies at understanding the mechanistic underpinning of a known patient phenotype, which could be readily measured clinically, we would have needed to pursue the time-consuming and costly process of developing a biomarker ourselves.

It is also worthy of note that our path to the clinic was aided in part by good fortune. A key driver of enthusiasm for trialing ezogabine in ALS patients, which was only partially in our control, was its well-known chemical and pharmacological properties as an approved drug. As an efficacious drug for epilepsy that acts through the opening of Kv7.2/3 channels, we could rely on preexisting clinical evidence that ezogabine engaged its target in the brain with therapeutic effect. In the absence of such data, expensive studies of compound toxicity, bioavailability, and in vivo half-life, likely coupled with additional cycles of chemical optimization and further testing, would have been needed before we could have considered a clinical trial. Our own experience suggests that many academic labs may find expertise in these pharmacological tests and medicinal chemistry either unavailable or unaffordable. Thus if it is the motivation of an academic investigator to rapidly test their iPSC-derived hypotheses clinically, it may be advisable for them to focus attention on libraries of already approved drugs or compounds that have made substantial progress in the clinic. Furthermore, if several molecules have been discovered with promise, it clearly would be most pragmatic to move forward using a compound with stronger pharmacological data even if another showed marginally better performance in vitro.

In short, we found that the strong scientific foundation we had produced using iPSCs, a clear clinical question that could be tested using an established biomarker, and a compound with strong pharmacological properties were each essential pieces in the puzzle of organizing partnerships between academics, clinicians, patient advocacy groups, and industry that were needed to mount a clinical trial. The basic research supported by Boston Children's Hospital, Harvard, HHMI, Target ALS, the ALS Association (ALSA), and GSK enabled the assembly of a consortium to fund the clinical study, which included the Harvard Stem Cell Institute (HSCI), Massachusetts General Hospital (MGH), GSK, and ALSA. It was our experience that, as has recently been suggested (Saha and Hogle, 2014), when a large federation of scientists, physicians, and drug development experts can be assembled, many of the clinical and regulatory challenges to mounting a clinical trial are more rapidly overcome.

Another critical accelerator of our study was our ability to work with a preexisting clinical research team. The Northeast ALS (NEALS) Consortium together with the MGH Neurological Clinical Research Institute provided robust yet adaptable infrastructure for the rapid translation to a trial of the scale we proposed. Working with a preexisting clinical network allowed us to streamline the processes of developing the clinical trial protocol, trial contracts, and essential measures for subject safety monitoring and obtaining the needed FDA IND exemption for testing Retigabine in ALS patients. Investigators interested in advancing toward the clinic would be well advised to familiarize themselves with similar clinical consortia operating in their indication of interest and then to build strong enabling relationships with such groups. If such a group does not exist, our experience suggests that energy expended to help organize one would be well allocated.

Another important and likely generalizable lesson we learned while preparing our trial design was that enthusiasm from our funding partners for making the needed investment was increased by incorporating provisions for collecting additional samples that would fuel future basic research on ALS. A key component of the ezogabine study is to derive hiPSC lines from participants. This type of parallel study represents a unique opportunity afforded by hiPSC research. It seems highly probable that similar approaches would be viewed as valuable by funders if implemented in other clinical investigations for which highly reproducible hiPSC-based disease models have been developed. We found that adopting this strategy was motivating to our entire consortium, which was eager to see whether the patients enrolling in the clinical trial reflected the biology of those patients that originally drove initiation of the study. The future availability of this iPSC resource to the community will mean that our clinical trial will become "evergreen." It will allow investigators to study correlations between patient outcomes and in vitro phenotypes in motor neurons or other cell types. In addition, with the iPSC resource in hand, variation in drug response could be investigated mechanistically. If the trial is a success,

the iPSCs could become a resource for evaluating compounds emerging from future high-throughput screens for novel Kv7.2/3 agonists. If our trial fails, these iPSC could nonetheless be useful for attempting to further stratify the patient population, for testing additional therapeutics, or for trying therapeutic combinations. Due to the substantial interest in the iPSC resources, a key feature of the partnership agreement that serves as a foundation of our trial is that it makes these stem cells available for both basic and commercial research following completion of the trial.

The combination of careful execution and good fortune outlined above placed us in position to file an Investigator-Sponsored IND-exemption request, which has now been approved by the FDA to evaluate ezogabine in a phase 2 multicenter, randomized, double-blind, placebo-controlled study of neuronal excitability in ALS subjects. Secondary outcomes of the study will include tolerability and safety of ezogabine in ALS patients. This trial will be conducted at 12 U.S. NEALS Consortia sites. Of note, this program progressed from initial discovery to phase 2a study initiation in less than 2 years. We hope that outlining what we feel were the key factors playing important roles in the successful translation of ezogabine, both fortuitous and carefully calculated, will be valuable to those interested in taking their own discoveries made with iPSCs to the clinic.

We recognize that the type of trial we are undertaking and propose could be useful in many other disease contexts, though it is not without its complications and potential limitations. For example, there are substantial challenges to reducing technical variability among more than 100 iPSC lines made from patient samples collected at a dozen sites. Even with iPSCs in hand, improved processes will be needed to efficiently and reproducibly differentiate, culture, and analyze motor neurons from this large number of patients. Still, we are optimistic that such challenges can be overcome and that additional clinical trials will emerge from the many studies of disease-relevant cell types being made from hiPSCs (Grskovic et al., 2011). The ever-expanding reporting of clinically relevant phenotypes in hiPSC disease models, as well as the pharmacological correction of pathologic disease features in these cells, is creating an exciting environment for the development of new medicines. As the reproducibility and robustness of stem cell technologies continues to improve so too will their utility in nominating hypothesis for clinical testing. We believe that interest in hiPSC technologies for applications in drug research and development will continue to grow and that these cells will eventually serve as surrogates for many clinical phenotypes and perhaps even provide a new form of companion diagnostic.

ACKNOWLEDGMENTS

We are grateful to our many colleagues at GlaxoSmithKline (GSK) and The Harvard Stem Cell Institute for helpful discussions that contributed to the experiences and lessons we share here. J.M. and J.G. are employees of GSK. K.E., C.W., and B.W. have received sponsored research support from GSK.

REFERENCES

Bae, J.S., Simon, N.G., Menon, P., Vucic, S., and Kiernan, M.C. (2013). J. Clin. Neurol. *9*, 65–74.

Bellin, M., Marchetto, M.C., Gage, F.H., and Mummery, C.L. (2012). Nature Rev. *13*, 713–726.

Bellingham, M.C. (2011). CNS Neurosci. Ther. *17*, 4–31.

Grskovic, M., Javaherian, A., Strulovici, B., and Daley, G.Q. (2011). Nat. Rev. Drug Discov. *10*, 915–929.

Kiskinis, E., Sandoe, J., Williams, L.A., Boulting, G.L., Moccia, R., Wainger, B.J., Han, S., Peng, T., Thams, S., Mikkilineni, S., et al. (2014). Cell Stem Cell *14*, 781–795.

McNeish, J. (2004). Nat. Rev. Drug Discov. *3*, 70–80.

Perrin, S. (2014). Nature *507*, 423–425.

Saha, K., and Hogle, L.F. (2014). Cell Stem Cell *14*, 559–560.

Wainger, B.J., Kiskinis, E., Mellin, C., Wiskow, O., Han, S.S., Sandoe, J., Perez, N.P., Williams, L.A., Lee, S., Boulting, G., et al. (2014). Cell Rep. *7*, 1–11.

Cell Stem Cell
Perspective

Programming and Reprogramming Cellular Age in the Era of Induced Pluripotency

Lorenz Studer,[1,*] Elsa Vera,[1] and Daniela Cornacchia[1]
[1]Developmental Biology and Center of Stem Cell Biology, Memorial Sloan Kettering Cancer Center, New York, NY 10003, USA
*Correspondence: studerl@mskcc.org
http://dx.doi.org/10.1016/j.stem.2015.05.004

The ability to reprogram adult somatic cells back to pluripotency presents a powerful tool for studying cell-fate identity and modeling human disease. However, the reversal of cellular age during reprogramming results in an embryonic-like state of induced pluripotent stem cells (iPSCs) and their derivatives, which presents specific challenges for modeling late onset disease. This age reset requires novel methods to mimic age-related changes but also offers opportunities for studying cellular rejuvenation in real time. Here, we discuss how iPSC research may transform studies of aging and enable the precise programming of cellular age in parallel to cell-fate specification.

Pluripotent stem cells (PSCs) are characterized by their ability to extensively self-renew and differentiate into all the cell types of the body. During normal development, human pluripotency is restricted to the earliest stages of somatic and germ cell development, stages that can be captured in human embryonic stem cells (ESCs) (Thomson et al., 1998) and embryonic germ cells (EGCs) (Shamblott et al., 1998). The discovery of induced PSCs (iPSCs) (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007) was the realization of a longstanding dream in biology, namely to access pluripotency starting from somatic cells of an adult organism. Human iPSC technology has opened up new frontiers in regenerative medicine and human disease modeling. Protocols for the directed differentiation of human PSCs (hPSCs) have been developed to generate an increasingly broad repertoire of differentiated human lineages, and a future is foreseeable wherein any human cell type can be generated in vitro—on demand and at scale. Although technologies for the programming and reprogramming of cell fate have evolved rapidly, our ability to control the maturation state and age of resulting pluripotent-derived lineages remains rudimentary at best. In fact, there is general agreement that human-pluripotent-derived lineages exhibit the properties of fetal-stage cells such as in the case of hPSC-derived neural, cardiac, or pancreatic lineages. Importantly, such fetal-like properties are observed in iPSC-derived lineages independent of the age of the initial somatic cell donor. The embryonic-like nature of hPSC-derivatives represents a potential barrier to the use of PSCs, which motivates the development of strategies for directing cellular age in vitro, in particular for applications in human disease modeling. On the other hand, those findings raise the intriguing question whether the reprogramming process resets not only cell fate (i.e., from specified to pluripotent) but also the chronological age characteristic of the donor cell population.

Here, we will discuss recent studies that address questions of age in pluripotent stem cells. Those include both efforts to study the apparent rejuvenation process during reprogramming as well as the development of techniques to artificially induce age in iPSC-derived lineages for modeling late-onset disorders. The long-term goal is to reliably program and re-program cellular age independently of cell fate and thereby recrate specific cell types of any age (e.g., 80-year-old neurons, 20-year-old pancreatic cells, or 40-year-old heart cells).

Current Strategies for Studying Aging

According to the World Health Organization, global life expectancy will increase from 48 years of age in 1950 to 73 years by 2025 (http://www.who.int/whr/1998/en/whr98_en.pdf?ua=1), and, in many developed countries, the average life expectancy is already >80 years and rising. The associated worldwide increase in the incidence of age-related disorders such as Alzheimer's (AD) and Parkinson's disease (PD) is expected to cause enormous social, economic, and medical challenges. The mounting problem of an aging society has triggered a race to find novel strategies to treat age-related disorders. A more radical proposition is the search of the "youth elixir," hence the development of techniques that would actively rejuvenate the human body. However, if such strategies were to succeed, they would likely further extend overall human lifespan with unknown consequences for society. In fact, the pursuit of preventing human aging has been touted by some as "egocentric efforts of rich people to live longer" (http://www.reddit.com/r/IAmA/comments/2tzjp7/hi_reddit_im_bill_gates_and_im_back_for_my_third/). On the other hand, anyone who has witnessed the suffering of loved ones can understand the motivation for treating or even better preventing serious age-related disorders and achieving longer and healthier lives. Independently of whether such efforts are considered as "chasing immortality" or simply as addressing the emerging challenge of potentially billions of elderly people facing aging-related consequences, the question of age and longevity is a fascinating scientific problem. The main challenge for tackling this issue is the identification of a unanimously accepted cause of aging.

Modern theories of aging can be classified into programmed versus damage or error-induced mechanisms. Theories of programmed aging argue for a developmentally controlled program that drives aging through the regulation of tissue homeostasis as well as repair and defense responses. In contrast, damage- or error-induced aging theories emphasize the temporal or stress-induced accumulation of damage caused by reactive oxygen species (ROS), cross-linked macromolecules,

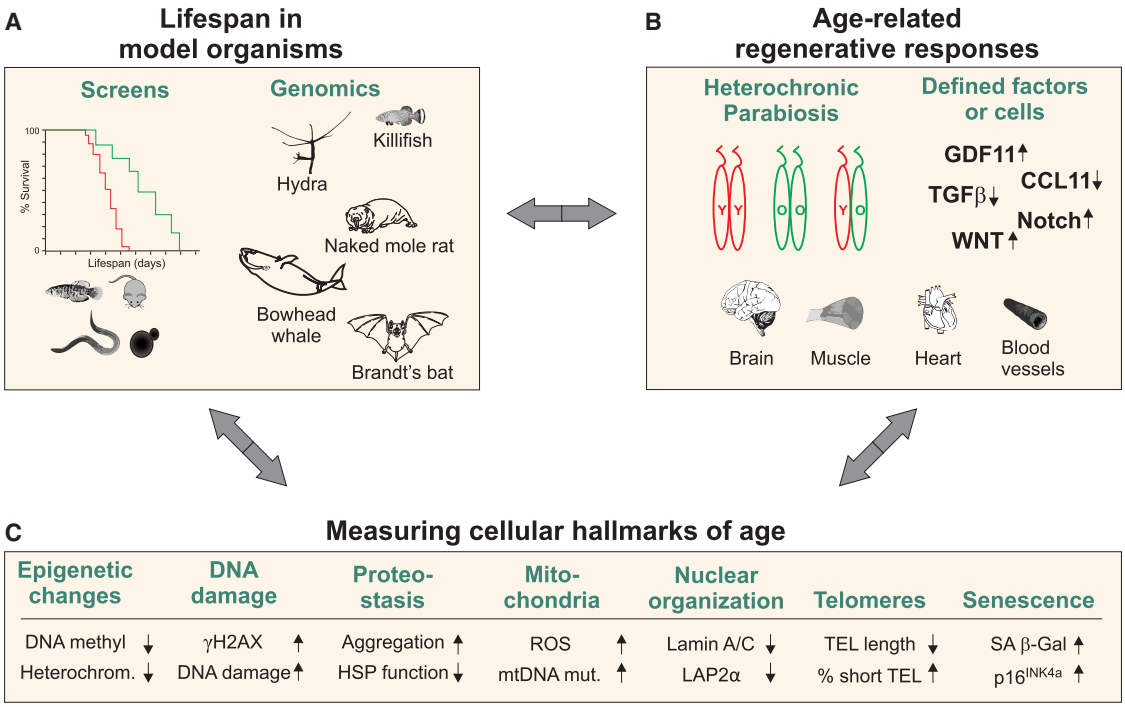


Figure 1. Traditional Approaches and Cellular Hallmarks of Aging

(A) Screens in model organisms were crucial for the identification of signaling pathways that increase organismal lifespan. Such pathways include insulin/IGF1 signaling (IIS) and nutrient-sensing pathways among others. An alternative strategy is focused on the identification of genomic determinants of longevity through the study of exceptionally long-lived or short-lived animals. Organisms with unusually long lifespans include the naked mole rat, the bowhead whale, or the Brandt's bat. Examples of short-lived animals of interest in aging research include the killifish and hydra.

(B) Heterochronic parabiosis has become an important tool for the discovery of systemic factors capable of restoring a youthful regenerative capacity in aged tissues through shared blood circulation with a young animal. Blood-borne factors capable of mediating the effects of parabiosis have been identified including factors promoting skeletal muscle, blood vessel, brain, and heart regeneration.

(C) An additional approach to study aging represents the investigation of cellular alterations that accompany organismal aging. Cellular hallmarks of age that can be readily measured can affect virtually any compartment of the cell. Additional hallmarks such as altered intracellular communication and defects in nutrient sensing are not shown here because they are difficult to quantify in vitro.

DNA damage, or altered energy metabolism as the main drivers. Aging has also been explained by some as direct consequence of the "intrinsic thermodynamic instability of most complex biological molecules" (Hayflick, 2007) leading to progressive loss of molecular fidelity. Although multiple mechanisms most likely contribute to the aging process, consensus on the key molecular and biochemical cause of organismal aging remains elusive (Jin, 2010; Liu et al., 2012a).

Various approaches have been proposed for studying and manipulating the rate of aging. Traditionally, those have included the use of model organisms that enable pharmacological and genetic screens for lifespan extension and the search for systemic factors able to promote regenerative responses aimed at repairing age-dependent tissue damage. Several questions are raised by these studies, such as whether prolonged lifespan and improved systemic fitness are reflective of a reversal of organismal age at the level of tissues and single cells or whether enhanced longevity and regeneration can occur independently of a true reset of cellular age. Furthermore, it remains unclear whether the manifestations of organismal aging at the levels of organs and tissues are directly correlated to aging hallmarks that can be measured at the level of individual cells.

Longevity Pathways in Model Organisms

To gain insight into the molecular mechanisms of aging, a major focus has been set on genetically tractable model organisms such as yeast, worms, and flies, which allow the direct testing of the impact of genetic or pharmacological interventions on overall lifespan (Guarente and Kenyon, 2000) (Figure 1A). Genetic screens in invertebrates and subsequent studies in mammals have identified the insulin/IGF1 signaling (IIS) pathway, and the equivalent mammalian somatotrophic axis as the most evolutionarily conserved longevity-controlling pathway. In fact, every component of the IIS pathway that connects nutrient sensing to gene-regulatory mechanisms was shown to directly affect organismal lifespan. These factors comprise, aside from the insulin/IGF-1 receptor and its downstream effectors, other nutrient-sensing regulators, such as mTOR, AMPK, and sirtuins (Kenyon, 2010). Interfering with the IIS pathway is believed to mimic the low-energy state brought about by caloric restriction (CR), currently the most robust intervention known to prolong lifespan across species. CR involves a reduction of food intake while avoiding malnutrition and is associated with lower IIS signaling. This regime was shown to extend lifespan, improve stem cell-based regeneration, and delay the onset of age-related diseases such as cancer, diabetes or neurodegenerative disease, thereby prolonging the so called "healthspan" (Cerletti

CellPress

et al., 2012; Cheng et al., 2014; Colman et al., 2009; Kenyon, 2010; Mattison et al., 2012). Consistent with the effect of low nutrient intake on longevity, genetic screens in the worm demonstrated that loss of function of the insulin receptor *daf-2* can lead to a dramatic extension of lifespan that is dependent on the activity of the FOXO forkhead transcription factor *daf-16*. Furthermore, combining *daf-2* loss of function with the removal of reproductive organs yields an astonishing increase in worm lifespan up to 6-fold from 20 days to over 120 days (Arantes-Oliveira et al., 2003).

Studies in yeast were first to identify sirtuins as major prolongevity factors. Sirtuins are NAD^+-dependent deacylases and mono-ADP-ribosyl transferases and serve as nutrient sensors that link metabolic changes to stress response and epigenetic regulation. Sirtuins were shown to mediate some of the beneficial effects of CR on lifespan extension in part by boosting cellular stress defense (Guarente, 2013). Overexpression of Sirt1 in mice did not generally improve lifespan but resulted in an increased healthspan (Herranz and Serrano, 2010). In contrast, male (but not female) mice overexpressing Sirt6 showed a significant extension in lifespan (Kanfi et al., 2012). Alternative pathways reported to promote longevity and improve overall fitness include the overexpression of telomerase in the context of a cancer-resistant mouse (Tomás-Loba et al., 2008). Finally, several currently FDA-approved drugs were also shown to increase lifespan in model organisms, including the type II diabetes drug metformin (Cabreiro et al., 2013; Martin-Montalvo et al., 2013) as well as several anticonvulsive agents (Evason et al., 2005). On the other hand, strategies have also been developed that lead to decreased lifespan and accelerated aging, such as impaired DNA repair by ATR deficiency (Murga et al., 2009) or telomerase-null mice (Blasco et al., 1997; Lee et al. 1998). An unresolved question remains whether longevity pathways identified to date act exclusively by slowing down the aging process or whether some of those manipulations indeed achieve a true rejuvenation of previously old cells. It will be interesting to revisit this question in the light of the detailed cellular and molecular hallmarks of age available (López-Otín et al., 2013).

Another interesting toolset for studying the genetic drivers of organismal aging is provided by animals with extraordinary long or unusually short lifespan (Figure 1). For example, naked mole rats have the body size of a mouse but a life expectancy that is 10-fold higher, reaching more than 30 years of age in captivity. Genetic studies are underway to define factors responsible for such longevity in moles, bats, and other long-lived animals (Kim et al., 2011a). Studies in hydra are intriguing as an example of an organism that is short-lived despite lacking evidence of cellular senescence. In conclusion, the ability to modulate lifespan in model organisms as well as the existence of extremely long- and short-lived organisms make a strong argument that aging might not simply be the sum of accumulated damage throughout life but indicate the existence of genetic programs affecting the rate of aging.

Systemic Factors Affecting Age-Related Response

A key question in aging research is the relative contribution of cell-autonomous versus non-autonomous, systemic factors on age-related phenotypes. There has been renewed interest in this issue recently with the use of parabiosis. Parabiosis was first described more than 150 years ago (Bert, 1864) and developed into a reliable technology for surgically linking the blood circulation of two animals. Heterochronic parabiosis involves linking animals of different age and has become an important tool to study the impact of blood-derived factors present in young animals on the age-dependent regenerative capacity of old animals (Conboy et al., 2013) (Figure 1B). Those experiments demonstrated improved regenerative potential of skeletal muscle cells (Brack et al., 2007), improved neural precursor cell proliferation, and cognitive function (Villeda et al., 2011) as well as improved heart function (Loffredo et al., 2013) in old animals following parabiosis with a young host. An important endeavor has been the molecular identification of factors responsible for this improved regenerative response. Several pathways have been implicated including factors modulating Notch (Conboy et al., 2005), WNT (Brack et al., 2007), or chemokine (Villeda et al., 2011) signaling. There has been particular recent excitement about the identification of the TGFβ-ligand GDF11, capable of enhancing regenerative responses across several tissues including skeletal muscle, heart, and brain (Loffredo et al., 2013; Sinha et al., 2014) (Figure 1B). The identification of blood-derived factors presents a promising avenue for defining candidate therapeutic targets that improve tissue regeneration. One interesting point is the limited convergence of the results obtained in parabiosis studies with those obtained in longevity studies. In this light, it will be interesting to determine whether improved regeneration through blood-borne molecules reflects actual rejuvenation of the host tissue and whether long-term manipulation of those pathways can extend lifespan. To this end, it would be of particular interest to assess whether GDF11 or any other secreted factor can reset molecular hallmarks of aging or trigger improved regenerative responses independent of cellular rejuvenation.

Cellular Hallmarks of Aging

There is no consensus on the specific molecular and biochemical cause of organismal aging. Nevertheless, several candidate mechanisms have been proposed, ranging from cellular processes like accumulation of genetic or epigenetic changes to telomere shortening, mitochondrial dysfunction, or the accumulation of damaged proteins to systemic age-dependent alterations such as deregulated nutrient sensing, impaired intercellular communication, and the exhaustion of tissue-specific stem cells (López-Otín et al., 2013). Independently of whether any of those candidate mechanisms are primary drivers of aging or mere side effects, they represent a set of molecular and cellular hallmarks that can be used as surrogate markers to measure cellular age across cell and tissue types. For the scope of this review, we will focus on aging hallmarks that manifest at the cellular level, which can provide simple quantitative assays for measuring age in vitro (Figure 1C), though other aging hallmarks such as deregulated nutrient sensing and altered intracellular communication should be considered as well. The use of a reliable marker platform is a prerequisite for extending aging studies beyond organismal longevity or tissue regeneration. In particular, it enables studies that monitor the process of rejuvenation during iPSC induction and facilitates ongoing efforts to induce age in iPSC-derived lineages.

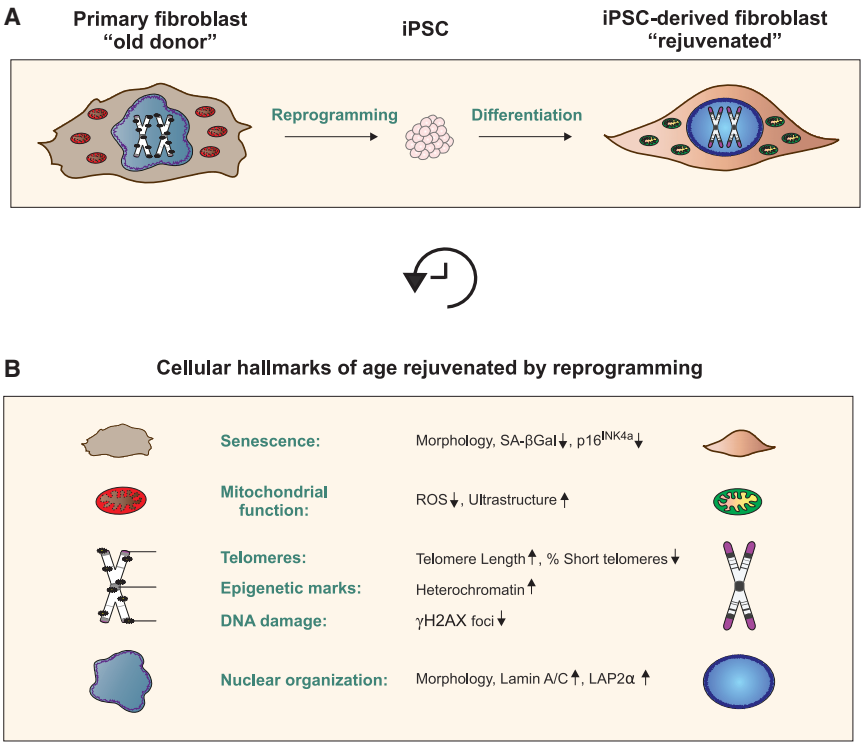


Figure 2. Phenotypic Rejuvenation during iPSC Induction

(A) Fibroblasts from an aged donor can be reprogrammed back to pluripotency and further differentiated into iPSC-derived fibroblasts. During this process, many cellular aging hallmarks are reset to a young-like state ranging from cellular senescence to nuclear morphology.

(B) Legend for each of the age-related hallmarks reset during reprogramming.

Pluripotency and Age Reset

The ability to induce pluripotency in adult somatic cells has been a major goal in biology. Somatic cell nuclear transfer (SCNT) studies were first to demonstrate the feasibility of epigenetically resetting a donor nucleus through transplantation into an enucleated egg. Starting with the pioneering studies in the frog by Gurdon (1962), many subsequent findings have confirmed that SCNT can result in the birth of healthy animals derived from an adult somatic cell nucleus. The most dramatic example for the power of SCNT is represented by Dolly the sheep (Campbell et al., 1996), whose birth immediately raised the question of whether the process of epigenetic reprogramming had reset the fate as well as the age of the original donor cell. Some initial controversy on whether age was indeed reset after nuclear transfer was based on analysis of telomere length in the cloned animal. However, the premature death of Dolly was likely unrelated to incomplete rejuvenation, and most follow-up studies indicate that telomere length is fully reset after nuclear transfer (Lanza et al., 2000).

SCNT provided an initial proof-of-concept that it is possible to reset the age-status of a somatic donor cell, but it did not yield a tractable platform for routinely studying mechanisms of reprogramming. The discovery of iPSCs by Shinya Yamanaka has been the "game changer" required to make such studies feasible. PSCs are characterized by the expression of pluripotency markers, indefinite self-renewal, and their ability to contribute to all the three germ layers of the embryo. However, many other features of pluripotent stem cells, such as high telomerase activity, mitochondrial status, and aspects of their epigenetic state, reflect features of young cells that are not restricted to the pluripotent state. Therefore, a key prediction is that the hallmarks of cellular aging present in adult fibroblasts (e.g., short telomeres, loss of heterochromatin, and impaired mitochondrial function) are reset during iPSC induction. Indeed, bona fide iPSCs are largely indistinguishable from ESCs in the expression of age-related markers such as telomere length (Marion et al., 2009; Suhr et al., 2009) or mitochondrial function (Prigione et al., 2011; Suhr et al., 2010). However, adult somatic cells, such as fibroblasts or adult blood-derived lineages, are differentiated cell types highly distinct from hPSCs. Thus, a direct comparison of adult somatic cells to their iPSC counterparts does not satisfactorily address whether age-related changes reflect a true rejuvenation or rather a simple switch in cell identity. Therefore, it was critical to confirm whether iPSC-derived lineages, fate-matched to their somatic equivalents, retain signs of rejuvenation. Initial studies in iPSC-derived fibroblasts were focused on a limited set of hallmarks that included mitochondrial stress response and cellular senescence (Lapasset et al., 2011). We have recently reported on the use of a more complete panel of age-associated cellular markers suitable for reliably distinguishing fibroblasts derived from young versus aged donors (Miller et al., 2013). Those include: telomere length, mitochondrial function, DNA-damage response, global loss of heterochromatin, nuclear lamina-associated changes, and increase in the fraction of senescent cells. Remarkably, every single age-related parameter expressed in aged donor fibroblast population was reset after iPSC induction and differentiation into iPSC-derived fibroblast-like cells (Miller et al., 2013) (Figure 2). Another interesting case is the rejuvenation of antigen-specific T cells following reprogramming into T-cell-derived iPSCs and further differentiation into iPSC-derived T cells (Nishimura et al., 2013; Themeli et al., 2013; Vizcardo et al., 2013). Such a strategy could counteract immune cell exhaustion such as in the context of adoptive immunotherapy.

Despite clear evidence on phenotypic rejuvenation during reprogramming, several questions remain unsolved. For example, in the case of reprogramming aged donor cells, it is possible that rejuvenated iPSCs selectively arose from a small fraction of phenotypically young cells within the original fibroblast population. The use of clonal studies to address this point is challenging, given that fibroblast-derived clones may re-establish heterogeneity of age-related marker expression. Hence, time-lapse studies and use of real-time reporters for age-related markers are needed to address the possibility of cellular selection during reprogramming. Another key question is whether

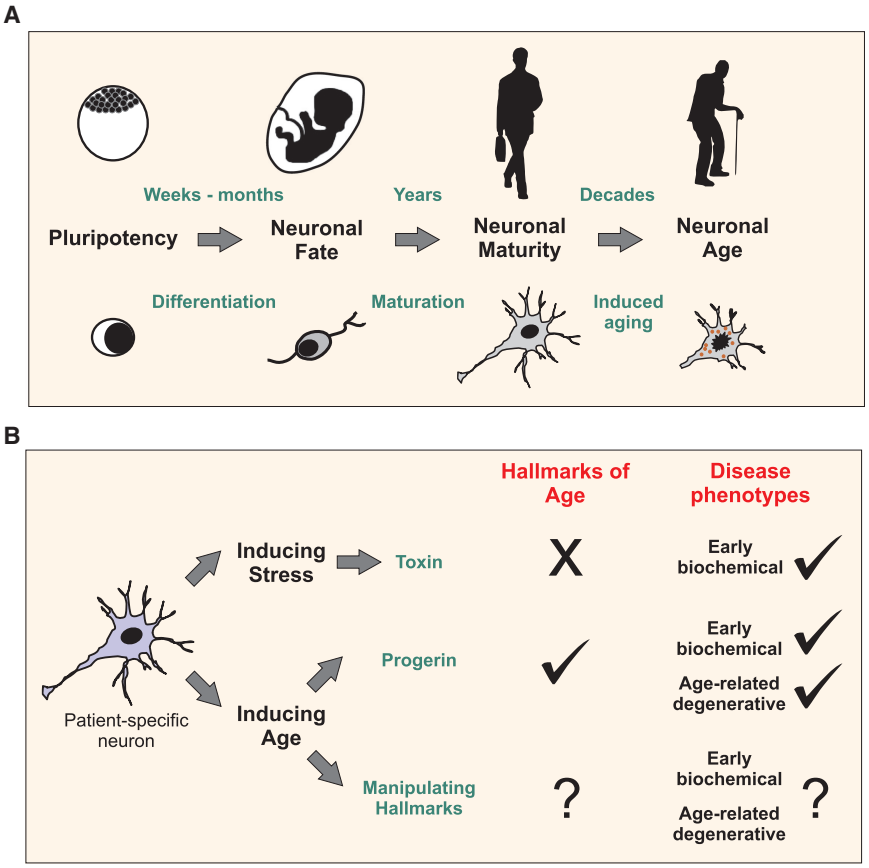


Figure 3. Manipulating Cellular Age in Pluripotent-Derived Cells

(A) The three stages of developmental timing: cell-fate specification, maturation, and aging. In vitro differentiation of hPSC to neuronal lineages follows a temporal program reminiscent of brain development in vivo. PSC-derived neurons commonly represent immature fetal-stage cells, and novel tools are under development to improve functional maturation and trigger expression of age-related markers in pluripotent-derived cells.

(B) Modeling late-onset disorders in patient-specific cells may require the induction of stress or age-related markers to uncover disease related phenotypes. Most studies so far focused on challenging cells with toxic stressors. We postulate that such paradigms are suitable to reveal disease-related susceptibility phenotypes that reflect early biochemical changes in the disease process "early biochemical." We propose "inducing aging" as an alternative strategy to "inducing stress" that may reveal additional disease phenotypes. As opposed to inducing stress, strategies aimed at inducing age are thought to trigger expression of the various molecular hallmarks of age. Induced aging strategies draw inspiration from naturally occurring mutations known to cause premature aging, such as the recently reported method based on the expression of progerin. An alternative strategy that has not yet been implemented for human disease modeling could involve the direct manipulation of one or a subset of molecular hallmarks of aging with the goal of eliciting cellular age in a more physiological manner.

rejuvenation of cellular phenotypes is accompanied by a complete reset of molecular markers of age. In particular, reports suggest that cells retain epigenetic memory of age at the level of DNA methylation (Kim et al., 2010; Kim et al., 2011b). It remains to be determined whether such memory is indeed retained over extended passages and can be reconciled with separate findings on the rapid loss of epigenetic memory in iPSCs (Polo et al., 2010).

Cellular Age in Pluripotent-Derived Cells

Directed Differentiation, Maturation, and Age—Examples of the Nervous System

There has been extensive progress in the directed differentiation of hPSCs to access a broad range of differentiated cell types. An area of particular interest is the derivation of lineages of the CNS for applications in regenerative medicine (Steinbeck and Studer, 2015) and human disease modeling (Bellin et al., 2012). Studies in the nervous system also highlight some of the remaining challenges of directed differentiation, such as the protracted timing of neuronal and glial fate specification, maturation, and age (Figure 3A). From a chronological perspective, cell specification, maturation, and aging represent contiguous temporal stages for each cell type. However, the specific cellular processes and underlying mechanisms characteristic of each stage are highly distinct. Although acquisition of cell fate is characterized by specific transcriptional programs and marker expression, cell maturation entails the development of functional features, such as electrical activity and synaptic connectivity, in the case of neuronal maturation. In contrast, aging is associated with a gradual decline in neuronal structure and function. Hallmarks of neuronal aging are often difficult to distinguish from disease-related degenerative processes. A convenient age-related marker is the pigment neuromelanin, which characterizes adult and aged dopamine neurons in the midbrain but is absent in matched cells at fetal and postnatal stages (Mann and Yates, 1974).

Challenges in neuronal fate specification include the generation of hPSC-derived cortical neurons. Current state-of-the-art protocols require up to 100 days of differentiation to generate upper layer cortical neurons (Espuny-Camacho et al., 2013; Shi et al., 2012). Cortical interneurons subtypes such as parvalbumin (PV+) neurons require even more extended time lines of in vitro or in vivo differentiation (Maroof et al., 2013; Nicholas et al., 2013). Similarly, the generation of glial lineages such as astrocytes (Krencik et al., 2011) or myelinating oligodendrocytes is highly protracted requiring differentiation periods ranging from 3–6 months (Wang et al., 2013). Lengthy differentiation protocols make hPSC studies laborious, costly, and difficult to standardize. Transcriptome studies show that neural human iPSC-derived tissues typically map with the first trimester of human development (Mariani et al., 2012). Interestingly, data obtained from 2D versus 3D organoid culture systems do not show major differences (Kadoshima et al., 2013; Lancaster et al., 2013) suggesting that mimicking normal tissue

architecture is not sufficient to advance the differentiation clock.

The timing of neuronal maturation presents additional challenges beyond cell fate specification. For example, grafted hPSC-derived midbrain dopamine neurons require 4-5 months of in vivo maturation to induce functional rescue in a rat model of Parkinson's disease (Kirkeby et al., 2012; Kriks et al., 2011). Using the same xeno-transplantation paradigms but starting with mouse ESC-derived DA neurons, behavioral recovery is typically achieved in 4-5 weeks (Barberi et al., 2003; Kim et al., 2002). Such species-specific differences in timing of dopamine neuron maturation is mirrored in studies using primary, stage-matched DA neurons from mouse, pig or human embryos (Isacson and Deacon, 1997). The most straightforward interpretation of those studies is the presence of a species-specific, cell autonomous "clock" that controls the rate of maturation even when exposed to the same host environment in the adult rodent brain. Protracted maturation is also observed in midbrain dopamine neurons in vitro as well as in many other neuronal cell types such as peripheral sensory or cortical neurons. Interestingly, cortical reprogramming strategies based on the use of lineage specific transcription factors, show similar species-specific differences in neuronal maturation between mouse and human cells (Pang et al., 2011; Vierbuchen et al., 2010).

There has been some success in speeding up the timing of neuronal cell-fate acquisition based on a combinatorial small-molecule screen that resulted in the rapid induction of human sensory neuron fates (Chambers et al., 2012). However, it remains to be determined whether the same strategy is applicable for the rapid derivation of other neuron types. Strategies for triggering neuronal maturation have been even more challenging. Improved cell-culture media formulations (Bardy et al., 2015) and the use of astrocyte co-culture can accelerate electrophysiological maturation of hPSC-derived neurons. However, there is no evidence that those conditions accelerate other aspects of neuronal maturation, such as the acquisition of late-stage neuronal markers including expression of parvalbumin (PV) in cortical interneuron cultures. Interestingly, the induction of PV was observed after co-culture of hPSC-derived interneurons with mouse excitatory neurons (Maroof et al., 2013). Those data suggest that activity and synaptic connectivity may contribute to PV expression. One interesting alternative for driving neuronal fate specification and maturation in hPSCs is the ectopic overexpression of transcription factors such as Ngn2 (Zhang et al., 2013). Under those conditions, hPSCs rapidly adopt the fate of functionally active neurons. However, full maturation of those neurons remains a challenge, and it remains unclear whether those Ngn2-induced neurons reflect a relevant human neuron subtype or an artificial neuron without a direct counterpart in the human brain. Therefore, better strategies are needed to trigger the maturation of hPSC-derived lineages in the CNS and beyond.

Disease Modeling via iPSCs

Human iPSC technology represents a powerful tool for both mechanistic studies and therapeutic development in human genetic disorders. Differences in patient-specific versus control iPSC-derived cells are defined as candidate disease phenotypes. However, the choice of the appropriate control iPSC lines for such studies is challenging given genetic heterogeneity in the general population. The use of isogenic, gene-corrected lines can increase confidence in the specificity of disease phenotypes. A major challenge remains the demonstration that a given phenotype is indeed disease-relevant, a process that may include lengthy follow-up studies in model organisms or relevant patient tissues. Over the past few years, there has been an increasing body of work reporting on disease-related phenotypes in patient-specific iPSCs (Bellin et al., 2012). In a much smaller set of studies iPSC-based disease modeling efforts have moved beyond the identification of disease phenotypes toward novel mechanistic insights. Examples from our team include the modeling of tissue-specific splicing defects in familial dysautonomia (FD) (Lee et al., 2009) or the cell-type-specific defects in innate immunity in the case of primary herpes simplex encephalitis (HSE) (Lafaille et al., 2012). There has also been success in using iPSC-based models for drug discovery (Lee et al., 2012), including studies wherein hit compounds have now moved forward into early-stage clinical trials, such as kinetin in the case of FD, the anticonvulsant compound retigabine for the treatment of amyotrophic lateral sclerosis (ALS) (Wainger et al., 2014), or IGF-1 in the treatment of autism-related disorders (Marchetto et al., 2010),(Shcheglovitov et al., 2013).

Many labs are pursuing iPSC-based disease modeling, and thousands of patient-specific iPSC lines are currently under production across the globe. In contrast, there is a much more limited effort to address remaining limitations of the technology. One such challenge is the question of how to introduce "age" into iPSC-based models of late-onset disorders. It is well known that age is the largest risk factor for many human disorders ranging from AD and PD to cancer, trumping the combined effect of both environment and genetics (Niccoli and Partridge, 2012). Most current iPSC studies model late-onset disease phenotypes in iPSC-derived neurons that represent first-trimester-stage cells. Despite such an age mismatch, some studies have successfully modeled aspects of human disease including AD and PD (reviewed in Liu et al. [2012a] and Srikanth and Young-Pearse [2014]). One interpretation is that such studies model susceptibility rather than normal disease or disease progression. Modeling disease susceptibility can be very powerful for studying biochemical changes directly dependent on a genetic defect. For example, changes in amyloid precursor protein processing, a-synuclein accumulation, or mitochondrial function have been reported in models of AD or PD.

Induced Aging Strategies

The most common strategy to induce age-like features in iPSC-derived lineages is the use of stress paradigms. Various studies have modeled late onset disorders in iPSC-derived cells following exposure to toxins, including compounds that trigger mitochondrial stress or ROS (Byers et al., 2011; Cooper et al., 2012; Liu et al., 2012b; Nguyen et al., 2011; Reinhardt et al., 2013; Seibler et al., 2011). Such in vitro stress paradigms may uncover disease-related susceptibility phenotypes (Figure 3B). However, it remains to be determined whether stressing "young" cells appropriately mimics age-related disease susceptibility. For example, it may be equally important to understand the nature of the protective mechanisms that stave off disease rather than to artificially force a disease phenotype in "young" cells. Another key concern is that most disease modeling studies have not assessed the impact of stress paradigms on the

CellPress

hallmarks of cellular aging (see Figure 1C). We postulate that an ideal in vitro aging strategy involves the induction of age-related hallmarks in both healthy and patient-specific cells without exerting overt toxicity. In vitro aging paradigms should be a tool for measuring the contribution of age-related changes and studying the interaction of genetic (isogenic lines) and age-related susceptibility on disease phenotypes.

Alternative strategies to manipulate cellular age in iPSC-derived lineages include insights from premature aging disorders such as Hutchinson Gildford progeria syndrome (HGPS), dyskeratosis congenita, and Werner and Cockayne syndrome, among others (Kipling et al., 2004). Interestingly, iPSC models of those progeroid syndromes show in large a reset of age-related phenotypes upon reprogramming followed by a rapid reacquisition of age-like features upon differentiation (Agarwal et al., 2010; Andrade et al., 2012; Batista et al., 2011; Liu et al., 2011; Zhang et al., 2011). These findings indicate that the induction of cellular age markers should be achievable by ectopically expressing progeroid gene products in wild-type iPSC-derived cells. In agreement with this, a recent report describes how de novo expression of mutant Werner syndrome gene in human-ESC-derived cells drives the appearance of cellular aging hallmarks, such as alterations in heterochromatin organization (Zhang et al., 2015). We previously reported ectopic expression of progerin, the mutant gene product causing HGPS, as a strategy to induce age-related hallmarks in iPSC-derived fibroblasts and neurons (Miller et al., 2013) (Figure 3B). We used the strategy to identify age- and disease-associated phenotypes in PD-iPSC-derived dopamine neurons. Among the most dramatic age-related phenotypes was the accumulation of neuromelanin, which is an age-related marker of midbrain dopamine neurons. Other findings included the presence of pathological intracellular structures by electron microscopy and the progressive loss of tyrosine-hydroxylase-positive cells after transplantation, phenotypes that required the interaction of genetic (Parkin mutation) and age-related susceptibility (progerin exposure). A key feature of the approach is the ability to ectopically drive progerin expression in iPSC-derived neurons of any genetic background, as endogenous progerin levels in HGPS-iPSC-derived neurons are likely too low for triggering age-related neuronal phenotypes. In fact, neurons appear largely spared in HGPS patients. This raises the question of how closely progerin exposure mimics normal aging. We demonstrate that progerin expression can induce phenotypic hallmarks of aging in iPSC-derived lineages, including neurons. However, beyond those phenotypic studies, we have not yet analyzed in-depth the molecular changes associated with progerin exposure. Therefore, it is conceivable that progerin may trigger age-related markers via a distinct molecular mechanism in comparison to normal aging.

Future strategies to induce aging may move beyond mimicking human progeroid disorders toward applying the increasing mechanistic understanding of human aging. The direct manipulation of any of the key cellular and molecular hallmarks of aging may be a particularly promising approach. Efficient tools for genome editing, such as the CRISPR/Cas9 system, will greatly facilitate such efforts and enable the systematic targeting of candidate pathways alone or in combination. The question whether manipulation of a single pathway can trigger broader changes in age-related hallmarks will be particularly important. Furthermore, it will be interesting to assess whether any given induced aging strategy may be more appropriate for a specific cell or tissue type (e.g., proliferating versus postmitotic tissue). Such efforts could go hand in hand with studies in model organisms. For example, recent efforts to systematically manipulate hallmarks of aging in the short-lived African killifish (Harel et al., 2015) using CRISPR/Cas9 technology could be highly complementary in determining the impact of aging strategies on individual organs.

Conclusions

We propose PSCs cells as a novel model for studying human aging. Unlike traditional aging paradigms that focus on endpoints such as longevity or the restoration of regenerative capacity, PSCs allow us to monitor and manipulate molecular and cellular hallmarks of aging during both reprogramming and cell differentiation. Capturing the timing and sequence of the steps involved in cellular rejuvenation offers a unique opportunity for subsequent mechanistic studies. The strong evidence for cellular rejuvenation during iPSC induction indicates that many aspects of aging are reversible and may represent epigenetic rather than genetic barriers in biology. Therefore, a future is conceivable wherein it will be possible to reliably rejuvenate somatic cells without the need to move them back to pluripotency. In addition to studying rejuvenation, it will be equally important to identify novel induced aging strategies. The ability to direct both cell fate and age in iPSC-derived lineages will allow modeling of human disorders at unprecedented precision. Such studies could yield more relevant disease phenotypes and define novel classes of therapeutic compounds targeting age-related cell behaviors. The ability to program and reprogram cellular age on demand will present an important step forward on the road to decoding the mystery of aging.

ACKNOWLEDGEMENTS

E.V. was supported by a NYSTEM postdoctoral fellowship. The work of the authors described in this review was supported by grants from the Starr Foundation, NYSTEM (C028503 and C026447), NINDS/NIH (NS072381), and NCI/NIH (P30CA008748).

REFERENCES

Agarwal, S., Loh, Y.H., McLoughlin, E.M., Huang, J., Park, I.H., Miller, J.D., Huo, H., Okuka, M., Dos Reis, R.M., Loewer, S., et al. (2010). Telomere elongation in induced pluripotent stem cells from dyskeratosis congenita patients. Nature *464*, 292–296.

Andrade, L.N., Nathanson, J.L., Yeo, G.W., Menck, C.F., and Muotri, A.R. (2012). Evidence for premature aging due to oxidative stress in iPSCs from Cockayne syndrome. Hum. Mol. Genet. *21*, 3825–3834.

Arantes-Oliveira, N., Berman, J.R., and Kenyon, C. (2003). Healthy animals with extreme longevity. Science *302*, 611.

Barberi, T., Klivenyi, P., Calingasan, N.Y., Lee, H., Kawamata, H., Loonam, K., Perrier, A.L., Bruses, J., Rubio, M.E., Topf, N., et al. (2003). Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in parkinsonian mice. Nat. Biotechnol. *21*, 1200–1207.

Bardy, C., van den Hurk, M., Eames, T., Marchand, C., Hernandez, R.V., Kellogg, M., Gorris, M., Galet, B., Palomares, V., Brown, J., et al. (2015). Neuronal medium that supports basic synaptic functions and activity of human neurons in vitro. Proc. Natl. Acad. Sci. USA. Published online April 13, 2015.

Batista, L.F., Pech, M.F., Zhong, F.L., Nguyen, H.N., Xie, K.T., Zaug, A.J., Crary, S.M., Choi, J., Sebastiano, V., Cherry, A., et al. (2011). Telomere

shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. Nature *474*, 399–402.

Bellin, M., Marchetto, M.C., Gage, F.H., and Mummery, C.L. (2012). Induced pluripotent stem cells: the new patient? Nat. Rev. Mol. Cell Biol. *13*, 713–726.

Bert, P.J. (1864). J Anatomie Physiologie *1*, 69–87.

Blasco, M.A., Lee, H.W., Hande, M.P., Samper, E., Lansdorp, P.M., DePinho, R.A., and Greider, C.W. (1997). Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. Cell *91*, 25–34.

Brack, A.S., Conboy, M.J., Roy, S., Lee, M., Kuo, C.J., Keller, C., and Rando, T.A. (2007). Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. Science *317*, 807–810.

Byers, B., Cord, B., Nguyen, H.N., Schüle, B., Fenno, L., Lee, P.C., Deisseroth, K., Langston, J.W., Pera, R.R., and Palmer, T.D. (2011). SNCA triplication Parkinson's patient's iPSC-derived DA neurons accumulate α-synuclein and are susceptible to oxidative stress. PLoS ONE *6*, e26159.

Cabreiro, F., Au, C., Leung, K.Y., Vergara-Irigaray, N., Cochemé, H.M., Noori, T., Weinkove, D., Schuster, E., Greene, N.D., and Gems, D. (2013). Metformin retards aging in C. elegans by altering microbial folate and methionine metabolism. Cell *153*, 228–239.

Campbell, K.H., McWhir, J., Ritchie, W.A., and Wilmut, I. (1996). Sheep cloned by nuclear transfer from a cultured cell line. Nature *380*, 64–66.

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.

Chambers, S.M., Qi, Y., Mica, Y., Lee, G., Zhang, X.J., Niu, L., Bilsland, 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.

Cheng, C.W., Adams, G.B., Perin, L., Wei, M., Zhou, X., Lam, B.S., Da Sacco, S., Mirisola, M., Quinn, D.I., Dorff, T.B., et al. (2014). Prolonged fasting reduces IGF-1/PKA to promote hematopoietic-stem-cell-based regeneration and reverse immunosuppression. Cell Stem Cell *14*, 810–823.

Colman, R.J., Anderson, R.M., Johnson, S.C., Kastman, E.K., Kosmatka, K.J., Beasley, T.M., Allison, D.B., Cruzen, C., Simmons, H.A., Kemnitz, J.W., and Weindruch, R. (2009). Caloric restriction delays disease onset and mortality in rhesus monkeys. Science *325*, 201–204.

Conboy, I.M., Conboy, M.J., Wagers, A.J., Girma, E.R., Weissman, I.L., and Rando, T.A. (2005). Rejuvenation of aged progenitor cells by exposure to a young systemic environment. Nature *433*, 760–764.

Conboy, M.J., Conboy, I.M., and Rando, T.A. (2013). Heterochronic parabiosis: historical perspective and methodological considerations for studies of aging and longevity. Aging Cell *12*, 525–530.

Cooper, O., Seo, H., Andrabi, S., Guardia-Laguarta, C., Graziotto, J., Sundberg, M., McLean, J.R., Carrillo-Reid, L., Xie, Z., Osborn, T., et al. (2012). Pharmacological rescue of mitochondrial deficits in iPSC-derived neural cells from patients with familial Parkinson's disease. Sci. Transl. Med. *4*, 41ra90.

Espuny-Camacho, I., Michelsen, K.A., Gall, D., Linaro, D., Hasche, A., Bonnefont, J., Bali, C., Orduz, D., Bilheu, A., Herpoel, A., et al. (2013). Pyramidal neurons derived from human pluripotent stem cells integrate efficiently into mouse brain circuits in vivo. Neuron *77*, 440–456.

Evason, K., Huang, C., Yamben, I., Covey, D.F., and Kornfeld, K. (2005). Anticonvulsant medications extend worm life-span. Science *307*, 258–262.

Guarente, L. (2013). Calorie restriction and sirtuins revisited. Genes Dev. *27*, 2072–2085.

Guarente, L., and Kenyon, C. (2000). Genetic pathways that regulate ageing in model organisms. Nature *408*, 255–262.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. *10*, 622–640.

Harel, I., Benayoun, B.A., Machado, B., Singh, P.P., Hu, C.K., Pech, M.F., Valenzano, D.R., Zhang, E., Sharp, S.C., Artandi, S.E., and Brunet, A. (2015). A platform for rapid exploration of aging and diseases in a naturally short-lived vertebrate. Cell *160*, 1013–1026.

Hayflick, L. (2007). Biological aging is no longer an unsolved problem. Ann. N Y Acad. Sci. *1100*, 1–13.

Herranz, D., and Serrano, M. (2010). SIRT1: recent lessons from mouse models. Nat. Rev. Cancer *10*, 819–823.

Isacson, O., and Deacon, T. (1997). Neural transplantation studies reveal the brain's capacity for continuous reconstruction. Trends Neurosci. *20*, 477–482.

Jin, K. (2010). Modern Biological Theories of Aging. Aging Dis. *1*, 72–74.

Kadoshima, T., Sakaguchi, H., Nakano, T., Soen, M., Ando, S., Eiraku, M., and Sasai, Y. (2013). Self-organization of axial polarity, inside-out layer pattern, and species-specific progenitor dynamics in human ES cell-derived neocortex. Proc. Natl. Acad. Sci. USA *110*, 20284–20289.

Kanfi, Y., Naiman, S., Amir, G., Peshti, V., Zinman, G., Nahum, L., Bar-Joseph, Z., and Cohen, H.Y. (2012). The sirtuin SIRT6 regulates lifespan in male mice. Nature *483*, 218–221.

Kenyon, C.J. (2010). The genetics of ageing. Nature *464*, 504–512.

Kim, J.H., Auerbach, J.M., Rodríguez-Gómez, J.A., Velasco, I., Gavin, D., Lumelsky, N., Lee, S.H., Nguyen, J., Sánchez-Pernaute, R., Bankiewicz, K., and McKay, R. (2002). Dopamine neurons derived from embryonic stem cells function in an animal model of Parkinson's disease. Nature *418*, 50–56.

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, E.B., Fang, X., Fushan, A.A., Huang, Z., Lobanov, A.V., Han, L., Marino, S.M., Sun, X., Turanov, A.A., Yang, P., et al. (2011a). Genome sequencing reveals insights into physiology and longevity of the naked mole rat. Nature *479*, 223–227.

Kim, K., Zhao, R., Doi, A., Ng, K., Unternaehrer, J., Cahan, P., Huo, H., Loh, Y.H., Aryee, M.J., Lensch, M.W., et al. (2011b). Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. Nat. Biotechnol. *29*, 1117–1119.

Kipling, D., Davis, T., Ostler, E.L., and Faragher, R.G. (2004). What can progeroid syndromes tell us about human aging? Science *305*, 1426–1431.

Kirkeby, A., Grealish, S., Wolf, D.A., Nelander, J., Wood, J., Lundblad, M., Lindvall, O., and Parmar, M. (2012). Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions. Cell Rep. *1*, 703–714.

Krencik, R., Weick, J.P., Liu, Y., Zhang, Z.J., and Zhang, S.C. (2011). Specification of transplantable astroglial subtypes from human pluripotent stem cells. Nat. Biotechnol. *29*, 528–534.

Kriks, S., Shim, J.W., Piao, J., Ganat, Y.M., Wakeman, D.R., Xie, Z., Carrillo-Reid, L., Auyeung, G., Antonacci, C., Buch, A., et al. (2011). Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature *480*, 547–551.

Lafaille, F.G., Pessach, I.M., Zhang, S.Y., Ciancanelli, M.J., Herman, M., Abhyankar, A., Ying, S.W., Keros, S., Goldstein, P.A., Mostoslavsky, G., et al. (2012). Impaired intrinsic immunity to HSV-1 in human iPSC-derived TLR3-deficient CNS cells. Nature *491*, 769–773.

Lancaster, M.A., Renner, M., Martin, C.A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature *501*, 373–379.

Lanza, R.P., Cibelli, J.B., Blackwell, C., Cristofalo, V.J., Francis, M.K., Baerlocher, G.M., Mak, J., Schertzer, M., Chavez, E.A., Sawyer, N., et al. (2000). Extension of cell life-span and telomere length in animals cloned from senescent somatic cells. Science *288*, 665–669.

Lapasset, L., Milhavet, O., Prieur, A., Besnard, E., Babled, A., Aït-Hamou, N., Leschik, J., Pellestor, F., Ramirez, J.M., De Vos, J., et al. (2011). Rejuvenating senescent and centenarian human cells by reprogramming through the pluripotent state. Genes Dev. *25*, 2248–2253.

Lee, H.W., Blasco, M.A., Gottlieb, G.J., Horner, J.W., 2nd, Greider, C.W., and DePinho, R.A. (1998). Essential role of mouse telomerase in highly proliferative organs. Nature *392*, 569–574.

Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., Ganat, Y.M., Menon, J., Shimizu, F., Viale, A., et al. (2009). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. Nature *461*, 402–406.

Lee, G., Ramirez, C.N., Kim, H., Zeltner, N., Liu, B., Radu, C., Bhinder, B., Kim, Y.J., Choi, I.Y., Mukherjee-Clavin, B., et al. (2012). Large-scale screening using familial dysautonomia induced pluripotent stem cells identifies compounds that rescue IKBKAP expression. Nat. Biotechnol. *30*, 1244–1248.

Liu, G.H., Barkho, B.Z., Ruiz, S., Diep, D., Qu, J., Yang, S.L., Panopoulos, A.D., Suzuki, K., Kurian, L., Walsh, C., et al. (2011). Recapitulation of premature ageing with iPSCs from Hutchinson-Gilford progeria syndrome. Nature *472*, 221–225.

Liu, G.H., Ding, Z., and Izpisua Belmonte, J.C. (2012a). iPSC technology to study human aging and aging-related disorders. Curr. Opin. Cell Biol. *24*, 765–774.

Liu, G.H., Qu, J., Suzuki, K., Nivet, E., Li, M., Montserrat, N., Yi, F., Xu, X., Ruiz, S., Zhang, W., et al. (2012b). Progressive degeneration of human neural stem cells caused by pathogenic LRRK2. Nature *491*, 603–607.

Loffredo, F.S., Steinhauser, M.L., Jay, S.M., Gannon, J., Pancoast, J.R., Yalamanchi, P., Sinha, M., Dall'Osso, C., Khong, D., Shadrach, J.L., et al. (2013). Growth differentiation factor 11 is a circulating factor that reverses age-related cardiac hypertrophy. Cell *153*, 828–839.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The hallmarks of aging. Cell *153*, 1194–1217.

Mann, D.M., and Yates, P.O. (1974). Lipoprotein pigments—their relationship to ageing in the human nervous system. II. The melanin content of pigmented nerve cells. Brain *97*, 489–498.

Marchetto, M.C., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H., and Muotri, A.R. (2010). A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell *143*, 527–539.

Mariani, J., Simonini, M.V., Palejev, D., Tomasini, L., Coppola, G., Szekely, A.M., Horvath, T.L., and Vaccarino, F.M. (2012). Modeling human cortical development in vitro using induced pluripotent stem cells. Proc. Natl. Acad. Sci. USA *109*, 12770–12775.

Marion, R.M., Strati, K., Li, H., Tejera, A., Schoeftner, S., Ortega, S., Serrano, M., and Blasco, M.A. (2009). Telomeres acquire embryonic stem cell characteristics in induced pluripotent stem cells. Cell Stem Cell *4*, 141–154.

Maroof, A.M., Keros, S., Tyson, J.A., Ying, S.W., Ganat, Y.M., Merkle, F.T., Liu, B., Goulburn, A., Stanley, E.G., Elefanty, A.G., et al. (2013). Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. Cell Stem Cell *12*, 559–572.

Martin-Montalvo, A., Mercken, E.M., Mitchell, S.J., Palacios, H.H., Mote, P.L., Scheibye-Knudsen, M., Gomes, A.P., Ward, T.M., Minor, R.K., Blouin, M.J., et al. (2013). Metformin improves healthspan and lifespan in mice. Nat. Commun. *4*, 2192.

Mattison, J.A., Roth, G.S., Beasley, T.M., Tilmont, E.M., Handy, A.M., Herbert, R.L., Longo, D.L., Allison, D.B., Young, J.E., Bryant, M., et al. (2012). Impact of caloric restriction on health and survival in rhesus monkeys from the NIA study. Nature *489*, 318–321.

Miller, J.D., Ganat, Y.M., Kishinevsky, S., Bowman, R.L., Liu, B., Tu, E.Y., Mandal, P.K., Vera, E., Shim, J.W., Kriks, S., et al. (2013). Human iPSC-based modeling of late-onset disease via progerin-induced aging. Cell Stem Cell *13*, 691–705.

Murga, M., Bunting, S., Montaña, M.F., Soria, R., Mulero, F., Cañamero, M., Lee, Y., McKinnon, P.J., Nussenzweig, A., and Fernandez-Capetillo, O. (2009). A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. Nat. Genet. *41*, 891–898.

Nguyen, H.N., Byers, B., Cord, B., Shcheglovitov, A., Byrne, J., Gujar, P., Kee, K., Schüle, B., Dolmetsch, R.E., Langston, W., et al. (2011). LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. Cell Stem Cell *8*, 267–280.

Niccoli, T., and Partridge, L. (2012). Ageing as a risk factor for disease. Curr. Biol. *22*, R741–R752.

Nicholas, C.R., Chen, J., Tang, Y., Southwell, D.G., Chalmers, N., Vogt, D., Arnold, C.M., Chen, Y.J., Stanley, E.G., Elefanty, A.G., et al. (2013). Functional maturation of hPSC-derived forebrain interneurons requires an extended timeline and mimics human neural development. Cell Stem Cell *12*, 573–586.

Nishimura, T., Kaneko, S., Kawana-Tachikawa, A., Tajima, Y., Goto, H., Zhu, D., Nakayama-Hosoya, K., Iriguchi, S., Uemura, Y., Shimizu, T., et al. (2013). Generation of rejuvenated antigen-specific T cells by reprogramming to pluripotency and redifferentiation. Cell Stem Cell *12*, 114–126.

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.

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.

Prigione, A., Hossini, A.M., Lichtner, B., Serin, A., Fauler, B., Megges, M., Lurz, R., Lehrach, H., Makrantonaki, E., Zouboulis, C.C., and Adjaye, J. (2011). Mitochondrial-associated cell death mechanisms are reset to an embryonic-like state in aged donor-derived iPS cells harboring chromosomal aberrations. PLoS ONE *6*, e27352.

Reinhardt, P., Schmid, B., Burbulla, L.F., Schöndorf, D.C., Wagner, L., Glatza, M., Höing, S., Hargus, G., Heck, S.A., Dhingra, A., et al. (2013). Genetic correction of a LRRK2 mutation in human iPSCs links parkinsonian neurodegeneration to ERK-dependent changes in gene expression. Cell Stem Cell *12*, 354–367.

Seibler, P., Graziotto, J., Jeong, H., Simunovic, F., Klein, C., and Krainc, D. (2011). Mitochondrial Parkin recruitment is impaired in neurons derived from mutant PINK1 induced pluripotent stem cells. J. Neurosci. *31*, 5970–5976.

Shamblott, M.J., Axelman, J., Wang, S., Bugg, E.M., Littlefield, J.W., Donovan, P.J., Blumenthal, P.D., Huggins, G.R., and Gearhart, J.D. (1998). Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc. Natl. Acad. Sci. USA *95*, 13726–13731.

Shcheglovitov, A., Shcheglovitova, O., Yazawa, M., Portmann, T., Shu, R., Sebastiano, V., Krawisz, A., Froehlich, W., Bernstein, J.A., Hallmayer, J.F., and Dolmetsch, R.E. (2013). SHANK3 and IGF1 restore synaptic deficits in neurons from 22q13 deletion syndrome patients. Nature *503*, 267–271.

Shi, Y., Kirwan, P., Smith, J., Robinson, H.P., and Livesey, F.J. (2012). Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. Nat. Neurosci. *15*, 477–486, S1.

Sinha, M., Jang, Y.C., Oh, J., Khong, D., Wu, E.Y., Manohar, R., Miller, C., Regalado, S.G., Loffredo, F.S., Pancoast, J.R., et al. (2014). Restoring systemic GDF11 levels reverses age-related dysfunction in mouse skeletal muscle. Science *344*, 649–652.

Srikanth, P., and Young-Pearse, T.L. (2014). Stem cells on the brain: modeling neurodevelopmental and neurodegenerative diseases using human induced pluripotent stem cells. J. Neurogenet. *28*, 5–29.

Steinbeck, J.A., and Studer, L. (2015). Moving Stem Cells to the Clinic: Potential and Limitations for Brain Repair. Neuron *86*, 187–206.

Suhr, S.T., Chang, E.A., Rodriguez, R.M., Wang, K., Ross, P.J., Beyhan, Z., Murthy, S., and Cibelli, J.B. (2009). Telomere dynamics in human cells reprogrammed to pluripotency. PLoS ONE *4*, e8124.

Suhr, S.T., Chang, E.A., Tjong, J., Alcasid, N., Perkins, G.A., Goissis, M.D., Ellisman, M.H., Perez, G.I., and Cibelli, J.B. (2010). Mitochondrial rejuvenation after induced pluripotency. PLoS ONE *5*, e14095.

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.

Themeli, M., Kloss, C.C., Ciriello, G., Fedorov, V.D., Perna, F., Gonen, M., and Sadelain, M. (2013). Generation of tumor-targeted human T lymphocytes from induced pluripotent stem cells for cancer therapy. Nat. Biotechnol. *31*, 928–933.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Tomás-Loba, A., Flores, I., Fernández-Marcos, P.J., Cayuela, M.L., Maraver, A., Tejera, A., Borrás, C., Matheu, A., Klatt, P., Flores, J.M., et al. (2008). Telomerase reverse transcriptase delays aging in cancer-resistant mice. Cell *135*, 609–622.

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.

Villeda, S.A., Luo, J., Mosher, K.I., Zou, B., Britschgi, M., Bieri, G., Stan, T.M., Fainberg, N., Ding, Z., Eggel, A., et al. (2011). The ageing systemic milieu negatively regulates neurogenesis and cognitive function. Nature *477*, 90–94.

Vizcardo, R., Masuda, K., Yamada, D., Ikawa, T., Shimizu, K., Fujii, S., Koseki, H., and Kawamoto, H. (2013). Regeneration of human tumor antigen-specific T cells from iPSCs derived from mature CD8(+) T cells. Cell Stem Cell *12*, 31–36.

Wainger, B.J., Kiskinis, E., Mellin, C., Wiskow, O., Han, S.S., Sandoe, J., Perez, N.P., Williams, L.A., Lee, S., Boulting, G., et al. (2014). Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. Cell Rep. *7*, 1–11.

Wang, S., Bates, J., Li, X., Schanz, S., Chandler-Militello, D., Levine, C., Maherali, N., Studer, L., Hochedlinger, K., Windrem, M., and Goldman, S.A. (2013). Human iPSC-derived oligodendrocyte progenitor cells can myelinate and rescue a mouse model of congenital hypomyelination. Cell Stem Cell *12*, 252–264.

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.

Zhang, J., Lian, Q., Zhu, G., Zhou, F., Sui, L., Tan, C., Mutalif, R.A., Navasankari, R., Zhang, Y., Tse, H.F., et al. (2011). A human iPSC model of Hutchinson Gilford Progeria reveals vascular smooth muscle and mesenchymal stem cell defects. Cell Stem Cell *8*, 31–45.

Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron *78*, 785–798.

Zhang, W., Li, J., Suzuki, K., Qu, J., Wang, P., Zhou, J., Liu, X., Ren, R., Xu, X., Ocampo, A., et al. (2015). A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. Science

Cell Stem Cell
A Decade of iPS Cells
Review

CellPress

Complex Tissue and Disease Modeling using hiPSCs

Robert Passier,[1,2] Valeria Orlova,[1] and Christine Mummery[1,*]
[1]Department of Anatomy and Embryology, Leiden University Medical Centre, Einthovenweg 20, 2333ZC Leiden, The Netherlands
[2]Department of Applied Stem Cell Technologies, MIRA Institute, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands
*Correspondence: c.l.mummery@lumc.nl
http://dx.doi.org/10.1016/j.stem.2016.02.011

Defined genetic models based on human pluripotent stem cells have opened new avenues for understanding disease mechanisms and drug screening. Many of these models assume cell-autonomous mechanisms of disease but it is possible that disease phenotypes or drug responses will only be evident if all cellular and extracellular components of a tissue are present and functionally mature. To derive optimal benefit from such models, complex multicellular structures with vascular components that mimic tissue niches will thus likely be necessary. Here we consider emerging research creating human tissue mimics and provide some recommendations for moving the field forward.

Introduction

It is increasingly clear that animal models fall short in predicting the pathophysiology of many human diseases. Aside from differences in physiology, the immune system, inflammation, and individual genetic backgrounds, there are important differences in liver metabolism compared with other species, even when compared with other primates. These differences impact severity of the disease phenotype and the effectiveness of new drugs in clinical trials. Some of these issues also contribute to the failure to identify potentially toxic side effects of new drugs in current safety pharmacology studies. Drug attrition rates are high not least because present preclinical assays do not always detect potential risk of damage to the heart, kidney, liver, and brain. The emergence of reprogramming as an approach to derive human induced pluripotent stem cells (hiPSCs) from patients and healthy individuals combined with efficient gene modification has led to unprecedented opportunities over the last several years to model human disease. Ten years after their first discovery, many researchers now produce hiPSC lines routinely and induce their differentiation efficiently into multiple somatic cell types that are affected by inherited diseases or are specific drug targets. Commercial providers have optimized reagents and differentiation protocols such that they are now widely applicable across many hiPSC lines. While regenerative medicine is still a long-term goal, other research is looking toward more immediate uses of hiPSCs in understanding mechanisms underlying disease and finding ways to delay or reverse its natural course. Repurposing of existing drugs through better understanding of their mechanisms of action on patient-derived hiPSCs has already resulted in some going directly into clinical trials without intervening animal experiments for severe conditions with no other treatment options. These include amyotrophic lateral sclerosis (ALS), spinal muscular atrophy (SMA), and Alzheimer's disease (AD) (Bright et al., 2015; McNeish et al., 2015; Naryshkin et al., 2014; Wainger et al., 2014). This is providing new perspectives for treating intractable conditions at an extraordinary rate.

However, many diseases have multicellular contributions and are not cell autonomous as often assumed. The next generation of disease models is therefore increasingly based on combinations of cell types, sometimes in "organ-on-chip" formats, microfluidic devices that integrate multiple cell types of various developmental lineages as complex synthetic human tissues in chips, or as "organoids," structures of one or more cell types that self-organize into organ subunits. These models can provide rapid readouts of disease pathology and allow the identification of compounds or drugs that could reverse the condition in vivo. They also support integration of various forms of "tissue stress:" inflammatory cytokines and cells, bacterial or viral challenge, biophysical stretch and strain, and microfluidic flow through synthetic vessels, mimicking interaction with the circulatory system.

In this Review we consider what has been learned from various models in which multicellular interaction has been shown to impact phenotype (Figure 1). The emerging complex models that capture the 3D tissue niche and promote cell maturation may represent the future in drug discovery and safety pharmacology models.

Disease Modeling and Safety Pharmacology in Monotypic hiPSC Cultures

hiPSC lines were derived from patients (Park et al., 2008) within 1 year of the first description of human somatic cell reprogramming (Takahashi et al., 2007; Yu et al., 2007). Since then, many hiPSC lines have been described and the cell type expressing the mutant gene has often been shown to exhibit the phenotype expected from the patient in simple monotypic cultures. These have been reviewed both for monogenic (Bellin et al., 2012; Merkle and Eggan, 2013) and complex (Glass et al., 2010; Zhu et al., 2011) genetic diseases so are only described briefly here. The major advantage of studying monotypic hiPSC cultures is that they usually have clear, defined readouts reflecting the pathophysiology of the target cells and the model can be tailored and scaled up as necessary for high-throughput screening. Of note, though, because the first patient hiPSC derivatives were studied before differentiation protocols had become efficient and surface markers available to select differentiated cell types, many studies were carried out in mixed but undefined and variable cell populations, which may have confounded precise phenotypic analysis. However, it is now often possible to create defined cell type combinations that can be controlled from experiment to experiment and are physiologically meaningful. In addition, combinations of cells from different developmental

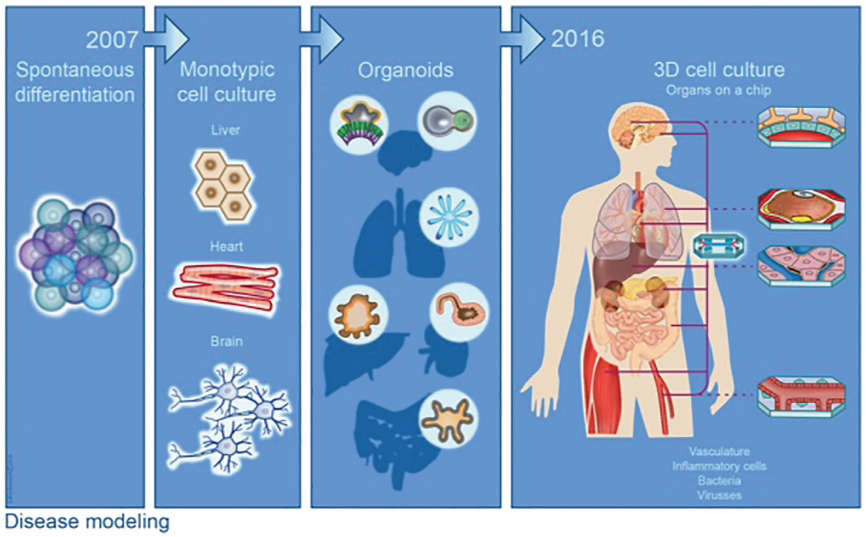


Figure 1. Timeline of Developments in the hiPSC Field

Schematic overview illustrating progress from spontaneous and uncontrolled differentiation of mixed cell types via "embryoid bodies" toward controlled differentiation of specialized cells types in defined conditions that can be used for disease modeling and, finally, the development of complex structures and systems that recapitulate human tissues on a small scale, such as organoids and organ-on-chip technologies for disease modeling applications. The illustration is courtesy of Bas Blankevoort.

lineages can be combined as in normal organs and tissues. We consider here the hiPSC models available for multiple organs but focus on heart and brain since it is now possible to generate different cell subtypes relevant to disease from these organs very specifically, phenotypic readouts are available, and states of cell maturity can be achieved that capture salient features of the diseases in humans.

Heart

Contractile cardiomyocytes represent approximately one-third of the total numbers of myocardial cells but constitute two-thirds of total myocardial volume and are responsible for providing permanent blood flow via coordinated electrical and contractile activity. Disturbances in these properties lead to impaired heart function and are major causes of cardiac disease. Since mutations in genes encoding ion channels (channelopathies) as well as cytoskeletal and sarcomeric proteins (cardiomyopathies) affect cardiomyocyte function in a cell-autonomous manner, these disease phenotypes can be studied both in single cardiomyocytes and in multicellular monotypic cultures.

In the last decade, major advances in standardizing and improving differentiation conditions mean that many hiPSC lines yield cultures containing 20%–80% cardiomyocytes depending on the methodology (Burridge et al., 2012; Mummery et al., 2012). Genetic marking of cardiac transcription factors (Den Hartogh and Passier, 2016) or cardiomyocyte surface markers, such as SIRP1 and VCAM1 (Elliott et al., 2011), enable further enrichment for functional analysis on purified cardiomyocyte populations (Bellin et al., 2013). Conventional single-cell patch-clamp electrophysiology has been widely used to study the effects of ion channel mutations in hiPSC-derived cardiomyocytes (hiPSC-CMs). These ion channels are expressed early during fetal development and in differentiating hiPSC-CMs; phenotypes have therefore been readily detectable despite cardiomyocyte immaturity. Various channelopathies, including long-QT syndrome (LQTS), Brugada syndrome, Timothy syndrome (also called LQT8), and catecholaminergic polymorphic ventricular tachycardia (CPVT) have been modeled in hiPSC-CMs; all faithfully recapitulate cardiac phenotypes observed in patients (reviewed in Karakikes et al., 2015). Cardiomyopathies are also severe heart diseases often caused by mutations in the structural sarcomeric proteins. They are characterized by systolic and/or diastolic dysfunction, sarcomere disarray, hypertrophy (or cardiomyocyte enlargement), and interstitial fibrosis. In single hiPSC-CMs from patients with dilated cardiomyopathy (Sun et al., 2012; Wu et al., 2015), hypertrophic cardiomyopathy (Han et al., 2014; Lan et al., 2013), LEOPARD syndrome, and arrhythmogenic right ventricular cardiomyopathy (Kim et al., 2013; Ma et al., 2013), the disease phenotype manifested as increased cardiomyocyte size in culture, with sarcomeric disorganization and contractile dysfunction.

Drugs can affect cardiomyocyte function in much the same way as inherited genetic disorders and induce cardiotoxicity. The majority of cardiotoxic drugs bind to the human Ether-à-go-go (hERG) potassium ion channel, responsible for the repolarizing I_{Kr} currents during the cardiac action potential. The hERG channel block prolongs the action potential (long-QT) and, just as genetic forms of long-QT, can cause life-threatening cardiac arrhythmias. How to predict which individuals might be sensitive to drug-induced long-QT is an important question in drug development. hiPSC-CMs are already proving useful in predicting drug-induced cardiotoxicity but just as importantly, they capture genetic variance and may help identify susceptible individuals for personalized medicine. This was exemplified by the observation that hiPSC-CMs from LQT patients with a mutation in *hERG* showed predisposed sensitivity to inhibitors of the I_{ks} channel (Braam et al., 2013).

Despite these successes of hiPSC-CMs in recapitulating specific cardiac diseases and predicting drug-induced cardiotoxicity, modeling complex multicellular cardiac disease phenotypes requires populations of all cardiac cell types. Besides ventricular, atrial, and pacemaker cardiomyocytes, vascular and epicardial cells are also needed to create complex human heart tissues and mimic their functions. Recent culture condition refinements based on recapitulating the sequence of signals that occur during heart development have enabled all of the major cell types of the heart to be derived from hiPSCs (Birket et al., 2015b; Devalla et al., 2015; Ionta et al., 2015; Iyer et al., 2015; Jung et al., 2014; Witty et al., 2014; Zhang et al., 2011). The stage is now set to start recombining these cells in 2D surfaces pre-patterned to force cardiomyocyte alignment or in 3D spheroids in organoid-like formats to see whether they will undergo maturation and morphogenic organization as in the heart itself.

In 2D, cardiomyocytes plated on polymers, such as polydimethylsiloxane (PDMS) or polyacrylamide, on which rectangles of different aspect ratios (ranging from 1:1 to 7:1) had been patterned, became "anisotropic;" this means they became aligned as in the heart and showed enhanced sarcomere organization (Bray et al., 2008; Ribeiro et al., 2015a). These flexible transparent polymers are relatively soft compared to culture plastic and are more like native heart tissue. They have been used to determine cardiomyocyte contractile force based on displacement of the short edges of the cells or of fluorescent beads on the polymer surface during contraction cycles by video imaging (Ribeiro et al., 2015b). Increased contractile force was recently demonstrated in this way in hiPSC-CMs following the addition of thyroid hormone (T3), insulin-like growth factor (IGF)1, and the corticosteroid dexamethasone. These factors have been implicated during late fetal/perinatal tissue maturation but are produced systemically and not by cardiomyocytes themselves. The increased contraction force was accompanied by increased upstroke velocity of the action potential and reduced resting membrane potential, indicating enhanced maturation under these conditions (Birket et al., 2015a). Of importance, these conditions were crucial for revealing reduced contraction force in hiPSC-CMs derived from patients with cardiomyopathy caused by a mutation in the cardiac protein MYBPC3.

In another approach, hiPSC-CMs can be plated on PDMS muscular thin films (MTFs) micropatterned with fibronectin so that they align. Upon release from the coverslip surface, shortening of cardiomyocytes during contraction causes MTFs to curl, with the displacement reflecting the force of contraction and thus allowing it to be calculated mathematically. MTFs were recently used to demonstrate lower forces of contraction in hiPSC-CMs derived from patients suffering from Barth syndrome, a mitochondrial disorder, that also causes severe cardiomyopathy (Wang et al., 2014).

It is becoming increasingly clear that interactions between multiple cardiac cell types benefit their survival, morphology, maturity, and function. Combined intramyocardial transplantation of hiPSC-derived cardiomyocytes, endothelial cells, and smooth muscle cells, for example, in an IGF-1-containing fibrin patch in a large animal model showed much higher survival of engrafted cells and greater improvement in cardiac function compared to cardiomyocytes alone (Ye et al., 2014). In another study hPSC-derived cardiomyocytes, endothelial cells, and human amniotic mesenchymal cells were combined in a 3D hydrogel. This improved survival and functional performance and changed the molecular profile after 4 and 6 weeks compared to cardiomyocytes alone (Burridge et al., 2014). Multicellular 3D aggregates have also been reported to be better in predicting cardiotoxicity than 2D configurations. 3D aggregates were combined in a microfluidic device and toxic effects of drugs were assessed as alterations in beating frequency by video imaging using an automated detection system (Bergström et al., 2015). Different compounds (doxorubicin, verapamil, and quinidine) were assessed for cardiotoxicity over a period of 6 hr, and although higher sensitivity could be reached using electrophysiology, these assays were efficient and less labor intensive.

Even more complex models are engineered heart tissues (EHTs) in which cardiac cells are combined with biomaterials and non-cardiac cells. EHTs based on collagen/Matrigel scaffolds with neonatal rat cardiomyocytes were first described more than a decade ago (Zimmermann et al., 2000), but have more recently been combined with hPSCs (Soong et al., 2012; Tulloch et al., 2011) or used fibrin/Matrigel scaffolds (Hirt et al., 2014). Although these 3D EHTs subjected to direct mechanical load display higher levels of maturity than standard 2D cardiomyocyte cultures, they are still not equivalent to adult cardiomyocytes. Various approaches reported to promote cardiac maturation in 2D, such as prolonged culture (for several months), the addition of growth factors like IGF and thyroid hormone T3, high oxygen levels, and various combinations of non-cardiomyocytes, also improved maturity of hPSC-CMs in EHTs. Recently, electrical stimulation of hiPSC-CM-derived EHTs for several weeks was shown to increase force generation by 50% and improve structural organization (Hirt et al., 2014). The benefit of 3D culture was demonstrated during the analysis of contractile force in hiPSC-CMs derived from a patient with mutations in the sarcomeric protein titin (Hinson et al., 2015). In 2D culture, no differences were observed between diseased and control cardiomyocytes, but in 3D the difference in contraction force, evidenced as the ability to displace the polymer in which the cardiomycoytes were suspended, was highly significant.

Brain

Neurons are the most important functional components of the brain. Abnormal behavior and function of neurons are considered primary causes of many neurological diseases and psychiatric disorders. Refinement of neural cell differentiation protocols from hPSCs over the last several years now means that pure populations of human forebrain (glutamatergic, presynaptic and postsynaptic cortical, and GABAergic interneurons), midbrain (dopaminergic) and hindbrain neurons, and their pathogenic counterparts are available for the study of mechanisms of (inherited) neural and neuropsychiatric disease initiation and progression (Bellin et al., 2012; Ho et al., 2015). Dysfunction of these neurons is implicated in the pathogenesis of Parkinson's disease (PD), schizophrenia (SZ), autism spectrum disorder (ASD)-like RETT syndrome, epilepsy, and seizure. Aspects of these disease phenotypes are cell autonomous. For example, peripheral neurons from hiPSCs of patients with familial dysautonomia show low expression of IKBKAP, a gene involved in transcriptional elongation. This manifests as defects in neurogenic differentiation and neuronal precursor migration and is corrected by the drug kinetin, which reduces the level of the mutant IKBKAP splice form through modification of mRNA splicing (Lee et al., 2009). In the case of PD, two genetic or familial forms have been described in which the hiPSC-derived dopaminergic neurons show a phenotype: in one, the neurons carried mutations in Leu-rich repeat kinase 2 (LRRK2) and showed increased susceptibility to oxidative stress (Batista et al., 2011), while in the other, there were three copies of the SNCA locus and the neurons showed increased alpha-synuclein protein levels (Devine et al., 2011). While the complex genetic nature of these PD models precludes the generation of isogenic controls by repair of the mutation, rare inherited forms of ALS have been amenable to this and provide a powerful example of the value of monotypic cultures, in this case of motor neurons. hiPSC lines from patients with a rare familial form of ALS caused by mutations in the SOD-1 gene (Wainger et al., 2014; Zhu et al., 2011), which encodes copper-zinc superoxide dismutase and protects cells against

reactive oxygen species, gave rise to derivative motor neurons that showed hyperexcitability compared with their genetically repaired controls. This mimicked the phenotype observed in patients in response to strong magnetic fields. More importantly though, the ALS motor neurons showed an imbalance in Na^+/K^+ by electrophysiology and this could be returned to the levels of that in the isogenic control by an anti-epileptic drug. This drug is already used clinically and is known to pass the blood-brain barrier (BBB). It then took less than 1 year for regulatory approval to repurpose this drug for trial in ALS patients—not just those with the genetic form but also forms with unknown origin because hiPSC motor neurons with an unrelated mutation also showed phenotypic rescue. Thus this approach represents an impressive example of the power of hiPSC disease modeling, particularly in drug repurposing.

Most recent developments include the ability to capture disease phenotypes and drug responses of patients in hiPSC-derived neurons for which the genetic cause is unknown and likely to be complex (Mertens et al., 2015b). However, neuropsychiatric disease is typically associated with deregulation of neuronal connectivity between diverse neuronal populations so that simply studying one neuronal subtype may fail to capture the phenotype of the disease (Spellman and Gordon, 2015). Even in the case of ALS in which the SOD1 gene seems to exert its effect autonomically in motor neurons, there was very early evidence in mice that the effects may be indirect and act by damaging adjacent astrocytes (Bruijn et al., 1997). Similar cell-non-autonomous effects appeared to underlie reduced synaptic puncti in hiPSC-derived motor neurons from patients with SMA (Ebert et al., 2009). SMA is a genetic disease evident in childhood, characterized by motor neuron loss thought to be due to reduced survival motor neuron (SMN). While SMN is expressed ubiquitously, reduced levels in astrocytes may cause their activation and result in the phenotype evident only in hiPSC motor neurons co-cultured with hiPSC astrocytes. The phenotype would again not be revealed without two different neuronal cell types being present. Neural cells from patient hiPSCs can also self-organize into defined neural circuits and 3D systems termed neural organoids. The first striking example of morphogenesis, cell polarity, and neural cell layer formation of neural organoids was seen in the optic cup from mouse ESCs (Eiraku et al., 2011) and later hESCs (Nakano et al., 2012), in which histological sections look remarkably like cross-sections of the eye with the multiple layers of the retina clearly formed. Cerebral organoids from hiPSCs have been used to examine the pathogenesis of neurodevelopmental disorders (Lancaster et al., 2013). These organoids consist of radial glia progenitor cells and neurons (Lancaster and Knoblich, 2014), which recapitulate the gene expression patterns and development of human fetal neocortex (Camp et al., 2015). They are thus potentially valuable models for disorders of neurodevelopment, particularly microencephaly, neurogenesis, and fate specification, conditions not well recapitulated in rodents. In hiPSCs generated from a microencephaly patient with a null mutation in centrosomal protein CDK5RAP2, it was reported that the cerebral organoids were depleted of neural progenitors and showed premature neuronal differentiation, recapitulating features of microencephaly. Finally, telencephalic and cortical organoids containing multiple neuronal cell types have been used to model early development of ASDs. ASD hiPSC-derived organoids were described as showing complex cellular phenotypes that included accelerated cell cycle, upregulation of genes directing gamma-amino butyric acid (GABA) neuron fate, increased synaptogenesis and dendrite outgrowth, and changes in synaptic activity (Mariani et al., 2015).

Future models will most likely incorporate neuronal circuits with, at minimum, two distinct neuronal cell types that form synapses: oligodendrocytes to provide myelination, and astrocytes and microglia to incorporate critical aspects of inflammation and synaptic pruning (reviewed in Haston and Finkbeiner, 2016). These circuits will need repeated stimulation electrically, with stress hormones or relevant drugs, to create clinically meaningful responses. At present hiPSC neurons mimic the molecular and cellular states before symptom onset, so they are presently better suited to study disease predisposition rather than the disease state itself. We refer readers to a recent discussion paper (Brennand et al., 2015) from a group of experts in this area that considers current challenges for creating meaningful patient-specific in vitro models to study brain disorders. The authors concluded that the convergence of findings between laboratories and patient cohorts provides optimism that hiPSC-based platforms will inform future drug discovery efforts, but critical technical challenges remain.

Recent studies demonstrated that reprogramming to the pluripotent state erases the memory of somatic cell origin or "epigenetic memory" but also eliminates the age-related features, such as telomere length and mitochondrial function (Studer et al., 2015). Modeling of age-associated diseases with late onset might be limited using hiPSC technology. Various approaches for "re-aging" hiPSC-derived cells are currently being explored (Cornacchia and Studer, 2015; Studer et al., 2015). The first successful attempts used the expression of progerin that induces a genetic form of premature aging; this effectively "aged" the cells and revealed the phenotype in neurons from PD patient hiPSCs (Miller et al., 2013). A recent study, however, showed that neurons derived by direct lineage reprogramming of somatic cells without an intermediate pluripotent state retained aged features, so age-related cellular defects were revealed without progerin expression (Mertens et al., 2015a).

Genetically engineered hiPSC lines can also be useful in neurotoxicity screening and this can potentially be expanded to other cell types. A powerful cytotoxicity assay has been described using neuronal (MAP2) and astrocytic (GFAP) lineage-specific knockin luciferase reporters (Pei et al., 2015a, 2015b). Interestingly, significant differences in responses were observed in neuronal cells and astrocytes. This further emphasizes the need to have different cell types in cytotoxicity screens for safe-pharmacology applications.

Emerging Principles: Advances from Heterotypic Cultures

As differentiation protocols and the ability to generate monotypic cell populations from hiPSCs improved, it became clear that the earlier mixed cell population that arose spontaneously in aggregates (or embryoid bodies) had several advantages: the differentiating cells created their own microenvironment or niche, which had appropriate organization and produced relevant extracellular matrix proteins. This was lost in the monotypic cultures. Therefore new types of heterotypic cultures began to emerge in which the combinations of cells were better controlled; these

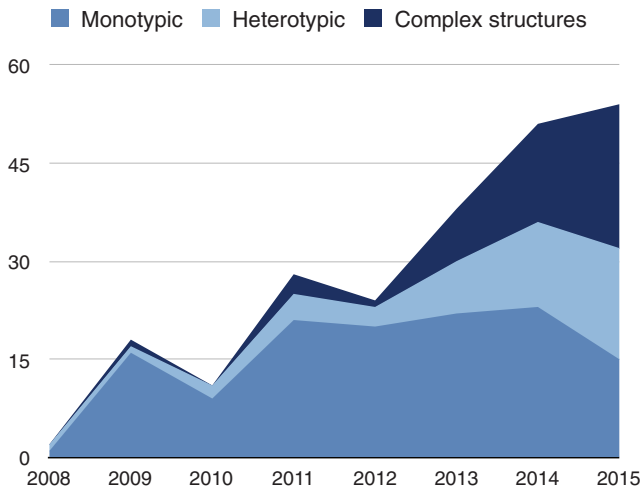


Figure 2. Number of Publications between 2008 and 2016 Describing the Use of Differentiated hiPSC for Disease and "Organ in a Dish" Modeling

PubMed Advanced Search Builder was used for the literature search with the following Builder: [(human pluripotent stem cells) AND differentiation AND (disease modeling)] OR [(human induced pluripotent stem cells) AND differentiation AND disease modeling] OR [(human pluripotent stem cells) AND organoids] OR [(human induced pluripotent stem cells) AND organoids] NOT review. Literature on lineage conversion or human embryonic stem cells was manually excluded. Publications that described the use of hiPSC for modeling neuronal, cardiovascular, and liver disease, as well as 3D organoid formation, were included. Citations were next exported to Papers citation manager and manuscripts were manually assigned a specific color-code for monotypic or heterotypic culture and an additional flag label for complex structures that combine micropatterned, microfluidic, and 2D and 3D microtissues and organoids. The numbers of publications per year for monotypic, heterotypic, and complex structures were then used to create the stacked area plot. The complete list of references included is available in the Supplemental Information. The authors would like to mention a limitation of this graphic representation related to the selection bias.

changes in the field are reflected in the number of publications each year over the last 10 years using embryoid body-, monotypic-, and heterotypic-type cultures (Figure 2).

This is particularly evident in tissues of organs that form later in development than the heart, brain, and vascular system such as the lung, kidney, and liver, which have been challenging to derive from hPSCs. Early successes relied on an exquisite knowledge of embryonic organ development to find the right protocol. The kidney, for example, is derived from intermediate mesoderm in the embryo and ultimately consists of up to 2 million nephrons, the structures that filter the blood, and 20 different cell types that collectively regulate excretion, regulation of pH, and the electrolyte and fluid balance. The liver is an example of how heterotypic cultures may lead to better liver tissue models based on hiPSCs. The liver is the largest internal organ in the human body. It regulates over 500 different functions that include metabolism of fats and proteins, bile production, excretion of drugs and hormones, and blood detoxification (Bhatia et al., 2014). Liver has a very complex architecture with hepatocytes organized as cords with other cells, including sinusoidal endothelial cells and biliary epithelial cells or cholangiocytes, positioned between the cords and organized as bile ducts.

Here we will discuss recent advances in generation of complex kidney and liver structures from hiPSCs.

Kidney

Drug nephrotoxicity is an important cause of acute kidney injury in hospitalized patients and currently there are no patient-specific models for assaying nephrotoxicity in vitro. An important advance in this context is the recent generation of complex kidney-like structures from hiPSCs (Morizane et al., 2015; Takasato et al., 2015). By controlling the timing of patterning within the hiPSC-derived intermediate mesoderm, it was recently shown that complex kidney organoids containing nephrons associated with a collecting duct network and surrounded by renal interstitium and endothelial cells could be generated. Comparison of transcriptional profiles of kidney organoids with human fetal tissues showed that they were very similar to first trimester human kidney (Roost et al., 2015). Since patient-derived hiPSCs could provide such opportunities, there has been significant interest in the observation that despite their immaturity, proximal tubules from hiPSCs showed the ability to endocytose dextran and differentially apoptose in response to cisplatin, an anticancer drug with proximal and distal tubular toxicity (Takasato et al., 2015), as well as to gentamicin, a commonly used antibiotic with proximal tubular toxicity (Morizane et al., 2015). In addition, several genetic kidney diseases have been introduced into hiPSCs using CRISPR/Cas9 to target relevant genes, creating isogenic pairs for direct comparison (Freedman et al., 2015). Cyst formation by kidney tubules in the hiPSC organoid model was disrupted when the polycystic kidney disease genes PKD1 or PKD2 were deleted. This was clearly distinct from effects on epithelium surrounding the lumen in epiblast spheroids from the same hiPSC lines; here the capacity to form lumen at these earlier stages of hiPSC differentiation was unaltered by the mutations. In addition, a clinical biomarker of proximal tubule injury, kidney injury molecule-1 (KIM-1), was upregulated only in the kidney organoids and not in hiPSC epiblast, indicating that the response was tissue specific and likely required the complex, multicellular 3D context. Although these structures exhibit lineage complexity that differs from conventional kidney cell lines and organoids, all cellular components of the developing proximal nephron (tubular cells, endothelial cells, nephron progenitors, and podocyte-like cells) are present in a well-organized way within each individual organoid.

Liver

During development, hepatic endoderm cells delaminate and form liver bud. Endothelial cells are essential for the initiation of the liver bud formation from hepatic endoderm, proliferation of the hepatic cells, and hepatic maturation (Matsumoto et al., 2001). In addition, inductive signals from endothelial cells also promote liver regeneration and hepatocyte proliferation upon injury (Ding et al., 2010, 2014; Hu et al., 2014). Furthermore, the presence of the vascular structures and blood flow is essential for the maturation of the hepatocytes in zebrafish (Korzh et al., 2008). Recently, Takebe et al. reported generation of functional human liver by co-culture of hiPSC-derived hepatic endoderm cells with the stromal cells composed of primary human umbilical vein endothelial cells (HUVECs) and human mesenchymal stem cells (hMSCs) (Takebe et al., 2013). The self-organized liver buds exhibited increase expression of mature hepatic markers as early as 6 days in co-culture. Furthermore, transplantation of these heterotypic structures into mouse resulted in rapid anastomosis with the host vasculature and formation of functional liver and extended expansion of

hiPSC-hepatocytes over a period of 2 months. More importantly, these hiPSC-derived mini-livers contained hepatocytes organized into the hepatic cord-like structures with characteristics of adult liver. Simultaneous induction of liver organoids upon pulse induction of GATA6 over a period of 5 days was reported recently (Guye et al., 2016). The GATA6 inductive approach resulted in the formation of definitive endodermal cells with hepatic characteristics, as well as mesoderm and neuroectoderm-derived vascular, hematopoietic, and neuronal cells. Prolonged culture over a period of 30 days resulted in the formation of functional tissues with a layer of hepatocytes and vascular cord-like structures. Therefore, this approach can be used for simultaneous induction of different cellular components from hiPSCs, which would be beneficial for transplantation studies and development of novel disease models.

Heterotypic interaction of NOTCH2 expressed on cholangiocytes with mesenchymal JAG1 is needed for the development of the bile duct structure (Hofmann et al., 2010). Recapitulation of these developmental principles facilitated successful differentiation of cholangiocytes from hiPSC-derived hapatoblasts via co-culture with JAG1-expressing stromal cells, OP9 (Ogawa et al., 2015). Hollow cyst-like structures were then generated by co-culture of aggregated hiPSC-derived hepatoblasts with OP9 cells in a 3D matrix of collagen and Matrigel. In addition, 3D conditions facilitated efficient tubulogenesis and growth of the aggregates. Interestingly, the cyst-like tubular structures were also generated without stromal cell layer if (adult stem cell) organoid-promoting culture conditions in 3D were used (Sampaziotis et al., 2015). In both cases, formation of the cyst-like structure was abrogated by NOTCH inhibition with γ-secretase inhibitor, indicating its importance in the maturation of the bile duct from hiPSC-derived cholangiocytes. Strikingly, both approaches resulted in functional cholangicytes with disease phenotypes when they were generated from patient hiPSCs with inherited polycystic liver disease or cystic fibrosis (CFTR F508del mutation). Administration of the drug (VX908), which is already in clinical trials for cystic fibrosis, resulted in the correction of the disease phenotype, demonstrating that this technology can be used to screen for potential therapeutic agents for bile disorders.

Other chronic liver diseases for which efforts are ongoing to create hiPSC-based models include cirrhosis, caused by alcohol abuse, drugs, virus infection, inflammation, or autoimmune and metabolic conditions, which accounts for more than 1 million deaths worldwide annually; and multiple inherited conditions that can cause liver damage such as alpha-1-antitrypsin deficiency, hemochromatosis, Wilson's disease, hereditary tyrosinemia, cystic fibrosis, and polycystic liver disease. A drawback at present to using hiPSC-derived liver cells in drug target discovery is that liver cell types other than hepatocytes, like oval cells, that would ideally be required for optimal disease and tissue modeling are not yet available. Progress here guided by developmental biology principles has a high likelihood in driving interest in adopting hiPSCs for this purpose.

Other Organs

Organoids are multicellular 3D structures or organs in miniature. They were until recently associated with adult stem cell cultures from tissues of the gastrointestinal tract (Huch and Koo, 2015; Johnson and Hockemeyer, 2015; Sato and Clevers, 2013). These organoids have now been derived from large and small intestine, pancreas, liver, stomach, and prostate. They generally contain the epithelial component of the tissue from which they are isolated in highly organized structures, but not the stromal tissue or vasculature. Organoids derived more recently from hiPSCs, however, often contain multiple tissue cell types as well as stromal cells and vasculature and are therefore considered heterotypic. Examples include (ectodermal) hiPSC-neural organoids, as discussed earlier, (mesodermal) hiPSC-kidney organoids, and (endodermal) hiPSC-intestinal, lung, gastric, and liver organoids (Dye et al., 2015; Lancaster et al., 2013; McCracken et al., 2014; Paşca et al., 2015; Spence et al., 2011; Takasato et al., 2015). If given the proper extracellular matrix, these structures show a remarkable ability to self-organize and develop the polarity seen in normal tissue. Organoids based on hiPSCs offer particular opportunities for disease modeling only partially met by adult-stem-cell-derived organoids: adult stem cells give rise to organoids with more mature phenotypes, but in general only the epithelial component of a tissue is represented.

What Are the Next Steps? Increasing Complexity and Multiple Integrated Readouts

Inclusion of Vasculature

Blood vessels not only supply tissues with nutrients and oxygen; they are also intimately involved in regulation of tissue morphogenesis, regeneration, and homeostasis, including the resolution of inflammation. During early development endothelial cells also play an important role in organogenesis. They provide instructive signals during heart morphogenesis and septation, are essential for the maturation of cardiomyocytes, and instruct development of endoderm-derived organs, such as liver, lung, kidney, and pancreas (Cleaver and Melton, 2003; Ding et al., 2014; Kao et al., 2015; Ramasamy et al., 2015).

Inclusion of the vasculature would therefore be an important next step in recreating complex tissue structures from hiPSCs. This was recently illustrated by a study in which "organ buds" made up of hPSC-derived tissue-specific progenitor cells were combined with endothelial cells and mesenchymal stromal cells (MSCs). The MSCs initiated condensation within the heterotypic cell mixtures. By defining optimal mechanical properties of the matrix, transplantable 3D organ buds could be formed from tissues that included kidney, pancreas, intestine, heart, lung, and brain. These organs buds were vascularized in vivo and self-organized into functional, tissue-specific structures (Takebe et al., 2015). Of note, each tissue and organ has its own type of endothelial cell, so tissue-specific blood vessel induction in these structures may be of importance (reviewed in Rafii et al., 2016; Ramasamy et al., 2015).

Aside from performing instructive roles in tissue morphogenesis, endothelial cells form vascular tubes, which require "mural cells" (pericytes and vSMCs) to develop into stable vasculature. Defective interactions between these cells underlie different genetic disorders that can cause hemorrhages and also be the cause of a spectrum of neurological conditions, such as hereditary hemorrhagic telangiectasia (HHT), cerebral autosomal-dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), and retinal vasculopathy with cerebral leukodystrophy (RVCL) (Yamamoto et al., 2011). Blood vessels "in a dish" that recapitulate endothelial-mural cell interactions would benefit investigation of mechanisms underlying these disorders, and they could also be used to study other

pathologies, such as thrombosis and vascular malformation disorders like cerebral cavernous malformations (CCMs) and others (Gibson et al., 2015; Storkebaum et al., 2011; Westein et al., 2013). Both 2D and 3D systems with multiple vascular cell types present would be essential in modeling these diseases.

Modeling Inflammation

Inflammation is a self-defense mechanism that protects organisms from infection and tissue injury. Chronic inflammation resulting from persistent infection and prolonged activation of the endothelium causes many pathological conditions, including cardiovascular, neurological, and neurodegenerative diseases. These are of particular concern because of their poor prognosis, significant morbidity, and the lack of effective treatments. In developing hiPSC disease models, especially for chronic and inflammatory conditions, it is therefore essential to take endothelial and inflammatory cells into account.

Modeling inflammatory responses is complex, and it requires assessment of interaction between different cells in the tissues, such as inflammatory cells, resident macrophages and microglia in the brain, endothelial cells, and epithelial cells in the lung and intestine. Through recent advances in hiPSC technology it has now become possible to differentiate many of the cellular components that would be useful to model inflammation in vitro. These include neutrophils, monocytes/macrophages, and microglia (Lachmann et al., 2015; Nayak et al., 2015; Schwartz et al., 2015). These cells serve as an initial wave of infiltrating cells at the site of tissue damage or infection and can exacerbate the inflammatory response. Interestingly, a comparison of hiPSC-derived macrophages with primary isogenic cells derived from peripheral blood demonstrated high phenotypic, functional, and transcriptomic similarity (Zhang et al., 2015).

Several groups have shown that endothelial cells from hiPSCs exhibit functional inflammatory responses (Adams et al., 2013; Belair et al., 2015; Patsch et al., 2015). hiPSC-derived endothelial cells also exhibit robust responses to proinflammatory stimuli (TNFα, IL-1β, and LPS) reflected in increased surface expression of adhesion receptors E-selectin, intercellular adhesion molecule 1 (ICAM-1), and vascular cell adhesion molecule 1 (VCAM-1). hiPSC-endothelial cells can bind human leukocytes under static conditions (Patsch et al., 2015) and under physiological flow that mimics blood in vivo (Adams et al., 2013). Derivation of autologous endothelial cells and leukocytes would be a major step forward in modeling uncontrolled inflammatory reactions with patient-specific cells. In particular, these might be useful for the pharmaceutical industry as assays for adverse drug responses (ADRs), much like those based on autologous primary human cells (Reed et al., 2015). The extreme ADR in the TGN1412 (an immunomodulatory drug) clinical trial that caused multiple organ failure in six young, healthy participants could perhaps have been avoided if a representative human in vitro assay had been available.

hiPSC-derived lung or intestine cells would be useful in modeling severe influenza reactions, Crohn's disease, and inflammatory bowel disease (IBD). A recent paper based on mouse intestinal stem cells would argue similarly for an important role of the immune system in promoting intestinal regeneration (Lindemans et al., 2015).The functionality of hiPSC-derived lung epithelial cells has recently been demonstrated in the case of life-threatening influenza reactions that occur in otherwise healthy children. The hypothesis that genetic factors could underlie influenza sensitivity was proposed some time ago and reviewed recently (Casanova, 2015). hiPSCs derived from a child with the severe influenza were used to demonstrate that the defective interferon (INF) response in lung-epithelial cells caused by a mutation in INF regulatory factor 7 (IFN7) is responsible for increased virus replication (Ciancanelli et al., 2015). Protocols to differentiate hiPSCs into appropriate lung cells made it possible to confirm that the severity of influenza might be due to inborn genetic errors in immunity, and that it might also be possible to develop therapies based on hiPSC models. Other conditions are similarly awaiting new developments in virus or bacterial hiPSC interaction.

Modeling the Blood-Brain Barrier

It is thought that a contributory factor to neurovascular disease may be initial vascular dysfunction that leads to the breakdown of the BBB. Vascular contributions to the development of cognitive impairment and dementia have become recognized over the last several years (Gorelick et al., 2011; Storkebaum et al., 2011; Zhao et al., 2015) and changes in the vasculature often precede neuronal defects. Patients with brain blood vessel malformations are at risk of developing neurological symptoms, usually because of hemorrhages, the incidence of which increases with age (Söderman et al., 2003). A vascular contribution is even becoming evident in AD. Deposition of β-amyloid (Aβ) peptide on the walls of brain capillary vessels, for example, is the common pathology of a condition called Cerebral Amyloid Angiopathy (CAA) (Hendriks et al., 1992), a major cause of hemorrhagic stroke in the elderly not associated with hypertension. Multiple recent studies also showed that CAA could cause further cognitive impairment, AD, or dementia. Other vascular pathologies have also been indicated to precede AD (Kanekiyo et al., 2014; Verghese et al., 2011). Understanding the mechanisms of neurovascular dysfunction and mechanisms that regulate BBB is therefore important for prevention and identification of potential drug targets for neurovascular diseases.

There are three major cellular components of the BBB: endothelial cells, pericytes/vSMCs, and astrocytes, which together form the specialized barrier that isolates and protects brain parenchyma from harmful components of the blood, facilitates active transport of nutrients, and mediates clearance of waste. Astrocytes and pericytes provide critical signals for the maturation and maintenance of the BBB and are thought to facilitate induction of BBB-like characteristics in endothelial cells. Therefore incorporation of all these components will be essential to recreate the BBB in vitro. Recently Lippmann et al. demonstrated that co-culture of rat neonatal astrocytes or human neuronal progenitor cells (NPCs) can be used to direct hiPSC endothelial cells to differentiate into BBB-like endothelial cells (Lippmann et al., 2012, 2014). Derivation of mature astrocytes from hiPSCs has been reported (Roybon et al., 2013; Sareen et al., 2014). Systems in which all of the cellular BBB components were derived from hiPSCs would be extremely valuable in identifying the causative cellular components in the disease pathology. Furthermore, inflammatory components could be incorporated via either reactive astrocytes stimulated with pro-inflammatory factors (TNFα, IL-1β, and IFNγ) (Roybon et al., 2013) or hiPSC-derived microglial or ''brain resident macrophages'' that could again be derived from hiPSCs (Schwartz et al., 2015).

Blood flow simulation and peripheral immune cells can be also incorporated as challenges by integration in micron-scale hollow tubes lined with endothelial cells separated from the pericyte/astrocyte or microglial components by thin porous membranes. Several prototypes of these "BBB-on-a-chip" devices have been described over the last 2 or 3 years (Booth and Kim, 2012; Brown et al., 2015; Cho et al., 2015; Deosarkar et al., 2015; Griep et al., 2013; Hyun Jo et al., 2015; Yeon et al., 2012). Future prototypes would incorporate not only the BBB compartment, but also a cerebral spinal fluid (CSF) compartment, currently being developed as part of the NIH-funded brain-on-chip program (Alcendor et al., 2013).

Additionally, sensors to measure transendothelial barrier resistance (TEER), electrical sensors to quantify infiltrating leukocytes, and label-free microbiosensors to measure amyloid beta isoforms, protein transport, or drug delivery across BBB are developing. These will transform the field and reduce the use of animal models in assessing the ability of drugs to cross the BBB.

Organs-on-Chips

Organs-on-chips are an emerging technology with excellent potentials for increasing tissue complexity of hiPSC tissue models and including vasculature. There are major initiatives in the US to promote the technology (https://ncats.nih.gov/; http://wyss.harvard.edu) as well as national initiatives in Switzerland (http://www.artorg.unibe.ch) and the Netherlands (http://www.hdmt.technology/). Organs-on-chips are microfluidic devices (or "chips") about the size of a microscope slide usually made of a transparent polymer and containing one or more small open or closed culture chambers (or micro-incubators) coupled to small (microfluidic) channels through which (culture) fluid or gas can flow. The cells, once seeded, proliferate or differentiate as in normal cell culture but may also mature or age because of a more physiological microenvironment than regular culture conditions. Organs-on-chips mimic the smallest functional subunits of human organ or tissue: the alveolus of a lung (lung-on-chip), synchronously contracting heart cells (heart-on-chip), intestine (gut-on-chip), and the like in a (micro)environment similar to that in vivo (Bhatia and Ingber, 2014; Wilmer et al., 2016). There is particular potential here for hiPSCs since these are now amenable to stable integration of reporter constructs (Den Hartogh and Passier, 2016). Most of the early organs-on-chips were based on primary cell cultures or transformed cell lines (Alonzo et al., 2015; Bertassoni et al., 2014; Bischel et al., 2013; Jeon et al., 2014; Kim et al., 2016; Moya et al., 2013; Nguyen et al., 2013; Sackmann et al., 2014; Wang et al., 2016; Zervantonakis et al., 2011, 2012; Zheng et al., 2012). Incorporation of hiPSC derivatives from patients or healthy individuals (or a combination of the two) is now widely considered. In particular, microfluidic models are now starting to be used to create vascular models with hiPSC derivatives (Belair et al., 2015; Mathur et al., 2015; Palpant et al., 2015; Theodoris et al., 2015; Wanjare et al., 2015). We recently used these microfluidic chambers to create 3D blood vessels from hPSC-derived endothelial cells and pericytes (van der Meer et al., 2013). Inclusion of cardiomyocytes into organ-on-chip devices that mimic blood flow and the endothelial-cardiomyocyte interface can improve prediction of drug-induced cardiotoxicity (Mathur et al., 2015). Incorporation of endothelial cells and the creation of heterotypic microphysiological systems are already dictating new directions in the field (Kurokawa and George, 2016). The open microfluidic systems also allow collection of the "flow through" for analysis of secreted proteins.

Finally, attempts are ongoing to differentiate hiPSCs directly in microfluidic devices. This depends on spatiotemporal control of the microenvironment by optimal delivery of exogenous factors and removal of cell-secreted factors under controlled perfusion frequency (Giobbe et al., 2015). This technology is again dependent on the intrinsic self-organization of hiPSCs to generate organotypic cultures and could mean that fewer cells would be needed to form functional microscale structures. Functional cardiac and hepatic cells have thus been obtained within 2 weeks of seeding and these showed expected drug responses in situ.

Organs-on-chips can thus yield unique biomedical data from hiPSCs through the integration of multicellular and multifactorial aspects of tissue physiology and disease.

Integrating Molecular and Functional Readouts

In all of these models, the greatest benefit will accrue if they support long-term, real-time analysis, including gene-based (fluorescent) reporters and electrical, mechanical, and bio-nanosensors for toxicity and disease (as both end-point and mechanistic readouts). Including genome-wide molecular analysis (genomics, transcriptomics, proteomics, and metabolomics) will be of importance but access will be required for appropriate (preferably repeated) cell sampling. Alternatively, as the sensitivity of detection methods for proteins increases, microfluidic devices in particular lend themselves to continuous monitoring of proteins in flow-through medium, and proteins secreted by cells developing disease phenotypes, challenged with drugs, or undergoing the effects of stress can be determined. Such (patho-)physiological information can normally only be obtained through animal testing. Systems biology approaches based on computational modeling will be required to integrate data from different hiPSC derivatives, establish relationships between the data, and identify clinically relevant endpoints. The importance of the multidisciplinary methodologies and expertise for this is beginning to be widely recognized.

Future Outlook

A key outstanding question for the field is whether hiPSCs will ultimately prove to be useful for disease modeling and drug discovery given what we now know about the challenges they pose. In addition and in light of rapidly evolving CRISPR technologies, human ESCs may turn out as the easier-to-standardize model system for studies of monogenetic diseases even though they have the disadvantage that information on the severity of the phenotype, age of onset, and drug responsiveness would not be known. Nevertheless, hiPSCs will likely be of special use as patient-specific reference models and/or be exploited in modeling complex diseases or diseases for which causative mutations are unknown.

Issues that still need to be addressed include the lack of mature phenotypes in both hESC and hiPSC derivatives, the extent of improvement over animal models for drug discovery, and the degree of complexity that can be recapitulated with in vitro stem cell models. To derive optimal benefit, it will be important for the field to focus on developing deeper complexity in these models and define useful readouts to enable their utility in drug discovery and safety pharmacology. An indirect outcome could be reduction in the use of laboratory animals because the alternatives represent better human mimics.

Questions and issues that need to be addressed in the future include:

- Do we need to integrate multiple organs on one chip? What are the advantages and the disadvantages?
- Will complex, highly advanced human models lead to a better predictability (compared to current models and human hPSC-derived high-throughput models)?
- Where will models fit in the process of drug discovery? It may be important to implement models at different stages during this process.
- How will possible findings from human hPSC-derived models be best linked with clinically relevant data?
- In the context of reproducibility and standardization, it is important to follow ''Good Cell Culture Practice''.
- Communication between the different stakeholders (scientists, industry, international regulatory bodies, and socially engaged organizations) is required for successful implementation of hPSC-derived models in the process of drug discovery and safety pharmacology and for the replacement of animal models.
- The relevance for regenerative medicine will need to be considered (for example, optogenetics in combination with hPSC-derived grafts in a Parkinson's disease model) (Steinbeck et al., 2015).

Multitypic cell cultures from hiPSCs could ultimately lead to better and safer drugs and a better understanding of human disease.

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes a detailed list of the references used to compile Figure 2 and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2016.02.011.

CONFLICTS OF INTEREST

Christine Mummery and Robert Passier are co-founders of Pluriomics bv. Christine Mummery is on the advisory board of Galapagos bv.

ACKNOWLEDGMENTS

This Review discusses a rapidly growing area of research and the authors apologize to those many other authors whose papers may not have been cited here. Research from the authors' laboratories is supported by the European Research Council (ERCAdG 323182 STEMCARDIOVASC), European Community's Seventh Framework Programme (FP7/2007-2013) grant agreement 602423, and ZonMw-MKMD-40-42600-98-036.

REFERENCES

Adams, W.J., Zhang, Y., Cloutier, J., Kuchimanchi, P., Newton, G., Sehrawat, S., Aird, W.C., Mayadas, T.N., Luscinskas, F.W., and García-Cardeña, G. (2013). Functional vascular endothelium derived from human induced pluripotent stem cells. Stem Cell Reports *1*, 105–113, http://dx.doi.org/10.1016/j.stemcr.2013.06.007.

Alcendor, D.J., Block, F.E., 3rd, Cliffel, D.E., Daniels, J.S., Ellacott, K.L.J., Goodwin, C.R., Hofmeister, L.H., Li, D., Markov, D.A., May, J.C., et al. (2013). Neurovascular unit on a chip: implications for translational applications. Stem Cell Res. Ther. *4* (*Suppl 1*), S18, http://dx.doi.org/10.1186/scrt379.

Alonzo, L.F., Moya, M.L., Shirure, V.S., and George, S.C. (2015). Microfluidic device to control interstitial flow-mediated homotypic and heterotypic cellular communication. Lab Chip *15*, 3521–3529, http://dx.doi.org/10.1039/C5LC00507H.

Batista, L.F.Z., Pech, M.F., Zhong, F.L., Nguyen, H.N., Xie, K.T., Zaug, A.J., Crary, S.M., Choi, J., Sebastiano, V., Cherry, A., et al. (2011). Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. Nature *474*, 399–402, http://dx.doi.org/10.1038/nature10084.

Belair, D.G., Whisler, J.A., Valdez, J., Velazquez, J., Molenda, J.A., Vickerman, V., Lewis, R., Daigh, C., Hansen, T.D., Mann, D.A., et al. (2015). Human vascular tissue models formed from human induced pluripotent stem cell derived endothelial cells. Stem Cell Rev. *11*, 511–525, http://dx.doi.org/10.1007/s12015-014-9549-5.

Bellin, M., Marchetto, M.C., Gage, F.H., and Mummery, C.L. (2012). Induced pluripotent stem cells: the new patient? Nat. Rev. Mol. Cell Biol. *13*, 713–726, http://dx.doi.org/10.1038/nrm3448.

Bellin, M., Casini, S., Davis, R.P., D'Aniello, C., Haas, J., Ward-van Oostwaard, D., Tertoolen, L.G.J., Jung, C.B., Elliott, D.A., Welling, A., et al. (2013). Isogenic human pluripotent stem cell pairs reveal the role of a KCNH2 mutation in long-QT syndrome. EMBO J. *32*, 3161–3175, http://dx.doi.org/10.1038/emboj.2013.240.

Bergström, G., Christoffersson, J., Schwanke, K., Zweigerdt, R., and Mandenius, C.-F. (2015). Stem cell derived in vivo-like human cardiac bodies in a microfluidic device for toxicity testing by beating frequency imaging. Lab Chip *15*, 3242–3249, http://dx.doi.org/10.1039/c5lc00449g.

Bertassoni, L.E., Cecconi, M., Manoharan, V., Nikkhah, M., Hjortnaes, J., Cristino, A.L., Barabaschi, G., Demarchi, D., Dokmeci, M.R., Yang, Y., and Khademhosseini, A. (2014). Hydrogel bioprinted microchannel networks for vascularization of tissue engineering constructs. Lab Chip *14*, 2202–2211, http://dx.doi.org/10.1039/C4LC00030G.

Bhatia, S.N., and Ingber, D.E. (2014). Microfluidic organs-on-chips. Nat. Biotechnol. *32*, 760–772, http://dx.doi.org/10.1038/nbt.2989.

Bhatia, S.N., Underhill, G.H., Zaret, K.S., and Fox, I.J. (2014). Cell and tissue engineering for liver disease. Science Translational Medicine *6*, 245sr2–245sr2. http://dx.doi.org/10.1126/scitranslmed.3005975.

Birket, M.J., Ribeiro, M.C., Kosmidis, G., Ward, D., Leitoguinho, A.R., van de Pol, V., Dambrot, C., Devalla, H.D., Davis, R.P., Mastroberardino, P.G., et al. (2015a). Contractile Defect Caused by Mutation in MYBPC3 Revealed under Conditions Optimized for Human PSC-Cardiomyocyte Function. Cell Rep. *13*, 733–745, http://dx.doi.org/10.1016/j.celrep.2015.09.025.

Birket, M.J., Ribeiro, M.C., Verkerk, A.O., Ward, D., Leitoguinho, A.R., den Hartogh, S.C., Orlova, V.V., Devalla, H.D., Schwach, V., Bellin, M., et al. (2015b). Expansion and patterning of cardiovascular progenitors derived from human pluripotent stem cells. Nat. Biotechnol. *33*, 970–979, http://dx.doi.org/10.1038/nbt.3271.

Bischel, L.L., Young, E.W.K., Mader, B.R., and Beebe, D.J. (2013). Tubeless microfluidic angiogenesis assay with three-dimensional endothelial-lined microvessels. Biomaterials *34*, 1471–1477, http://dx.doi.org/10.1016/j.biomaterials.2012.11.005.

Booth, R., and Kim, H. (2012). Characterization of a microfluidic in vitro model of the blood-brain barrier (μBBB). Lab Chip *12*, 1784–1792, http://dx.doi.org/10.1039/c2lc40094d.

Braam, S.R., Tertoolen, L., Casini, S., Matsa, E., Lu, H.R., Teisman, A., Passier, R., Denning, C., Gallacher, D.J., Towart, R., and Mummery, C.L. (2013). Repolarization reserve determines drug responses in human pluripotent stem cell derived cardiomyocytes. Stem Cell Res. (Amst.) *10*, 48–56, http://dx.doi.org/10.1016/j.scr.2012.08.007.

Bray, M.-A., Sheehy, S.P., and Parker, K.K. (2008). Sarcomere alignment is regulated by myocyte shape. Cell Motil. Cytoskeleton *65*, 641–651, http://dx.doi.org/10.1002/cm.20290.

Brennand, K.J., Marchetto, M.C., Benvenisty, N., Brüstle, O., Ebert, A., Izpisua Belmonte, J.C., Kaykas, A., Lancaster, M.A., Livesey, F.J., McConnell, M.J., et al. (2015). Creating Patient-Specific Neural Cells for the In Vitro Study of Brain Disorders. Stem Cell Reports *5*, 933–945, http://dx.doi.org/10.1016/j.stemcr.2015.10.011.

Bright, J., Hussain, S., Dang, V., Wright, S., Cooper, B., Byun, T., Ramos, C., Singh, A., Parry, G., Stagliano, N., and Griswold-Prenner, I. (2015). Human secreted tau increases amyloid-beta production. Neurobiol. Aging *36*, 693–709, http://dx.doi.org/10.1016/j.neurobiolaging.2014.09.007.

Brown, J.A., Pensabene, V., Markov, D.A., Allwardt, V., Neely, M.D., Shi, M., Britt, C.M., Hoilett, O.S., Yang, Q., Brewer, B.M., et al. (2015). Recreating

blood-brain barrier physiology and structure on chip: A novel neurovascular microfluidic bioreactor. Biomicrofluidics *9*, 054124, http://dx.doi.org/10.1063/1.4934713.

Bruijn, L.I., Becher, M.W., Lee, M.K., Anderson, K.L., Jenkins, N.A., Copeland, N.G., Sisodia, S.S., Rothstein, J.D., Borchelt, D.R., Price, D.L., and Cleveland, D.W. (1997). ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. Neuron *18*, 327–338.

Burridge, P.W., Keller, G., Gold, J.D., and Wu, J.C. (2012). Production of de novo cardiomyocytes: human pluripotent stem cell differentiation and direct reprogramming. Cell Stem Cell *10*, 16–28, http://dx.doi.org/10.1016/j.stem.2011.12.013.

Burridge, P.W., Metzler, S.A., Nakayama, K.H., Abilez, O.J., Simmons, C.S., Bruce, M.A., Matsuura, Y., Kim, P., Wu, J.C., Butte, M., et al. (2014). Multicellular interactions sustain long-term contractility of human pluripotent stem cell-derived cardiomyocytes. Am. J. Transl. Res. *6*, 724–735.

Camp, J.G., Badsha, F., Florio, M., Kanton, S., Gerber, T., Wilsch-Bräuninger, M., Lewitus, E., Sykes, A., Hevers, W., Lancaster, M., et al. (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. Proc. Natl. Acad. Sci. USA *112*, 15672–15677, http://dx.doi.org/10.1073/pnas.1520760112.

Casanova, J.-L. (2015). Severe infectious diseases of childhood as monogenic inborn errors of immunity. Proc. Natl. Acad. Sci. USA *112*, E7128–E7137, http://dx.doi.org/10.1073/pnas.1521651112.

Cho, H., Seo, J.H., Wong, K.H.K., Terasaki, Y., Park, J., Bong, K., Arai, K., Lo, E.H., and Irimia, D. (2015). Three-Dimensional Blood-Brain Barrier Model for in vitro Studies of Neurovascular Pathology. Sci. Rep. *5*, 15222, http://dx.doi.org/10.1038/srep15222.

Ciancanelli, M.J., Huang, S.X.L., Luthra, P., Garner, H., Itan, Y., Volpi, S., Lafaille, F.G., Trouillet, C., Schmolke, M., Albrecht, R.A., et al. (2015). Infectious disease. Life-threatening influenza and impaired interferon amplification in human IRF7 deficiency. Science *348*, 448–453, http://dx.doi.org/10.1126/science.aaa1578.

Cleaver, O., and Melton, D.A. (2003). Endothelial signaling during development. Nat. Med. *9*, 661–668, http://dx.doi.org/10.1038/nm0603-661.

Cornacchia, D., and Studer, L. (2015). Back and forth in time: Directing age in iPSC-derived lineages. Brain Res. http://dx.doi.org/10.1016/j.brainres.2015.11.013, S0006-8993(15)00857-4.

Deosarkar, S.P., Prabhakarpandian, B., Wang, B., Sheffield, J.B., Krynska, B., and Kiani, M.F. (2015). A Novel Dynamic Neonatal Blood-Brain Barrier on a Chip. PLoS ONE *10*, e0142725–e21. http://dx.doi.org/10.1371/journal.pone.0142725.

Devalla, H.D., Schwach, V., Ford, J.W., Milnes, J.T., El-Haou, S., Jackson, C., Gkatzis, K., Elliott, D.A., Chuva de Sousa Lopes, S.M., Mummery, C.L., et al. (2015). Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology. EMBO Mol. Med. *7*, 394–410, http://dx.doi.org/10.15252/emmm.201404757.

Devine, M.J., Ryten, M., Vodicka, P., Thomson, A.J., Burdon, T., Houlden, H., Cavaleri, F., Nagano, M., Drummond, N.J., Taanman, J.-W., et al. (2011). Parkinson's disease induced pluripotent stem cells with triplication of the α-synuclein locus. Nat. Commun. *2*, 440, http://dx.doi.org/10.1038/ncomms1453.

Ding, B.-S., Nolan, D.J., Butler, J.M., James, D., Babazadeh, A.O., Rosenwaks, Z., Mittal, V., Kobayashi, H., Shido, K., Lyden, D., et al. (2010). Inductive angiocrine signals from sinusoidal endothelium are required for liver regeneration. Nature *468*, 310–315, http://dx.doi.org/10.1038/nature09493.

Ding, B.-S., Cao, Z., Lis, R., Nolan, D.J., Guo, P., Simons, M., Penfold, M.E., Shido, K., Rabbany, S.Y., and Rafii, S. (2014). Divergent angiocrine signals from vascular niche balance liver regeneration and fibrosis. Nature *505*, 97–102, http://dx.doi.org/10.1038/nature12681.

Den Hartogh, S.C., and Passier, R. (2016). Concise Review: Fluorescent Reporters in Human Pluripotent Stem Cells: Contributions to Cardiac Differentiation and Their Applications in Cardiac Disease and Toxicity. Stem Cells *34*, 13–26, http://dx.doi.org/10.1002/stem.2196.

Dye, B.R., Hill, D.R., Ferguson, M.A.H., Tsai, Y.-H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., et al. (2015). In vitro generation of human pluripotent stem cell derived lung organoids. eLife *4*, 1999, http://dx.doi.org/10.7554/eLife.05098.

Ebert, A.D., Yu, J., Rose, F.F.J., Jr., Mattis, V.B., Lorson, C.L., Thomson, J.A., and Svendsen, C.N. (2009). Induced pluripotent stem cells from a spinal muscular atrophy patient. Nature *457*, 277–280, http://dx.doi.org/10.1038/nature07677.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. Nature *472*, 51–56, http://dx.doi.org/10.1038/nature09941.

Elliott, D.A., Braam, S.R., Koutsis, K., Ng, E.S., Jenny, R., Lagerqvist, E.L., Biben, C., Hatzistavrou, T., Hirst, C.E., Yu, Q.C., et al. (2011). NKX2-5(eGFP/w) hESCs for isolation of human cardiac progenitors and cardiomyocytes. Nat. Methods *8*, 1037–1040, http://dx.doi.org/10.1038/nmeth.1740.

Freedman, B.S., Brooks, C.R., Lam, A.Q., Fu, H., Morizane, R., Agrawal, V., Saad, A.F., Li, M.K., Hughes, M.R., Werff, R.V., et al. (2015). Modelling kidney disease with CRISPR-mutant kidney organoids derived from human pluripotent epiblast spheroids. Nat. Commun. *6*, 8715, http://dx.doi.org/10.1038/ncomms9715.

Gibson, C.C., Zhu, W., Davis, C.T., Bowman-Kirigin, J.A., Chan, A.C., Ling, J., Walker, A.E., Goitre, L., Delle Monache, S., Retta, S.F., et al. (2015). Strategy for identifying repurposed drugs for the treatment of cerebral cavernous malformation. Circulation *131*, 289–299, http://dx.doi.org/10.1161/CIRCULATIONAHA.114.010403.

Giobbe, G.G., Michielin, F., Luni, C., Giulitti, S., Martewicz, S., Dupont, S., Floreani, A., and Elvassore, N. (2015). Functional differentiation of human pluripotent stem cells on a chip. Nat. Methods *12*, 637–640, http://dx.doi.org/10.1038/nmeth.3411.

Glass, C.K., Saijo, K., Winner, B., Marchetto, M.C., and Gage, F.H. (2010). Mechanisms underlying inflammation in neurodegeneration. Cell *140*, 918–934, http://dx.doi.org/10.1016/j.cell.2010.02.016.

Gorelick, P.B., Scuteri, A., Black, S.E., Decarli, C., Greenberg, S.M., Iadecola, C., Launer, L.J., Laurent, S., Lopez, O.L., Nyenhuis, D., et al.; American Heart Association Stroke Council, Council on Epidemiology and Prevention, Council on Cardiovascular Nursing, Council on Cardiovascular Radiology and Intervention, and Council on Cardiovascular Surgery and Anesthesia (2011). Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the american heart association/american stroke association. Stroke *42*, 2672–2713, http://dx.doi.org/10.1161/STR.0b013e3182299496.

Griep, L.M., Wolbers, F., de Wagenaar, B., ter Braak, P.M., Weksler, B.B., Romero, I.A., Couraud, P.O., Vermes, I., van der Meer, A.D., and van den Berg, A. (2013). BBB on chip: microfluidic platform to mechanically and biochemically modulate blood-brain barrier function. Biomed. Microdevices *15*, 145–150, http://dx.doi.org/10.1007/s10544-012-9699-7.

Guye, P., Ebrahimkhani, M.R., Kipniss, N., Velazquez, J.J., Schoenfeld, E., Kiani, S., Griffith, L.G., and Weiss, R. (2016). Genetically engineering self-organization of human pluripotent stem cells into a liver bud-like tissue using Gata6. Nat. Commun. *7*, 10243, http://dx.doi.org/10.1038/ncomms10243.

Han, L., Li, Y., Tchao, J., Kaplan, A.D., Lin, B., Li, Y., Mich-Basso, J., Lis, A., Hassan, N., London, B., et al. (2014). Study familial hypertrophic cardiomyopathy using patient-specific induced pluripotent stem cells. Cardiovasc. Res. *104*, 258–269, http://dx.doi.org/10.1093/cvr/cvu205.

Haston, K.M., and Finkbeiner, S. (2016). Clinical Trials in a Dish: The Potential of Pluripotent Stem Cells to Develop Therapies for Neurodegenerative Diseases. Annu. Rev. Pharmacol. Toxicol. *56*, 489–510, http://dx.doi.org/10.1146/annurev-pharmtox-010715-103548.

Hendriks, L., van Duijn, C.M., Cras, P., Cruts, M., Van Hul, W., van Harskamp, F., Warren, A., McInnis, M.G., Antonarakis, S.E., Martin, J.J., et al. (1992). Presenile dementia and cerebral haemorrhage linked to a mutation at codon 692 of the beta-amyloid precursor protein gene. Nat. Genet. *1*, 218–221, http://dx.doi.org/10.1038/ng0692-218.

Hinson, J.T., Chopra, A., Nafissi, N., Polacheck, W.J., Benson, C.C., Swist, S., Gorham, J., Yang, L., Schafer, S., Sheng, C.C., et al. (2015). HEART DISEASE. Titin mutations in iPS cells define sarcomere insufficiency as a cause of dilated cardiomyopathy. Science *349*, 982–986, http://dx.doi.org/10.1126/science.aaa5458.

Hirt, M.N., Boeddinghaus, J., Mitchell, A., Schaaf, S., Börnchen, C., Müller, C., Schulz, H., Hubner, N., Stenzig, J., Stoehr, A., et al. (2014). Functional improvement and maturation of rat and human engineered heart tissue by

chronic electrical stimulation. J. Mol. Cell. Cardiol. *74*, 151–161, http://dx.doi.org/10.1016/j.yjmcc.2014.05.009.

Ho, S.-M., Topol, A., and Brennand, K.J. (2015). From "directed differentiation" to "neuronal induction": modeling neuropsychiatric disease. Biomark. Insights *10* (*Suppl 1*), 31–41, http://dx.doi.org/10.4137/BMI.S20066.

Hofmann, J.J., Zovein, A.C., Koh, H., Radtke, F., Weinmaster, G., and Iruela-Arispe, M.L. (2010). Jagged1 in the portal vein mesenchyme regulates intrahepatic bile duct development: insights into Alagille syndrome. Development *137*, 4061–4072, http://dx.doi.org/10.1242/dev.052118.

Hu, J., Srivastava, K., Wieland, M., Runge, A., Mogler, C., Besemfelder, E., Terhardt, D., Vogel, M.J., Cao, L., Korn, C., et al. (2014). Endothelial cell-derived angiopoietin-2 controls liver regeneration as a spatiotemporal rheostat. Science *343*, 416–419, http://dx.doi.org/10.1126/science.1244880.

Huch, M., and Koo, B.K. (2015). Modeling mouse and human development using organoid cultures. Development *142*, 3113–3125, http://dx.doi.org/10.1242/dev.118570.

Hyun Jo, D., Lee, R., Hyoung Kim, J., Oh Jun, H., Geol Lee, T., and Hun Kim, J. (2015). Real-time estimation of paracellular permeability of cerebral endothelial cells by capacitance sensor array. Sci. Rep. *5*, 11014, http://dx.doi.org/10.1038/srep11014.

Ionta, V., Liang, W., Kim, E.H., Rafie, R., Giacomello, A., Marbán, E., and Cho, H.C. (2015). SHOX2 overexpression favors differentiation of embryonic stem cells into cardiac pacemaker cells, improving biological pacing ability. Stem Cell Reports *4*, 129–142, http://dx.doi.org/10.1016/j.stemcr.2014.11.004.

Iyer, D., Gambardella, L., Bernard, W.G., Serrano, F., Mascetti, V.L., Pedersen, R.A., Talasila, A., and Sinha, S. (2015). Robust derivation of epicardium and its differentiated smooth muscle cell progeny from human pluripotent stem cells. Development *142*, 1528–1541, http://dx.doi.org/10.1242/dev.119271.

Jeon, J.S., Bersini, S., Whisler, J.A., Chen, M.B., Dubini, G., Charest, J.L., Moretti, M., and Kamm, R.D. (2014). Generation of 3D functional microvascular networks with human mesenchymal stem cells in microfluidic systems. Integr Biol (Camb) *6*, 555–563, http://dx.doi.org/10.1039/C3IB40267C.

Johnson, J.Z., and Hockemeyer, D. (2015). Human stem cell-based disease modeling: prospects and challenges. Curr. Opin. Cell Biol. *37*, 84–90, http://dx.doi.org/10.1016/j.ceb.2015.10.007.

Jung, J.J., Husse, B., Rimmbach, C., Krebs, S., Stieber, J., Steinhoff, G., Dendorfer, A., Franz, W.-M., and David, R. (2014). Programming and isolation of highly pure physiologically and pharmacologically functional sinus-nodal bodies from pluripotent stem cells. Stem Cell Reports *2*, 592–605, http://dx.doi.org/10.1016/j.stemcr.2014.03.006.

Kanekiyo, T., Xu, H., and Bu, G. (2014). ApoE and Aβ in Alzheimer's disease: accidental encounters or partners? Neuron *81*, 740–754, http://dx.doi.org/10.1016/j.neuron.2014.01.045.

Kao, D.-I., Lacko, L.A., Ding, B.-S., Huang, C., Phung, K., Gu, G., Rafii, S., Stuhlmann, H., and Chen, S. (2015). Endothelial cells control pancreatic cell fate at defined stages through EGFL7 signaling. Stem Cell Reports *4*, 181–189, http://dx.doi.org/10.1016/j.stemcr.2014.12.008.

Karakikes, I., Ameen, M., Termglinchan, V., and Wu, J.C. (2015). Human induced pluripotent stem cell-derived cardiomyocytes: insights into molecular, cellular, and functional phenotypes. Circ. Res. *117*, 80–88, http://dx.doi.org/10.1161/CIRCRESAHA.117.305365.

Kim, C., Wong, J., Wen, J., Wang, S., Wang, C., Spiering, S., Kan, N.G., Forcales, S., Puri, P.L., Leone, T.C., et al. (2013). Studying arrhythmogenic right ventricular dysplasia with patient-specific iPSCs. Nature *494*, 105–110, http://dx.doi.org/10.1038/nature11799.

Kim, S., Chung, M., and Jeon, N.L. (2016). Three-dimensional biomimetic model to reconstitute sprouting lymphangiogenesis in vitro. Biomaterials *78*, 115–128, http://dx.doi.org/10.1016/j.biomaterials.2015.11.019.

Korzh, S., Pan, X., Garcia-Lecea, M., Winata, C.L., Pan, X., Wohland, T., Korzh, V., and Gong, Z. (2008). Requirement of vasculogenesis and blood circulation in late stages of liver growth in zebrafish. BMC Dev. Biol. *8*, 84, http://dx.doi.org/10.1186/1471-213X-8-84.

Kurokawa, Y.K., and George, S.C. (2016). Tissue engineering the cardiac microenvironment: Multicellular microphysiological systems for drug screening. Adv. Drug Deliv. Rev. *96*, 225–233, http://dx.doi.org/10.1016/j.addr.2015.07.004.

Lachmann, N., Ackermann, M., Frenzel, E., Liebhaber, S., Brennig, S., Happle, C., Hoffmann, D., Klimenkova, O., Lüttge, D., Buchegger, T., et al. (2015). Large-scale hematopoietic differentiation of human induced pluripotent stem cells provides granulocytes or macrophages for cell replacement therapies. Stem Cell Reports *4*, 282–296, http://dx.doi.org/10.1016/j.stemcr.2015.01.005.

Lan, F., Lee, A.S., Liang, P., Sanchez-Freire, V., Nguyen, P.K., Wang, L., Han, L., Yen, M., Wang, Y., Sun, N., et al. (2013). Abnormal calcium handling properties underlie familial hypertrophic cardiomyopathy pathology in patient-specific induced pluripotent stem cells. Cell Stem Cell *12*, 101–113, http://dx.doi.org/10.1016/j.stem.2012.10.010.

Lancaster, M.A., and Knoblich, J.A. (2014). Generation of cerebral organoids from human pluripotent stem cells. Nat. Protoc. *9*, 2329–2340, http://dx.doi.org/10.1038/nprot.2014.158.

Lancaster, M.A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. Nature *501*, 373–379, http://dx.doi.org/10.1038/nature12517.

Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., Ganat, Y.M., Menon, J., Shimizu, F., Viale, A., et al. (2009). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. Nature *461*, 402–406, http://dx.doi.org/10.1038/nature08320.

Lindemans, C.A., Calafiore, M., Mertelsmann, A.M., O'Connor, M.H., Dudakov, J.A., Jenq, R.R., Velardi, E., Young, L.F., Smith, O.M., Lawrence, G., et al. (2015). Interleukin-22 promotes intestinal-stem-cell-mediated epithelial regeneration. Nature *528*, 560–564, http://dx.doi.org/10.1038/nature16460.

Lippmann, E.S., Azarin, S.M., Kay, J.E., Nessler, R.A., Wilson, H.K., Al-Ahmad, A., Palecek, S.P., and Shusta, E.V. (2012). Derivation of blood-brain barrier endothelial cells from human pluripotent stem cells. Nat. Biotechnol. *30*, 783–791, http://dx.doi.org/10.1038/nbt.2247.

Lippmann, E.S., Al-Ahmad, A., Azarin, S.M., Palecek, S.P., and Shusta, E.V. (2014). A retinoic acid-enhanced, multicellular human blood-brain barrier model derived from stem cell sources. Sci. Rep. *4*, 4160, http://dx.doi.org/10.1038/srep04160.

Ma, D., Wei, H., Lu, J., Ho, S., Zhang, G., Sun, X., Oh, Y., Tan, S.H., Ng, M.L., Shim, W., et al. (2013). Generation of patient-specific induced pluripotent stem cell-derived cardiomyocytes as a cellular model of arrhythmogenic right ventricular cardiomyopathy. Eur. Heart J. *34*, 1122–1133, http://dx.doi.org/10.1093/eurheartj/ehs226.

Mariani, J., Coppola, G., Zhang, P., Abyzov, A., Provini, L., Tomasini, L., Amenduni, M., Szekely, A., Palejev, D., Wilson, M., et al. (2015). FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. Cell *162*, 375–390, http://dx.doi.org/10.1016/j.cell.2015.06.034.

Mathur, A., Loskill, P., Shao, K., Huebsch, N., Hong, S., Marcus, S.G., Marks, N., Mandegar, M., Conklin, B.R., Lee, L.P., and Healy, K.E. (2015). Human iPSC-based cardiac microphysiological system for drug screening applications. Sci. Rep. *5*, 8883, http://dx.doi.org/10.1038/srep08883.

Matsumoto, K., Yoshitomi, H., Rossant, J., and Zaret, K.S. (2001). Liver organogenesis promoted by endothelial cells prior to vascular function. Science *294*, 559–563, http://dx.doi.org/10.1126/science.1063889.

McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.-H., Mayhew, C.N., Spence, J.R., Zavros, Y., and Wells, J.M. (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. Nature *516*, 400–404, http://dx.doi.org/10.1038/nature13863.

McNeish, J., Gardner, J.P., Wainger, B.J., Woolf, C.J., and Eggan, K. (2015). From Dish to Bedside: Lessons Learned While Translating Findings from a Stem Cell Model of Disease to a Clinical Trial. Cell Stem Cell *17*, 8–10, http://dx.doi.org/10.1016/j.stem.2015.06.013.

Merkle, F.T., and Eggan, K. (2013). Modeling human disease with pluripotent stem cells: from genome association to function. Cell Stem Cell *12*, 656–668, http://dx.doi.org/10.1016/j.stem.2013.05.016.

Mertens, J., Paquola, A.C.M., Ku, M., Hatch, E., Böhnke, L., Ladjevardi, S., McGrath, S., Campbell, B., Lee, H., Herdy, J.R., et al. (2015a). Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. Cell Stem Cell *17*, 705–718, http://dx.doi.org/10.1016/j.stem.2015.09.001.

Mertens, J., Wang, Q.-W., Kim, Y., Yu, D.X., Pham, S., Yang, B., Zheng, Y., Diffenderfer, K.E., Zhang, J., Soltani, S., et al.; Pharmacogenomics of Bipolar Disorder Study (2015b). Differential responses to lithium in hyperexcitable neurons from patients with bipolar disorder. Nature *527*, 95–99, http://dx.doi.org/10.1038/nature15526.

Miller, J.D., Ganat, Y.M., Kishinevsky, S., Bowman, R.L., Liu, B., Tu, E.Y., Mandal, P.K., Vera, E., Shim, J.-W., Kriks, S., et al. (2013). Human iPSC-based modeling of late-onset disease via progerin-induced aging. Cell Stem Cell *13*, 691–705, http://dx.doi.org/10.1016/j.stem.2013.11.006.

Morizane, R., Lam, A.Q., Freedman, B.S., Kishi, S., Valerius, M.T., and Bonventre, J.V. (2015). Nephron organoids derived from human pluripotent stem cells model kidney development and injury. Nat. Biotechnol. *33*, 1193–1200, http://dx.doi.org/10.1038/nbt.3392.

Moya, M.L., Hsu, Y.-H., Lee, A.P., Hughes, C.C.W., and George, S.C. (2013). In vitro perfused human capillary networks. Tissue Eng. Part C Methods *19*, 730–737, http://dx.doi.org/10.1089/ten.tec.2012.0430.

Mummery, C.L., Zhang, J., Ng, E.S., Elliott, D.A., Elefanty, A.G., and Kamp, T.J. (2012). Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. Circ. Res. *111*, 344–358, http://dx.doi.org/10.1161/CIRCRESAHA.110.227512.

Nakano, T., Ando, S., Takata, N., Kawada, M., Muguruma, K., Sekiguchi, K., Saito, K., Yonemura, S., Eiraku, M., and Sasai, Y. (2012). Self-formation of optic cups and storable stratified neural retina from human ESCs. Cell Stem Cell *10*, 771–785, http://dx.doi.org/10.1016/j.stem.2012.05.009.

Naryshkin, N.A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., Ling, K.K.Y., Karp, G.M., Qi, H., Woll, M.G., et al. (2014). Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. Science *345*, 688–693, http://dx.doi.org/10.1126/science.1250127.

Nayak, R.C., Trump, L.R., Aronow, B.J., Myers, K., Mehta, P., Kalfa, T., Wellendorf, A.M., Valencia, C.A., Paddison, P.J., Horwitz, M.S., et al. (2015). Pathogenesis of ELANE-mutant severe neutropenia revealed by induced pluripotent stem cells. J. Clin. Invest. *125*, 3103–3116, http://dx.doi.org/10.1172/JCI80924.

Nguyen, D.-H.T., Stapleton, S.C., Yang, M.T., Cha, S.S., Choi, C.K., Galie, P.A., and Chen, C.S. (2013). Biomimetic model to reconstitute angiogenic sprouting morphogenesis in vitro. Proc. Natl. Acad. Sci. USA *110*, 6712–6717, http://dx.doi.org/10.1073/pnas.1221526110.

Ogawa, M., Ogawa, S., Bear, C.E., Ahmadi, S., Chin, S., Li, B., Grompe, M., Keller, G., Kamath, B.M., and Ghanekar, A. (2015). Directed differentiation of cholangiocytes from human pluripotent stem cells. Nat. Biotechnol. *33*, 853–861, http://dx.doi.org/10.1038/nbt.3294.

Palpant, N.J., Pabon, L., Roberts, M., Hadland, B., Jones, D., Jones, C., Moon, R.T., Ruzzo, W.L., Bernstein, I., Zheng, Y., and Murry, C.E. (2015). Inhibition of β-catenin signaling respecifies anterior-like endothelium into beating human cardiomyocytes. Development *142*, 3198–3209, http://dx.doi.org/10.1242/dev.117010.

Park, I.-H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008). Disease-specific induced pluripotent stem cells. Cell *134*, 877–886, http://dx.doi.org/10.1016/j.cell.2008.07.041.

Paşca, A.M., Sloan, S.A., Clarke, L.E., Tian, Y., Makinson, C.D., Huber, N., Kim, C.H., Park, J.-Y., O'Rourke, N.A., Nguyen, K.D., et al. (2015). Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. Nat. Methods *12*, 671–678, http://dx.doi.org/10.1038/nmeth.3415.

Patsch, C., Challet-Meylan, L., Thoma, E.C., Urich, E., Heckel, T., O'Sullivan, J.F., Grainger, S.J., Kapp, F.G., Sun, L., Christensen, K., et al. (2015). Generation of vascular endothelial and smooth muscle cells from human pluripotent stem cells. Nat. Cell Biol. *17*, 994–1003, http://dx.doi.org/10.1038/ncb3205.

Pei, Y., Peng, J., Behl, M., Sipes, N.S., Shockley, K.R., Rao, M.S., Tice, R.R., and Zeng, X. (2015a). Comparative neurotoxicity screening in human iPSC-derived neural stem cells, neurons and astrocytes. Brain Res. http://dx.doi.org/10.1016/j.brainres.2015.07.048, S0006-8993(15)00593-4.

Pei, Y., Sierra, G., Sivapatham, R., Swistowski, A., Rao, M.S., and Zeng, X. (2015b). A platform for rapid generation of single and multiplexed reporters in human iPSC lines. Sci. Rep. *5*, 9205, http://dx.doi.org/10.1038/srep09205.

Rafii, S., Butler, J.M., and Ding, B.-S. (2016). Angiocrine functions of organ-specific endothelial cells. Nature *529*, 316–325, http://dx.doi.org/10.1038/nature17040.

Ramasamy, S.K., Kusumbe, A.P., and Adams, R.H. (2015). Regulation of tissue morphogenesis by endothelial cell-derived signals. Trends Cell Biol. *25*, 148–157, http://dx.doi.org/10.1016/j.tcb.2014.11.007.

Reed, D.M., Paschalaki, K.E., Starke, R.D., Mohamed, N.A., Sharp, G., Fox, B., Eastwood, D., Bristow, A., Ball, C., Vessillier, S., et al. (2015). An autologous endothelial cell:peripheral blood mononuclear cell assay that detects cytokine storm responses to biologics. FASEB J. *29*, 2595–2602, http://dx.doi.org/10.1096/fj.14-268144.

Ribeiro, A.J.S., Ang, Y.-S., Fu, J.-D., Rivas, R.N., Mohamed, T.M.A., Higgs, G.C., Srivastava, D., and Pruitt, B.L. (2015a). Contractility of single cardiomyocytes differentiated from pluripotent stem cells depends on physiological shape and substrate stiffness. Proc. Natl. Acad. Sci. USA *112*, 12705–12710, http://dx.doi.org/10.1073/pnas.1508073112.

Ribeiro, M.C., Tertoolen, L.G., Guadix, J.A., Bellin, M., Kosmidis, G., D'Aniello, C., Monshouwer-Kloots, J., Goumans, M.-J., Wang, Y.-L., Feinberg, A.W., et al. (2015b). Functional maturation of human pluripotent stem cell derived cardiomyocytes in vitro–correlation between contraction force and electrophysiology. Biomaterials *51*, 138–150, http://dx.doi.org/10.1016/j.biomaterials.2015.01.067.

Roost, M.S., van Iperen, L., Ariyurek, Y., Buermans, H.P., Arindrarto, W., Devalla, H.D., Passier, R., Mummery, C.L., Carlotti, F., de Koning, E.J.P., et al. (2015). KeyGenes, a Tool to Probe Tissue Differentiation Using a Human Fetal Transcriptional Atlas. Stem Cell Reports *4*, 1112–1124, http://dx.doi.org/10.1016/j.stemcr.2015.05.002.

Roybon, L., Lamas, N.J., Garcia-Diaz, A., Yang, E.J., Sattler, R., Jackson-Lewis, V., Kim, Y.A., Kachel, C.A., Rothstein, J.D., Przedborski, S., et al. (2013). Human stem cell-derived spinal cord astrocytes with defined mature or reactive phenotypes. Cell Rep. *4*, 1035–1048, http://dx.doi.org/10.1016/j.celrep.2013.06.021.

Sackmann, E.K., Fulton, A.L., and Beebe, D.J. (2014). The present and future role of microfluidics in biomedical research. Nature *507*, 181–189, http://dx.doi.org/10.1038/nature13118.

Sampaziotis, F., Cardoso de Brito, M., Madrigal, P., Bertero, A., Saeb-Parsy, K., Soares, F.A.C., Schrumpf, E., Melum, E., Karlsen, T.H., Bradley, J.A., et al. (2015). Cholangiocytes derived from human induced pluripotent stem cells for disease modeling and drug validation. Nat. Biotechnol. *33*, 845–852, http://dx.doi.org/10.1038/nbt.3275.

Sareen, D., Gowing, G., Sahabian, A., Staggenborg, K., Paradis, R., Avalos, P., Latter, J., Ornelas, L., Garcia, L., and Svendsen, C.N. (2014). Human induced pluripotent stem cells are a novel source of neural progenitor cells (iNPCs) that migrate and integrate in the rodent spinal cord. J. Comp. Neurol. *522*, 2707–2728, http://dx.doi.org/10.1002/cne.23578.

Sato, T., and Clevers, H. (2013). Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. Science *340*, 1190–1194, http://dx.doi.org/10.1126/science.1234852.

Schwartz, M.P., Hou, Z., Propson, N.E., Zhang, J., Engstrom, C.J., Santos Costa, V., Jiang, P., Nguyen, B.K., Bolin, J.M., Daly, W., et al. (2015). Human pluripotent stem cell-derived neural constructs for predicting neural toxicity. Proc. Natl. Acad. Sci. USA *112*, 12516–12521, http://dx.doi.org/10.1073/pnas.1516645112.

Söderman, M., Andersson, T., Karlsson, B., Wallace, M.C., and Edner, G. (2003). Management of patients with brain arteriovenous malformations. Eur. J. Radiol. *46*, 195–205.

Soong, P.L., Tiburcy, M., and Zimmermann, W.-H. (2012). Cardiac differentiation of human embryonic stem cells and their assembly into engineered heart muscle. Curr. Protoc. Cell Biol. *Chapter 23*, Unit 23.8–23.8.21. http://dx.doi.org/10.1002/0471143030.cb2308s55.

Spellman, T.J., and Gordon, J.A. (2015). Synchrony in schizophrenia: a window into circuit-level pathophysiology. Curr. Opin. Neurobiol. *30*, 17–23, http://dx.doi.org/10.1016/j.conb.2014.08.009.

Spence, J.R., Mayhew, C.N., Rankin, S.A., Kuhar, M.F., Vallance, J.E., Tolle, K., Hoskins, E.E., Kalinichenko, V.V., Wells, S.I., Zorn, A.M., et al. (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. Nature *470*, 105–109, http://dx.doi.org/10.1038/nature09691.

Steinbeck, J.A.., Choi, S.J., Mrejeru, A., Ganat, Y., Deisseroth, K., Sulzer, D., Mosharov, E.V., and Studer, L. (2015). Optogenetics enables functional analysis of human embryonic stem cell-derived grafts in a Parkinson's disease model. Nat. Biotechnol. *33*, 204–209, http://dx.doi.org/10.1038/nbt.3124.

Storkebaum, E., Quaegebeur, A., Vikkula, M., and Carmeliet, P. (2011). Cerebrovascular disorders: molecular insights and therapeutic opportunities. Nat. Neurosci. *14*, 1390–1397, http://dx.doi.org/10.1038/nn.2947.

Studer, L., Vera, E., and Cornacchia, D. (2015). Programming and Reprogramming Cellular Age in the Era of Induced Pluripotency. Cell Stem Cell *16*, 591–600, http://dx.doi.org/10.1016/j.stem.2015.05.004.

Sun, N., Yazawa, M., Liu, J., Han, L., Sanchez-Freire, V., Abilez, O.J., Navarrete, E.G., Hu, S., Wang, L., Lee, A., et al. (2012). Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy. Science Translational Medicine *4*, 130ra47–130ra47. http://dx.doi.org/10.1126/scitranslmed.3003552.

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, http://dx.doi.org/10.1016/j.cell.2007.11.019.

Takasato, M., Er, P.X., Chiu, H.S., Maier, B., Baillie, G.J., Ferguson, C., Parton, R.G., Wolvetang, E.J., Roost, M.S., Chuva de Sousa Lopes, S.M., and Little, M.H. (2015). Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. Nature *526*, 564–568, http://dx.doi.org/10.1038/nature15695.

Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.-R., Ueno, Y., Zheng, Y.-W., Koike, N., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature *499*, 481–484, http://dx.doi.org/10.1038/nature12271.

Takebe, T., Enomura, M., Yoshizawa, E., Kimura, M., Koike, H., Ueno, Y., Matsuzaki, T., Yamazaki, T., Toyohara, T., Osafune, K., et al. (2015). Vascularized and Complex Organ Buds from Diverse Tissues via Mesenchymal Cell-Driven Condensation. Cell Stem Cell *16*, 556–565, http://dx.doi.org/10.1016/j.stem.2015.03.004.

Theodoris, C.V., Li, M., White, M.P., Liu, L., He, D., Pollard, K.S., Bruneau, B.G., and Srivastava, D. (2015). Human disease modeling reveals integrated transcriptional and epigenetic mechanisms of NOTCH1 haploinsufficiency. Cell *160*, 1072–1086, http://dx.doi.org/10.1016/j.cell.2015.02.035.

Tulloch, N.L., Muskheli, V., Razumova, M.V., Korte, F.S., Regnier, M., Hauch, K.D., Pabon, L., Reinecke, H., and Murry, C.E. (2011). Growth of engineered human myocardium with mechanical loading and vascular coculture. Circ. Res. *109*, 47–59, http://dx.doi.org/10.1161/CIRCRESAHA.110.237206.

van der Meer, A.D., Orlova, V.V., ten Dijke, P., van den Berg, A., and Mummery, C.L. (2013). Three-dimensional co-cultures of human endothelial cells and embryonic stem cell-derived pericytes inside a microfluidic device. Lab Chip *13*, 3562–3568, http://dx.doi.org/10.1039/c3lc50435b.

Verghese, P.B., Castellano, J.M., and Holtzman, D.M. (2011). Apolipoprotein E in Alzheimer's disease and other neurological disorders. Lancet Neurol. *10*, 241–252, http://dx.doi.org/10.1016/S1474-4422(10)70325-2.

Wainger, B.J., Kiskinis, E., Mellin, C., Wiskow, O., Han, S.S.W., Sandoe, J., Perez, N.P., Williams, L.A., Lee, S., Boulting, G., et al. (2014). Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. Cell Rep. *7*, 1–11, http://dx.doi.org/10.1016/j.celrep.2014.03.019.

Wang, G., McCain, M.L., Yang, L., He, A., Pasqualini, F.S., Agarwal, A., Yuan, H., Jiang, D., Zhang, D., Zangi, L., et al. (2014). Modeling the mitochondrial cardiomyopathy of Barth syndrome with induced pluripotent stem cell and heart-on-chip technologies. Nat. Med. *20*, 616–623, http://dx.doi.org/10.1038/nm.3545.

Wang, X., Phan, D.T.T., Sobrino, A., George, S.C., Hughes, C.C.W., and Lee, A.P. (2016). Engineering anastomosis between living capillary networks and endothelial cell-lined microfluidic channels. Lab Chip *16*, 282–290, http://dx.doi.org/10.1039/C5LC01050K.

Wanjare, M., Agarwal, N., and Gerecht, S. (2015). Biomechanical strain induces elastin and collagen production in human pluripotent stem cell derived vascular smooth muscle cells. Am. J. Physiol. Cell Physiol. *309*, C271–C281, http://dx.doi.org/10.1152/ajpcell.00366.2014.

Westein, E., van der Meer, A.D., Kuijpers, M.J.E., Frimat, J.-P., van den Berg, A., and Heemskerk, J.W.M. (2013). Atherosclerotic geometries exacerbate pathological thrombus formation poststenosis in a von Willebrand factor-dependent manner. Proc. Natl. Acad. Sci. USA *110*, 1357–1362, http://dx.doi.org/10.1073/pnas.1209905110.

Wilmer, M.J., Ng, C.P., Lanz, H.L., Vulto, P., Suter-Dick, L., and Masereeuw, R. (2016). Kidney-on-a-Chip Technology for Drug-Induced Nephrotoxicity Screening. Trends Biotechnol. *0*, http://dx.doi.org/10.1016/j.tibtech.2015.11.001.

Witty, A.D., Mihic, A., Tam, R.Y., Fisher, S.A., Mikryukov, A., Shoichet, M.S., Li, R.-K., Kattman, S.J., and Keller, G. (2014). Generation of the epicardial lineage from human pluripotent stem cells. Nat. Biotechnol. *32*, 1026–1035, http://dx.doi.org/10.1038/nbt.3002.

Wu, H., Lee, J., Vincent, L.G., Wang, Q., Gu, M., Lan, F., Churko, J.M., Sallam, K.I., Matsa, E., Sharma, A., et al. (2015). Epigenetic Regulation of Phosphodiesterases 2A and 3A Underlies Compromised β-Adrenergic Signaling in an iPSC Model of Dilated Cardiomyopathy. Cell Stem Cell *17*, 89–100, http://dx.doi.org/10.1016/j.stem.2015.04.020.

Yamamoto, Y., Craggs, L., Baumann, M., Kalimo, H., and Kalaria, R.N. (2011). Review: molecular genetics and pathology of hereditary small vessel diseases of the brain. Neuropathol. Appl. Neurobiol. *37*, 94–113, http://dx.doi.org/10.1111/j.1365-2990.2010.01147.x.

Ye, L., Chang, Y.-H., Xiong, Q., Zhang, P., Zhang, L., Somasundaram, P., Lepley, M., Swingen, C., Su, L., Wendel, J.S., et al. (2014). Cardiac repair in a porcine model of acute myocardial infarction with human induced pluripotent stem cell-derived cardiovascular cells. Cell Stem Cell *15*, 750–761, http://dx.doi.org/10.1016/j.stem.2014.11.009.

Yeon, J.H., Na, D., Choi, K., Ryu, S.-W., Choi, C., and Park, J.-K. (2012). Reliable permeability assay system in a microfluidic device mimicking cerebral vasculatures. Biomed. Microdevices *14*, 1141–1148, http://dx.doi.org/10.1007/s10544-012-9680-5.

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, http://dx.doi.org/10.1126/science.1151526.

Zervantonakis, I.K., Kothapalli, C.R., Chung, S., Sudo, R., and Kamm, R.D. (2011). Microfluidic devices for studying heterotypic cell-cell interactions and tissue specimen cultures under controlled microenvironments. Biomicrofluidics *5*, 13406, http://dx.doi.org/10.1063/1.3553237.

Zervantonakis, I.K., Hughes-Alford, S.K., Charest, J.L., Condeelis, J.S., Gertler, F.B., and Kamm, R.D. (2012). Three-dimensional microfluidic model for tumor cell intravasation and endothelial barrier function. Proc. Natl. Acad. Sci. USA *109*, 13515–13520, http://dx.doi.org/10.1073/pnas.1210182109.

Zhang, Q., Jiang, J., Han, P., Yuan, Q., Zhang, J., Zhang, X., Xu, Y., Cao, H., Meng, Q., Chen, L., et al. (2011). Direct differentiation of atrial and ventricular myocytes from human embryonic stem cells by alternating retinoid signals. Cell Res. *21*, 579–587, http://dx.doi.org/10.1038/cr.2010.163.

Zhang, H., Xue, C., Shah, R., Bermingham, K., Hinkle, C.C., Li, W., Rodrigues, A., Tabita-Martinez, J., Millar, J.S., Cuchel, M., et al. (2015). Functional analysis and transcriptomic profiling of iPSC-derived macrophages and their application in modeling Mendelian disease. Circ. Res. *117*, 17–28, http://dx.doi.org/10.1161/CIRCRESAHA.117.305860.

Zhao, Z., Nelson, A.R., Betsholtz, C., and Zlokovic, B.V. (2015). Establishment and Dysfunction of the Blood-Brain Barrier. Cell *163*, 1064–1078, http://dx.doi.org/10.1016/j.cell.2015.10.067.

Zheng, Y., Chen, J., Craven, M., Choi, N.W., Totorica, S., Diaz-Santana, A., Kermani, P., Hempstead, B., Fischbach-Teschl, C., López, J.A., and Stroock, A.D. (2012). In vitro microvessels for the study of angiogenesis and thrombosis. Proc. Natl. Acad. Sci. USA *109*, 9342–9347, http://dx.doi.org/10.1073/pnas.1201240109.

Zhu, H., Lensch, M.W., Cahan, P., and Daley, G.Q. (2011). Investigating monogenic and complex diseases with pluripotent stem cells. Nat. Rev. Genet. *12*, 266–275, http://dx.doi.org/10.1038/nrg2951.

Zimmermann, W.H., Fink, C., Kralisch, D., Remmers, U., Weil, J., and Eschenhagen, T. (2000). Three-dimensional engineered heart tissue from neonatal rat cardiac myocytes. Biotechnol. Bioeng. *68*, 106–114.

CellPress

Review

Transcriptional Control of Somatic Cell Reprogramming

Yan Xu,[1,2,5] Meng Zhang,[1,2,3,5] Wenjuan Li,[1,2,3] Xihua Zhu,[1,2] Xichen Bao,[1,2] Baoming Qin,[2,4] Andrew P. Hutchins,[2,*] and Miguel A. Esteban[1,2,*]

Somatic cells and pluripotent cells display remarkable differences in most aspects of cell function. Accordingly, somatic cell reprogramming by exogenous factors requires comprehensive changes in gene transcription to induce a forced pluripotent state, which is encompassed by a simultaneous transformation of the epigenome. Nevertheless, how the reprogramming factors and other endogenous regulators coordinate to suppress the somatic cell gene program and activate the pluripotency gene network, and why the conversion is multi-phased and lengthy, remain enigmatic. We summarize the current knowledge of transcriptional regulation in somatic cell reprogramming, and highlight new perspectives that may help to reshape existing paradigms.

Resetting Gene Expression on the Road to Induced Pluripotency

The reprogramming of somatic cells to induced pluripotent stem cells (iPSCs) by exogenous factors, including the original cocktail devised by Takahashi and Yamanaka (OCT4, SOX2, KLF4, and c-MYC: OSKM), demonstrates the power of transcription factors to determine cell fate [1]. Nevertheless, our understanding of reprogramming mechanisms remains incomplete, raising concerns regarding potential applications [2]. Substantial progress has been made in defining the functional events necessary for reprogramming [3]: (i) a proliferation burst coupled to the bypass of apoptosis and cell senescence [4], (ii) loss of somatic cell characteristics coupled to a process of organelle remodeling and epithelialization (or mesenchymal-to-epithelial transition, MET) [5,6], and (iii) a metabolic shift that facilitates quicker – albeit less efficient – ATP production [7,8] (Box 1). These phenomena are linked to the progressive appearance of stem cell markers [e.g., alkaline phosphatase activity and surface antigens including SSEA-1 (for mouse), SSEA-3 and TRA-1-60 (for human reprogramming in primed conditions)], and finally the complete activation of the core pluripotency network [9–11]. However, while the reduction of somatic cell markers (e.g., *Thy1*, *S100a4*, and collagen family members in mouse fibroblasts) happens rapidly and in the majority of reprogramming cells, the acquisition of stem cell markers takes place in a significantly smaller fraction. Consequently, the final number of cells fully activating the pluripotency network is limited and most cells in the original population become trapped in an early intermediate cell state [11,12] or in a partially reprogrammed state termed the pre-iPSC state [13]. Gene expression studies of bulk populations with DNA microarrays have confirmed these findings in the mouse system, contributing to establishing the dogma that reprogramming is phased and stochastic, and hence inefficient [6,14,15]. In addition, analysis of specific cell intermediates in mouse reprogramming has defined two major waves of gene transcription separated by a period of relative quiescence [12] (Box 1), suggesting that the multi-phased nature of reprogramming is partly determined by the inability of OSKM to rearrange global gene transcription simultaneously. The first transcriptional wave of mouse reprogramming is induced in almost every cell in culture, and ensures the activation of a proliferative and cell remodeling response that is concomitant with the reduction of somatic gene expression. Conversely, the

Trends

Somatic cells confront massive barriers on the way to induced pluripotency, which impairs reprogramming efficiency and may induce abnormalities.

OSK act as pioneer transcription factors, interacting with distal regions in closed chromatin. This recruits coactivators and corepressors, inducing successive rounds of chromatin remodeling that make reprogramming permissive. Conversely, the role of c-MYC in reprogramming is complex and poorly understood.

Higher-order chromatin architecture is reorganized in reprogramming. This requires the reassembly of enhancer/promoter loops, interactions within and between topologically associated chromatin domains (TADs) and reorganization of lamina-associated domains (LADs).

Transcriptional pause release of pluripotency genes is rate-limiting for reprogramming. Gene body elongation is also regulated in reprogramming.

[1]Laboratory of RNA, Chromatin, and Human Disease, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China
[2]Key Laboratory of Regenerative Biology of the Chinese Academy of Sciences and Guangdong Provincial Key Laboratory of Stem Cells and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China
[3]University of Chinese Academy of Sciences, Beijing 100049, China

 http://dx.doi.org/10.1016/j.tcb.2015.12.003
© 2015 Elsevier Ltd. All rights reserved.

CrossMark

second wave takes place in SSEA-1^{+} cells and is enriched for pluripotency regulators [12]. Analysis of gene expression in human reprogramming using immortalized secondary fibroblasts has also shown distinct waves of gene transcription, but there are differences in the number and nature of these waves compared with the mouse system [11] (Box 1). Understanding how the exogenous factors control gene activation/repression in the different transcriptional waves of reprogramming may clarify why the above-mentioned checkpoints need to exist at all, and consequently why the entire process takes so long (typically ~20 and ~30 days in mouse and human systems, respectively) compared to somatic cell nuclear transfer and cell fusion [16].

Initial Accessibility of OSKM to Target Sites

At the onset of reprogramming, exogenous OSKM face millions of potential binding sites among the myriad bases of the genome. However, much of this DNA, in particular those sites bound by pluripotency transcription factors in embryonic stem cells (ESCs), is packaged into nucleosomes and occluded by repressor complexes, non-permissive histone modifications, and higher-order chromatin structures [17]. To overcome these barriers, OSK (but not c-MYC) act as pioneer factors: transcription factors that bind to closed chromatin – but not to refractory heterochromatin – and then progressively endow competence for the activation of cell type-specific programs [18]. For this purpose, in the first 48 h of reprogramming OSK interact with distal genomic regions located in DNase I-resistant chromatin that lacks evident pre-existing histone marks [19] (Figure 1A, Key Figure). To access this non-permissive chromatin, OSK co-bind to degenerate DNA recognition motifs partially exposed on nucleosomes [20]. This suggests, paradoxically, that there is no obvious initial target preselection and it is instead the random engagement of the exogenous factors with chromatin that initiates the cascade of reprogramming. The high frequency of OSK co-bound DNA regions genome-wide, as opposed to regions bound by only one or two exogenous factors, also supports the idea that transcription factor cooperativity at target sites is important for reprogramming in the same

Box 1. Key Cellular Events during Reprogramming

To achieve pluripotency, reprogramming cells must traverse through multiple roadblocks/checkpoints, including the apoptosis and cell senescence barrier, the MET, a metabolic switch, acquisition of early pluripotency genes, and finally the activation of the full pluripotency gene network (Figure I). In mouse, traversal across these roadblocks is mediated by two major transcriptional waves, named early and late waves [12]. Human reprogramming also shows distinct waves of gene transcription and, although many of the same mouse roadblocks are present, the order and timing are substantially altered [11] (Figure I).

Reprogramming is a stressful process that triggers the production of reactive oxygen species and a DNA damage response, leading to apoptosis via c-MYC-dependent activation of p53 and BAX [114], as well as OCT4-dependent CASPASE-3/8 activation [115], in the early phase of mouse/human OSKM reprogramming. The cell cycle is also crucial because quiescent cells are refractory to reprogramming [116], while rapidly cycling cells are capable of highly-efficient reprogramming [111]. Consequently, inhibition of the *Ink4/Arf/Cdkn2a* locus by ablation of JMJD3 (KDM6B) or over-expression of JHDM1A/B overcomes cell senescence and reprogramming efficiency is improved [24,94,117]. Importantly, the cell cycle of ESCs is different compared to somatic cells [118] and, by imposing ESC-specific cell cycle features on somatic cells (e.g., overexpressing cyclin D1), human iPSCs can be produced more efficiently [119].

Fibroblasts are mesenchymal and ESCs are epithelial-like, and as cells reprogram a near-uniform MET occurs [5,6]. OSKM are directly involved in the MET: OCT4, SOX2, and c-MYC inhibit TGF-β signaling and therefore repress SNAIL, the master transcription factor regulator of the epithelial-to-mesenchymal transition [120], while KLF4 directly activates epithelial genes including *Cdh1* [5]. Of interest, in human OSKM reprogramming the MET occurs much later than in mouse, and this may partly explain why reprogramming human cells is so lengthy [11].

Reprogramming involves a metabolic switch from a mitochondrial-based oxidative, to a glycolytic, metabolic program, a process seemingly independent of exogenous c-MYC [8]. This metabolic switch achieves the goal of quicker energy production but also aims to suppress reactive oxygen species generation by mitochondria because these are mostly detrimental for reprogramming [121]. Consequently, genes and chemical compounds that enhance glycolysis promote human reprogramming, such as the transcription factor hypoxia-inducible factor 1/2α (HIF1/2α) [122,123], although paradoxically in the late phase HIF2α is inhibitory because it induces the pro-apoptotic gene *TRAIL* [123].

[4]Laboratory of Metabolism and Cell Fate, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China
[5]These authors contributed equally

*Correspondence: andrew@gibh.ac.cn (A.P. Hutchins) and miguel@gibh.ac.cn (M.A. Esteban).

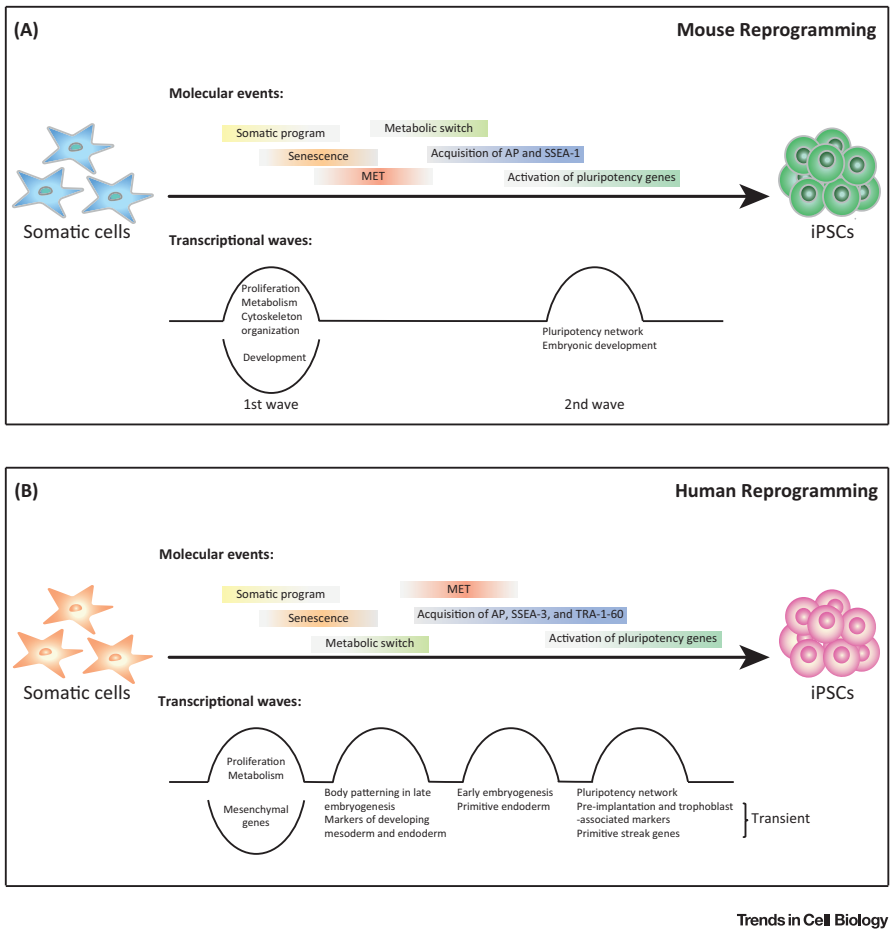


Figure I. Schematic Representation of Reprogramming. Abbreviation: AP, alkaline phosphatase.

way that it is relevant for ESC function [21]. However, this may not be absolutely necessary because, for example, TGF-β inhibitors can substitute for SOX2 (and c-MYC) without significantly affecting reprogramming efficiency [22]. After the initial engagement of OSKM with the genome, it is thought that OSK recruit chromatin modifiers to turn on a 'spring' mechanism that exposes additional binding sites (ultimately at pluripotency gene enhancers and promoters) in successive rounds of epigenetic remodeling (Figure 1B). Consequently, somatic cell chromatin becomes progressively and comprehensively remodeled to mimic the pluripotent state over the course of several weeks of reprogramming, eventually allowing the complete suppression of the somatic cell program and the stable reestablishment of the pluripotency gene network. In support of this model, genetic (Table 1) or chemical [22] manipulation of chromatin regulators greatly enhances reprogramming efficiency and/or accelerates the kinetics. For example, vitamin C improves reprogramming at least in part by facilitating the function of H3K36me2/3 histone demethylases JHDM1A/1B (KDM2A/B) [23,24], in agreement with the extensive histone modifications required for cell fate conversion and vitamin C being a cofactor for dioxygenases [25]. Of note, the role of OSK as pioneer transcription factors has only been tested in human reprogramming but it is expected to be shared by other species including mouse. Similarly, it remains unclear whether this function of OSK is necessary after the first 48 h of reprogramming.

Key Figure

Exogenous Transcription Factor Binding and Chromatin Remodeling During Reprogramming

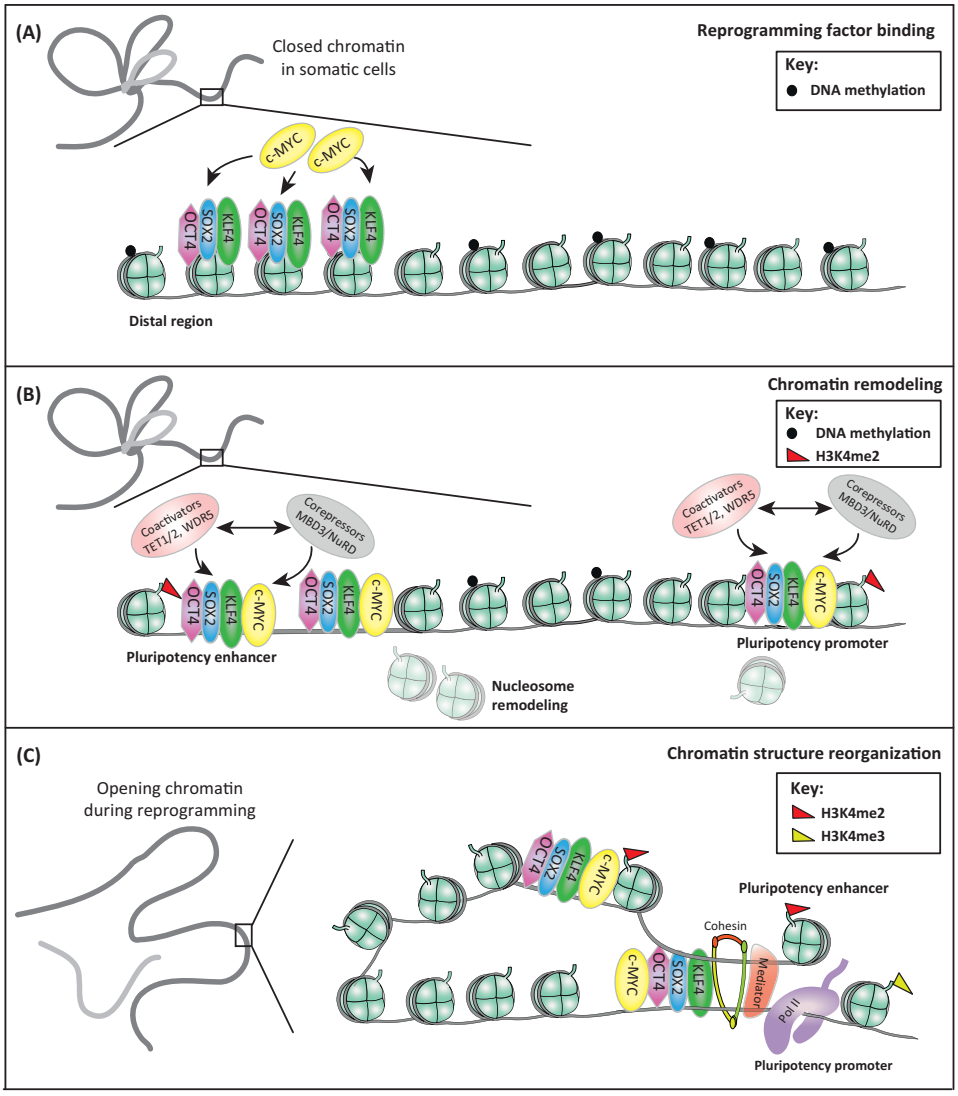


Figure 1. (A) In reprogramming, OSKM binding first takes places in distal regions. For this purpose, OSK act as pioneer transcription factors binding to DNA packed around nucleosomes, and c-MYC facilitates their binding. (B) OSKM bound to chromatin recruit both coactivators and corepressors to initiate chromatin remodeling. (C) Progressive chromatin reorganization during reprogramming induces the acquisition of an open ESC-like chromatin state at pluripotency loci. Pluripotency enhancers juxtapose over their corresponding promoters to initiate gene expression in the late phase of reprogramming, and this needs the recruitment of the SMC1 subunit of the cohesin complex by KLF4. OSK also help to recruit components of the mediator complex to these loci.

Table 1. Coregulators Involved in Somatic Cell Reprogramming[a]

Name	Function		Phenotype	Species	Refs
Chromatin Remodelers					
BRG1 and BAF155	Components of SWI/SNF complex	Enhance OCT4 binding to pluripotency promoters	OE +	M	[124]
BRM and BAF170		Their inhibition improves reprogramming by facilitating the activation of the pluripotency circuitry	KD +	M	[125]
CHD1L	Together with PARP1 remodels chromatin at pluripotency loci		OE +/KD –	M	[126]
CHD4	Components of NuRD complex	–	KD +	M	[49]
MBD2		Blocks reprogramming through direct binding to *NANOG* promoter elements and preventing transcriptional activation	KD +/OE –	H	[127]
		Participates in a positive feedback loop with OCT4 and SFRS2	OE +	H	[104]
MBD3		Establishes heterochromatic features and silences pluripotency genes; recruited by the reprogramming factors, potently restrains the reactivation of OSKM downstream target genes	OE –/KD and KO +	M/H	[49,51,52]
		Facilitates reprogramming in certain contexts	KD and KO –/OE +	M/H	[53]
INO80	Core component of INO80 complex; co-occupies pluripotency gene promoters with pluripotency transcription factors, maintains an open chromatin architecture and licenses recruitment of the mediator complex and Pol II		KD –	M	[128]
Histone Modifiers					
BMI1	H2AK119 ubiquitylation enzymes		KD –/OE +	M/H	[93,129]
RING1			KD –	H	[93]
WDR5	Core component of the TrxG complex mediating H3K4me3; interacts with OCT4 and shares overlapping gene regulatory functions with OCT4		KD –	M	[48]
LSD1	H3K4/K9 demethylase; its inhibition promotes the MET and pluripotency gene activation		IN and KD +	H	[11,130]
EHMT1 and SETDB1	H3K9 methyltransferases	–	KD –	H	[93]
EHMT1/2 and SETDB1		Load H3K9 methylation on pluripotency loci	KD and IN +	M	[19,49, 92,131, 132]
SUV39H1/2			KD +	M/H	[19,93, 131]
JMJD1A/1B	H3K9 demethylases	Remove H3K9 methylation on pluripotency loci	KD –	M/H	[92,131, 133,134, 117]
JMJD2B			KD –/OE +		
JMJD2C			KD –		
EED and SUZ12	H3K27 methyltransferases	–	KD –	H	[93,135]
			KD +	M	[49]
EZH1			KD +	H	[11]
EZH2		Silences the TGF-β signaling pathway and negatively regulates the expression of pro-mesenchymal microRNAs	KD and KO –/OE +	M/H	[36,46, 93,117]
JARID1B			KD +	M	[136]

Table 1. (continued)

Name	Function		Phenotype	Species	Refs
	H3K4 demethylases	Its inhibition downregulates mesenchymal master regulators			
JARID1B/1D		–	KD –	M	[117]
JMJD3	H3K27 demethylases	Targets PHF20 for ubiquitination and degradation via recruitment of TRIM26	OE –/KD and KO +	M/H	[117]
UTX		Interacts with OSK to activate potent pluripotency-promoting gene modules	KO and KD –	M/H	[117,135]
JHDM1A/1B	H3K36 demethylases; accelerate cell cycle progression and suppress cell senescence, also activate the microRNA cluster 302–367		OE +/ KD –	M	[24,94, 117,137]
DOT1L	H3K79 methyltransferase; its inhibition facilitates loss of H3K79me2 from genes that are fated to be repressed in the pluripotent state		KD and KO +	M/H	[93]
PRMT5*	Protein arginine methyltransferase 5; facilitates reprogramming via downregulation of p53		OE +	G	[138]
GCN5	Histone acetyltransferases	Forms a positive feed-forward loop with c-MYC, activating a distinct alternative splicing network and the early acquisition of pluripotency-associated splicing events	KD –	M/H	[106]
HDAC2	Histone deacetylase; facilitates TET1 binding and DNA demethylation at the promoters of iPSC maturation-related genes		KD +	M	[139]
SIRT1*	Deacetylases	Regulates reprogramming through deacetylating SOX2 and p53	KO –/OE +	M	[140,141]
SIRT6*		Regulates microRNA-766 transcription via a feedback regulatory loop	OE +	H	[142]
DNA Modifiers					
DNMT1	DNA methyltransferase; its inhibition promotes DNA demethylation in the late phase of reprogramming		KD and IN +	M/H	[15,49]
AID	Cytidine deaminase	Contributes to active DNA demethylation in reprogramming	KD and KO –/OE +	M/H	[143–145]
TET1	DNA demethylases	Promotes *Oct4* demethylation and reactivation	KD –/OE +	M	[62,133, 146–148]
		Regulates 5hmC formation at loci critical for MET in a Vc-dependent fashion	Vc –/no Vc +	M	[149]
TET2		Contributes together with PARP1 to an epigenetic program that directs subsequent transcriptional induction at pluripotency loci	KD –	M/H	[57,133, 134]
TET1/2		Synergize with NANOG to enhance reprogramming	KD –/OE +	M	[148]
TET1/2/3		Mediate demethylation and activation of microRNAs essential for the MET transition in reprogramming	KO –	M	[146]
Other Epigenetic Regulators					
ASF1A	Histone chaperones	Affects the expression of core pluripotency genes	KD –/OE +	H	[150]
CHAF1A and CHAF1B		Components of chromatin assembly factor-1 (CAF-1) complex; optimized suppression leads to a more accessible chromatin structure at pluripotency enhancers early during reprogramming	KD +	M	[55]

Table 1. (continued)

Name	Function		Phenotype	Species	Refs
BRD4	H4K5/8/12/16ac and H3K27ac reader; induces pause release of pluripotency genes		KD −/OE +	M/H	[74]
CBX3	H3K9 methylation reader; associates with the transcription initiation complex in a cell type- and activator-dependent manner		KD +	M	[92]
MacroH2A	Histone variants	Prevents the regain of H3K4me2 on pluripotency genes	KD and KO +/OE −	M/H	[151–153]
TH2A and TH2B		Increase the DNase I sensitivity of chromatin	KO −/OE +	M	[154]
RCOR2	Component of the LSD1 complex; facilitates nucleosomal demethylation activity of LSD1		KD −/OE +	M/H	[155]
SIN3A	Component of sin3A/HDAC complex		KD +	M	[49]

[a]Symbols and abbreviations: +, enhance; −, inhibit; 5hmC, 5-hydroxymethylcytosine; G, goat; H, human; IN, inhibitor; KD, knockdown; KO, knockout; M, mouse; OE, overexpression; Vc, vitamin C; *, their targets are not histones during reprogramming.

The Atypical Role of c-MYC in OSKM Reprogramming

A relevant issue brought to the spotlight by the 'pioneer factor' model of OSKM reprogramming is the specific role of exogenous c-MYC. Originally, exogenous c-MYC was thought to act by promoting survival and boosting cell proliferation in the early phase of OSKM reprogramming [3]. This mechanism helps to overcome the senescence barrier in reprogramming cells, while simultaneously increasing the chance of probabilistic events that promote chromatin remodeling in the desired direction [26,27]. However, overexpressing c-MYC together with OSK increases OSK binding to nucleosomal DNA in the first 48 h of human reprogramming [19] (Figure 1A). c-MYC interacts with partially degenerate E-boxes near OSK clusters bound to DNA [20], indicating that c-MYC lacks pioneer transcription factor activity on its own. This binding pattern is different from mouse ESCs, where c-MYC mostly locates at promoters separately from OSK [19], further complicating how OSK binding to DNA is enhanced in OSKM reprogramming. One interesting possibility is that c-MYC recruits some of its well-known partners including SWI/SNF nucleosome remodelers, histone demethylases/methyltransferases, and histone acetyltransferases to OSKM clusters [28]. These chromatin modifiers could subsequently reduce nucleosome compaction, allowing access of OSKM to additional target sites and/or stabilizing OSKM binding to DNA. Opposing this idea is, however, the general assumption that recruitment of chromatin modifiers is a known important feature of pioneer transcription factors [18], and therefore one would expect that OSK do not need c-MYC to achieve this goal. Yet, OSK might be exposed to different post-translational modifications in reprogramming cells compared to ESCs, in which case OSK function would not be optimal in the early stages of reprogramming and would instead require fine-tuning. Such a mechanism could explain why adding transactivation domains (e.g., VP16 or the MyoD transactivation domain, which are known to induce the recruitment of transcription factors, chromatin remodelers, and histone acetyltransferases) to OCT4 and SOX2 drastically enhances reprogramming [29,30]. Interestingly, c-MYC also interacts with the transcription factor SREBP1, a regulator of lipid metabolism, and this enhances mouse reprogramming efficiency by facilitating OSK binding to pluripotency loci [31]. Together, this suggests that OSKM possibly bind to chromatin in the form of large multiprotein complexes, implying that the incorrect stoichiometry of different coregulators could negatively modulate OSKM function and thus be detrimental for reprogramming.

Implications of the 'Pioneer Factor' Model of OSKM Reprogramming

Because the number of putative OSKM binding sites genome-wide is immense, and likely only a small fraction can repurpose the epigenome towards a pluripotent state, there are several

important implications of the 'pioneer factor' model of reprogramming. First, the exogenous transcription factors need to be expressed at high levels to effectively produce iPSCs, which is supported by the reprogramming of otherwise refractory mouse intermediate cells with additional OSKM overexpression [12]. This possibility may explain why reprogramming using recombinant OSKM proteins is inefficient in both mouse and human [32–34]. However, this could also be caused by the inability of the recombinant factors to induce an innate immune response [35]. Second, the initial and middle phases of reprogramming are largely stochastic at the chromatin reorganization level, and generally also at the transcriptional level, an observation validated by single-cell quantitative RT-PCR analysis of 48 genes in mouse [36]. Third, there are off-ESC-target OSKM binding events that may facilitate the activation of different cell type-specific programs in both mouse and human reprogramming [11,12,15,37,38], and could likewise be the source of low basal reprogramming efficiency and frequent epigenetic abnormalities in iPSCs [2]. In fact, it has been suggested that activation of lineage specifiers in mouse and human reprogramming intermediates underlies the ability of OSKM to directly produce cardiomyocytes, neurons, and other lineages upon changing the reprogramming culture conditions [39]. However, recent studies have argued that direct lineage conversion of mouse somatic cells by pluripotency transcription factors involves transient acquisition of pluripotency [40,41]. Fourth, the 'pioneer factor' model provides a rationale as to why the GATA family of transcription factors, which act as pioneer factors during development [18], can substitute for OCT4 in mouse and human OSKM reprogramming [42,43]. Instead, it has been proposed that the mutual counterbalancing of alternative cell fates by the exogenous transcription factors drives reprogramming cells to a default pluripotent state [42–44]. In this 'see-saw' model of mouse reprogramming, GATA factors can substitute for OCT4 because they all promote an endodermal cell fate, while in opposition SOX2 promotes an ectodermal cell fate and can be replaced by other ectoderm-related transcription factors (e.g., GMNN, SOX1, and SOX3) [42]. Despite these findings, it remains to be clarified why pluripotency is an automatic default state of a cell.

A 'Gas and Brakes' Model of Reprogramming

While OSKM remodel distal chromatin to facilitate access to target sites at pluripotent loci, they also dismantle the somatic cell gene program [3]. This process is initiated by the depletion of somatic cell master transcription factors and regulators (e.g., *Snai1* and *Snai2* in mouse fibroblasts) [5,15], implying that the reprogramming factors not only function as pioneer factors/transactivators but also as repressors. Indeed, comparative analysis of KLF4 targets in ESCs and transcriptomic data in reprogramming cells suggested that KLF4 is mainly a repressor in the first transcriptional wave of reprogramming and an activator in the second wave [12]. The same study predicted that c-MYC is an activator acting primarily in the first wave, while OCT4 and SOX2 are activators in the second wave. Yet, overexpressing either KLF4 or c-MYC, and in particular their combination, strongly represses mouse fibroblast gene expression without the need to undergo full reprogramming [45]. Moreover, analysis of DNA-binding events in the early phase of human reprogramming have shown that KLF4 and c-MYC bind to DNase I hypersensitive regions in active (somatic-like) gene promoters, while OCT4 and SOX2 do not [19]. These experiments suggest that KLF4 and c-MYC coordinately repress the somatic cell gene program in the early phase of reprogramming, but it remains unclear how they discern gene repression from activation. One interesting possibility is that KLF4 and c-MYC recruit different types of coregulators at specific loci. In this regard, c-MYC binds to the H3K27 methyltransferase EZH2 to repress loci encoding TGF-β pathway components in human OSKM reprogramming [46], and it also interacts with WDR5 [core component of the trithorax group (TrxG) complex with H3K4 methyltransferase activity] to activate other loci in mouse reprogramming [47,48]. A relevant corollary of this model is, however, that segregation of corepressors and coactivators in reprogramming may not always be clear-cut; therefore, both types of coregulators could arbitrarily compete for binding to OSKM at target loci (Figure 1B). In this 'gas and brakes'

scenario [49], the right genes would only become properly modulated in reprogramming cells after stochastic events, and the same phenomenon should apply to chromatin remodeling. The progressive change in the expression of OSKM coregulators from a somatic-like to an ESC-like pattern would then induce a more hierarchical and deterministic transition to a pluripotent state as reprogramming advances [36,49,50]. Reinforcing the idea that corepressors are a major 'brake' for reprogramming, optimized depletion of two components of the MBD3/NuRD complex, MBD3 or CHD4, which interact with OSKM, greatly enhances mouse reprogramming efficiency and kinetics [49,51,52]. It has also been proposed that the effect of suppressing MBD3 is context-dependent [53], but this might be related to differences in cell proliferation because complete, but not partial, ablation of MBD3 in somatic cells leads to accelerated cell senescence [49,53].

Opening Pluripotency Enhancers and Promoters

Chromatin remodeling at pluripotency loci is not a late event in reprogramming because the active histone mark H3K4me2 is acquired along with a concomitant reduction in the repressive H3K27me3 at pluripotency enhancers in the first 48 h of mouse OSKM reprogramming [54]. Accessibility to these enhancers can be drastically enhanced by optimized suppression of components of the histone chaperone complex CAF-1 [55], which mediates nucleosome assembly on newly synthesized DNA. Pluripotency promoters are also remodeled during the first days of reprogramming [54], but full opening to allow transcription factor binding and gene activation is generally considered to be a much later event [56]. This discrepancy can be explained by frequent hypermethylation of CpG sites at pluripotency gene promoters in somatic cells, and the observation that, as opposed to histone modifications, the biggest change in global DNA methylation takes place in the late phase of both mouse and human reprogramming [11,12]. Moreover, it has been reported that some pluripotency gene promoters already become partly accessible and allow exogenous transcription factor binding in the early phase of reprogramming [56–59]. Another relevant consideration is that, despite primarily targeting promoters, DNA methylation also occurs at enhancers to control gene expression [60], making it tempting to speculate that, in contrast to promoters, DNA demethylation of at least some pluripotency gene enhancers is an early event in reprogramming. In this regard, active TET-mediated DNA demethylation occurs primarily at enhancers in ESCs [61] and TET1 can substitute for OCT4 in mouse OSKM reprogramming [62], suggesting a causal relationship. Similarly, it has been postulated that TET2-induced cytosine 5-hydroxymethylation of pluripotency loci mediates the 'path-breaking' effect of coexpressing C/EBPα with OSKM in mouse primary B cell reprogramming [63]. Conversely, another study showed that TET2 acts by promoting deposition of 5-hydroxymethylcytosine at pluripotency loci in the early phase of mouse OSKM reprogramming [57], instead of progressing to induce CpG demethylation.

Changes in the 3D Genome Precede the Reactivation of the Pluripotency Network

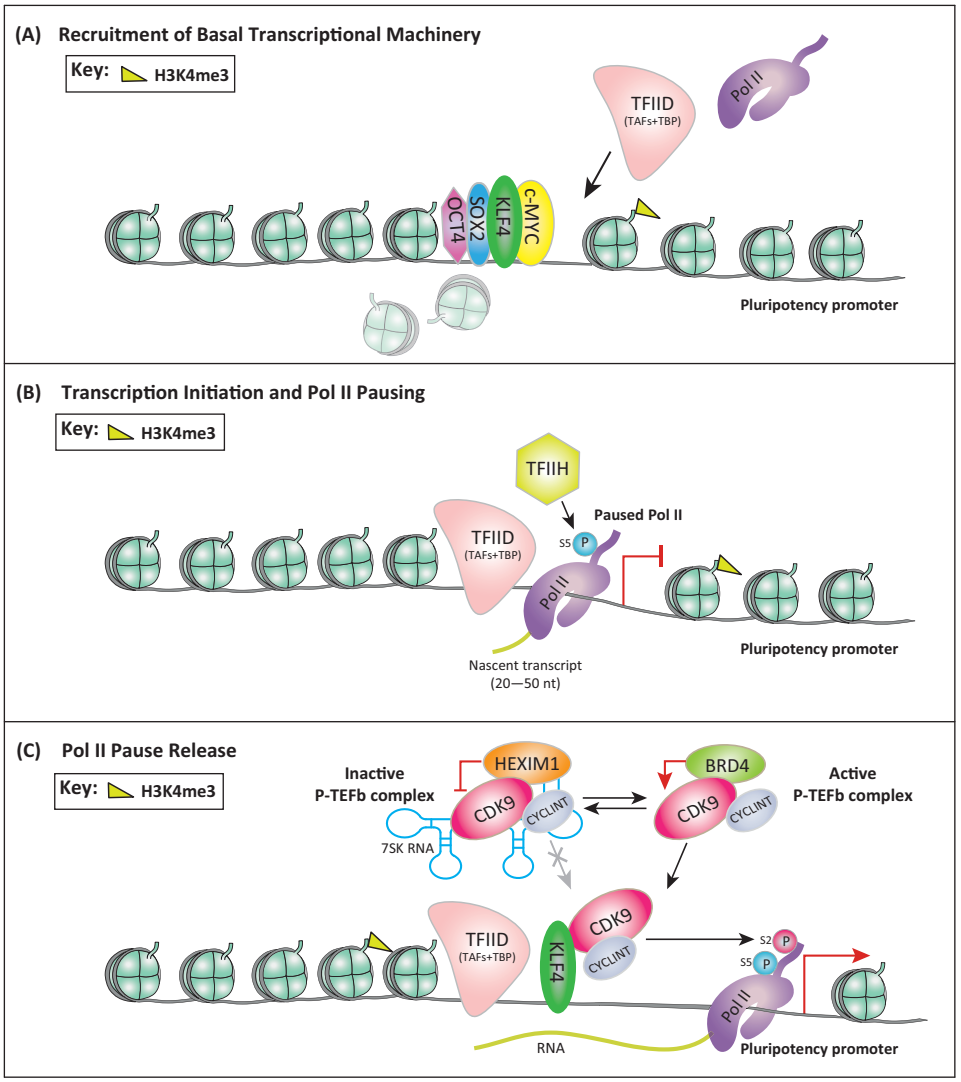
The metazoan genome is three dimensional [17], adding another layer of complexity to reprogramming. At the smallest scale (within hundreds of kb) there are enhancer/promoter interactions (a process often referred to as looping), which are cell type-specific and necessary for activating gene expression. At a larger scale, there are topologically associated chromatin domains (TADs), typically spanning 1 Mb, with boundaries enriched for the insulator protein CTCF. Although TAD boundaries are relatively stable among cell types, and are mostly conserved between human and mouse, long-range interactions within and between TADs are often cell type-specific. In addition to these short- and long-range interactions, specific chromatin regions contact the nuclear periphery to form lamina-associated domains (LADs), which mostly have a repressive function [64]. Although LADs are globally similar between cell types, key genes related to cell identity (e.g., developmental genes in ESCs) that are placed within the LADs become 'loose', and hence activated, as cell fate changes [65].

Logically, enhancer/promoter looping and long-range chromatin interactions change during reprogramming because they are different between fibroblasts and ESCs [58,59,66,67] (Figure 1C). The looping of pluripotency enhancers over their respective promoters (e.g., *Nanog* locus) starts relatively early in mouse reprogramming (in SSEA-1$^+$ cells), long before the respective loci become activated [59,66]. However, this phenomenon does not seem to be a general feature of all pluripotency loci [59,66]. Interestingly, these long-range interacting loci are enriched for exogenous KLF4 and mostly involve actively transcribed genes in mouse ESCs that are heavily enriched for RNA polymerase II (Pol II) [59]. By bridging pluripotency loci with newly-forming transcriptional factories at distant regions, KLF4 may contribute to the coordinated activation of many coregulated pluripotency genes. In addition, KLF4 is necessary for pluripotency enhancer–promoter looping [59] (Figure 1C). These findings correlate with the observation that KLF4 mainly works as an activator in the second transcriptional wave of mouse OSKM reprogramming [12]. This architectural role of KLF4 in reprogramming seems to be independent of its function as a transcription factor, requiring recruitment of the SMC1 subunit of the cohesin complex [59], which is in agreement with the role of other KLF family members in unrelated cell contexts [68]. OSK also interact with components of the mediator complex (e.g., MED1), which further facilitates enhancer/promoter looping and long-range interactions in mouse reprogramming [59,66] (Figure 1C). Mediator and cohesin complexes are necessary for connecting pluripotency enhancers and promoters in mouse ESCs [69]. Consequently, knocking down components of these two complexes, including SMC1 and MED12, impairs mouse/human OSKM reprogramming and ESC pluripotency [58,66,69]. Concomitant with the acquisition of pluripotency-specific modifications in the 3D genome, somatic cell-specific long-range interactions involving pluripotency loci, which presumably help to preserve a somatic cell fate, start being removed in the early phase of mouse reprogramming [66]. Likewise, the expression of the nuclear lamina component lamin A decreases during mouse OSKM reprogramming and its knockdown enhances iPSC generation [70], suggesting that this facilitates the reactivation of ESC-specific genes contained within LADs [71].

Assembly of Pol II Transcription Pre-Initiation Complexes at Pluripotency Promoters

The changes in long-range interactions of the 3D genome along the entire reprogramming process coordinate with parallel changes in histone and DNA modifications to promote successive rounds of OSKM binding events and further chromatin modifications. Ultimately, this ensures access of OSKM to a subset of pluripotency enhancers and promoters that have also experienced looping [58,66]. The subsequent activation of these key endogenous pluripotency regulators provides another boost to reprogramming, likely by promoting a more complete and widespread occupancy of target loci (specifically enhancers). Indeed, single-cell quantitative RT-PCR of mouse OSKM reprogramming has shown that reactivation of *Sox2* initiates the late hierarchical phase of gene expression, followed by *Sall4* and *Lin28* [36], although a different mouse study using DNA microarray analysis concluded that *Nr0b1* and *Etv5* become activated first [52]. Eventually, these events create a chain reaction that establishes the entire pluripotency gene network in the late phase of reprogramming [36], but the resulting mouse iPSCs require a few more divisions to be fully reprogrammed and transcriptionally stable [72]. However, Pol II must be recruited to these loci and adequately activated to allow transcription initiation and elongation before pluripotency genes can be fully reactivated [73].

Notably, hypophosphorylated Pol II is recruited relatively early to many pluripotency loci in mouse OSKM reprogramming [74], raising the question as to whether full promoter opening is required for Pol II binding and how OSKM orchestrate this process. Pol II recruitment is facilitated by the general transcription factor TFIID (composed of the TATA-binding protein and 13 TATA-associated factors or TAFs), resulting in the formation of the transcription pre-initiation complex [75] (Figure 2A). Because TAF subunits are differentially expressed in somatic cells and ESCs,



Trends in Cell Biology

Figure 2. Transcription Initiation and Elongation of Pluripotency Genes in Reprogramming. (A) OSKM factors bind to pluripotency loci and recruit chromatin modifiers to elicit changes in nucleosome distribution, DNA methylation, and histone modifications. Remodeled chromatin at pluripotency promoters allows binding of the TFIID complex and Pol II. (B) Transcription initiation at pluripotency loci needs phosphorylation of serine 5 (S5) on Pol II by CDK7. However, after transcribing 20–50 nt, activated Pol II becomes paused. (C) Pause-release factor CDK9 phosphorylates serine 2 (S2) of Pol II, stimulating Pol II pause release and productive transcriptional elongation. BRD4 helps to promote transcriptional pause release of pluripotency genes by dissociating CDK9 from an inactive complex containing HEXIM1 and 7SK snRNA. KLF4 helps to recruit CDK9 to pluripotency promoters.

overexpressing specific TAFs enriched in ESCs (in particular TAF4, but also TAF5 and TAF6) significantly enhances mouse and human OSKM reprogramming efficiency [76]. This suggests that Pol II recruitment and transcription initiation of pluripotency genes are rate-limiting for reprogramming. However, Pol II inhibition or TAF knockdown in mouse ESCs result in reduced growth and differentiation, respectively, and it has therefore been proposed that ESC-specific TAFs work independently of Pol II to maintain ESC identity and promote reprogramming [76]. Supporting this idea, TAF3 participates in long-range chromatin interactions that maintain the pluripotency transcriptional network in mouse ESCs [77].

Transcriptional Elongation is Rate-Limiting for Reprogramming

After formation of the pre-initiation complex in reprogramming cells, Pol II is not yet competent for transcription initiation and needs to be phosphorylated on serine 5 by CDK7 within the general transcription factor complex, TFIIH [78]. In addition, serine 5-phosphorylated Pol II stops 20–50 nt after initiating the transcription of pluripotency genes in mouse and human reprogramming [74], and in ESCs [73], as a result of interaction with the negative regulators of elongation, NELF and DSIF [78] (Figure 2B). Productive elongation of pluripotency genes then occurs only in some reprogramming cells through a mechanism, termed transcriptional pause release, which requires additional phosphorylation of Pol II on serine 2 by the P-TEFb complex [78]. Of note, Pol II pausing and pause-release are not exclusive to pluripotency genes in reprogramming or ESCs, but represent a general mechanism for regulating gene networks that require rapid activation or repression of high levels of gene expression upon the presence or absence of environmental stimuli, respectively [78]. Hence, it is expected that transcriptional pause release is necessary for reprogramming, but it is surprising that it only affects the late phase [74].

During Pol II pausing, CDK9, the catalytic subunit of P-TEFb complex, is kept inactive as part of a complex containing HEXIM1 and 7SK RNA [79] (Figure 2C). To release paused Pol II bound to pluripotency loci in mouse and human reprogramming cells, CDK9 is activated by BRD4, a bromodomain protein that interacts with acetylated histones in regulatory DNA regions, including enhancers [74,80] (Figure 2C). KLF4 also interacts with CDK9 to recruit it to pluripotency promoters [74], while OCT4 recruits BRD4 to promoters and enhancers [81]. The role of KLF4 in stimulating Pol II pause release in reprogramming contrasts with previous observations that c-MYC binds to CDK9 to promote transcriptional amplification in mouse ESCs [82]. However, the targets of c-MYC in mouse ESCs are mostly related to proliferation and metabolism rather than to the pluripotency gene network [83–85]. Moreover, recent reports have questioned the function of c-MYC in transcriptional amplification and have proposed a role in transcription initiation [86,87]. In agreement with these findings, suppressing CDK9 or BRD4 blocks mouse OSKM reprogramming and produces colonies resembling pre-iPSCs, although inhibiting CDK9 (and to a lesser extent BRD4) exclusively in the early phase of reprogramming paradoxically enhances efficiency [74]. The latter may be related to a putative role of CDK9 and BRD4 in promoting transcriptional pause release of fibroblast-specific networks; however, it could also be related to a negative effect of CDK9 inhibition on the cell cycle [88]. These findings suggest that using CDK9/BRD4 inhibitors to suppress highly specialized cell-specific transcriptional programs [89], many of which are subjected to transcriptional pause release, could enhance the reprogramming of highly specialized cell types (e.g., B cells and cancer cells). Supporting this possibility, BRD4 inhibitors facilitate the generation of neurons from somatic cells using chemicals [90]. Notably, suppressing KLF4 in the late phase of mouse and human OSKM reprogramming also produces pre-iPSCs-like colonies [91], reinforcing the idea that KLF4 regulates the second transcriptional wave of reprogramming at least in part through CDK9. Conversely, overexpressing BRD4 enhances the efficiency of mouse and human reprogramming and accelerates the kinetics [74], further confirming that transcriptional pause release of pluripotency genes is a rate-limiting step for the late phase of reprogramming.

In addition to promoter-proximal pause release, gene body elongation, which ensures the production of full-length RNA transcripts, is also modulated in reprogramming. In this regard, two histone marks associated with transcriptional elongation, H3K36me3 and H3K79me2, are depleted during mouse OSKM reprogramming and are displayed at a relatively low level in fully reprogrammed mouse iPSCs and ESCs [92]. This is consistent with the enhancing effects on reprogramming of overexpressing JHDM1A/1B or knocking down DOT1L (a H3K79 methyltransferase) [24,93,94], suggesting a negative role for gene body transcriptional elongation in reprogramming, as opposed to the positive role of Pol II pause release. However, JHDM1B and

DOT1L mostly regulate the early phase of reprogramming [93,94], while CDK9 and BRD4 are necessary for the late phase [74].

Concluding Remarks

Reprogramming by exogenous factors has been mostly studied using bulk populations or specific reprogramming intermediates with a focus on the role of OSKM, which leaves unclear whether other cocktails of reprogramming factors [36,95,96] share similar mechanisms (Box 2). These studies have laid the foundation for understanding the transcriptional and epigenetic responses that OSKM employ to enforce a pluripotent cell fate. Yet, we are likely looking only at the tip of the iceberg, and a multitude of newly discovered cell regulators, such as novel RNA species (large intergenic non-coding RNAs or lincRNAs, enhancer RNAs, circular RNAs, etc.) [97], as well as other layers of regulation including RNA modifications [98,99] and alternative splicing, may soon bring more complexity to the reprogramming field. For example, two p53-induced lincRNAs modulate reprogramming efficiency: overexpression of *lincRNA-RoR* enhances iPSC generation [100], while overexpression of *lincRNA-p21* does the opposite [101]. Likewise, reprogramming is associated with a reversion of the global alternative-splicing pattern to a pluripotent state, a process driven by splicing factors such as SFRS2 and MBNL1/2 [102–105]. In this regard, GCN5 cooperates with c-MYC to activate a pluripotency-like alternative-splicing network in reprogramming [106]. Another important consideration is whether the inherent heterogeneity of the reprogramming process affects current mechanistic interpretations. Advanced single-cell technologies including RNA-sequencing [107], DNA methylation [108], multiplex profiling of chromatin accessibility [109], and 3D genome studies [110] will help to clarify this issue. However, reprogramming under standard conditions is lengthy and single-cell sequencing procedures remain expensive, limiting the number of profiled cells and the time-points in the analysis, and hence the conclusions. A potential solution could be to study specific cell types highly amenable to reprogramming [111], modify further the culture conditions or the cocktails of exogenous factors [23,29,112], and use optimized secondary reprogramming systems [113] or genetically modified cells, including *Mbd3* hypomorphic cells [49], because these approaches enhance reprogramming efficiency, shorten its length, and likely reduce intercellular variability. A caveat is that these advances have been applied mostly to mouse cells,

Box 2. Alternative Reprogramming Factors

Since the discovery by Yamanaka and Takahashi that OSKM can reprogram somatic cells to pluripotency [1], many researchers have sought alternative factors that can perform the same feat. It was quickly discovered that c-MYC is dispensable [26,27] and that related family members of OSKM could also be swapped. Specifically, OCT4 can be replaced by the related POU homeodomain transcription factor BRN4 [40], SOX2 can be replaced with the SOX family members SOX1, SOX3, SOX15, or SOX18, KLF4 can be replaced with KLF1, KLF2, or KLF5, while c-MYC can be replaced by L-MYC or N-MYC [27]. Interestingly, OSKM members can also be replaced with factors that are essential for pluripotency, but are otherwise unrelated to OSKM. These include the orphan nuclear receptors ESRRB and NR5A1/NR5A2, which can substitute for KLF4 and OCT4, respectively [156,157], GLIS1 that can replace c-MYC [158], and TCL-1A that can replace OCT4 [159]. Perhaps more surprisingly, lineage specifiers that antagonize the pluripotent state are capable of substituting for OSKM members in reprogramming: the GATA family, SOX7, PAX1, CEBPA, HNF4A, and GRB2 can each replace OCT4, while GMNN can replace SOX2 [42,43].

Methods involving not only transcription factors have also been developed and include the RNA-binding protein LIN28A which, when coexpressed with NANOG, OCT4, and SOX2, can replace KLF4 and c-MYC [96]. The epigenetic modifier and corepressor RCOR2 can replace SOX2 [155], and the DNA demethylase TET1, or the epithelial junction molecule E-CADHERIN, can both replace OCT4 [62,160], while one of JHDM1B or BMI1 in combination with OCT4 can reprogram mouse cells [24,129]. OCT4 fused to the VP16 domain can also reprogram somatic cells alone [29]. These lines of research have culminated in two completely non-OSKM reprogramming cocktails, the first consisting of combinations of ESRRB, SALL4, LIN28A, DPPA2, and NANOG [36], and the second consisting of JDP2, JHDM1B, ID1/ID3, GLIS1, SALL4, and LRH1 [95]. In addition, suppressing DOT1L or knocking down *p53* can replace c-MYC and KLF4 [93,161], and the knockdown of *Dlx3* can replace OCT4 [42]. Similarly, it has been proposed that the microRNA cluster 302–367 and a cocktail of microRNA-200c, microRNA-302s, and microRNA-369s can entirely replace OSKM [162–164]. Finally, small-molecule chemical compounds alone can induce reprogramming in the absence of exogenous factors [165,166].

Outstanding Questions

Why are some transcription factors capable of reprogramming somatic cells to pluripotency? What are the key biochemical criteria that make a transcription factor competent for reprogramming?

Why do the OSKM transcription factors first need to engage binding sites unrelated to their ESC-binding sites to promote chromatin reorganization during reprogramming? Is there a preference in OSKM binding-site target selection at the onset of reprogramming that can lead to increased reprogramming efficiency? If so, can it be controlled experimentally?

What domains of OSK confer pioneer transcription factor activity in reprogramming and how?

What is the precise nature and order of epigenetic remodeling events in reprogramming, and in what way do OSKM drive these events or combine with them? Do variations in the culture conditions (including the use of chemical inhibitors) or differences in the stoichiometry of the reprogramming factors influence this?

Is the sequence of epigenetic remodeling in reprogramming different when using alternative cocktails of transcription factors, regulators, or chemicals? Are there mechanistic differences in reprogramming between different cell sources (e.g., fibroblasts versus epithelial cells)? What are the major interspecies (e.g., mouse versus human) differences in reprogramming, and what causes them?

How does the silencing of somatic genes in reprogramming take place at the epigenetic level? Can pluripotency genes be activated without proper somatic gene silencing?

Why is the late phase of reprogramming hierarchical under standard conditions? What is the role of the earliest-induced endogenous pluripotency transcription factors in pushing cells towards a fully reprogrammed state?

How do the cellular checkpoints of reprogramming impact on the transcriptional and epigenetic events of reprogramming? For example: why is the MET phase important for acquiring a pluripotent state? In addition, does the metabolic shift of reprogramming

and reprogramming human cells (particularly primary cells) remains so far substantially more challenging. Nevertheless, these upcoming high-resolution maps of mouse and human reprogramming may provide answers to many unresolved questions in the field (see Outstanding Questions). These future findings could help in devising protocols that activate the two transcriptional waves of reprogramming simultaneously, potentially allowing the generation of both mouse and human iPSCs without discernible phases and with the same speed as somatic cell nuclear transfer or cell fusion.

Acknowledgments

Work in the Laboratory of RNA, Chromatin, and Human Disease is funded by the Ministry of Science and Technology of China 973 program (2011CB965201), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA01020106), the National Natural Science Foundation of China (31371513, 31461143011, and 81261160506), the Bureau of Science, Technology and Information of Guanghzou Municipality (201508030027), the International Science and Technology Cooperation Program of China (2013DFE33080), and the Science and Technology Planning Project of Guangdong Province, China (2013B050800010). A.P.H. is funded by the National Natural Science Foundation of China (31471242 and 31550110206) and by postdoctoral grants (2014M552250 and 2015T80923).

References

1. 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
2. Liang, G. and Zhang, Y. (2013) Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell* 13, 149–159
3. Apostolou, E. and Hochedlinger, K. (2013) Chromatin dynamics during cellular reprogramming. *Nature* 502, 462–471
4. Banito, A. and Gil, J. (2010) Induced pluripotent stem cells and senescence: learning the biology to improve the technology. *EMBO Rep.* 11, 353–359
5. Li, R. *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
6. Samavarchi-Tehrani, P. *et al.* (2010) Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 7, 64–77
7. Pfeiffer, T. *et al.* (2001) Cooperation and competition in the evolution of ATP-producing pathways. *Science* 292, 504–507
8. Zhang, J. *et al.* (2012) Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. *Cell Stem Cell* 11, 589–595
9. Brambrink, T. *et al.* (2008) Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* 2, 151–159
10. Stadtfeld, M. *et al.* (2008) Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2, 230–240
11. Cacchiarelli, D. *et al.* (2015) Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell* 162, 412–424
12. Polo, J.M. *et al.* (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151, 1617–1632
13. Silva, J. *et al.* (2008) Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* 6, e253
14. Golipour, A. *et al.* (2012) A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell* 11, 769–782
15. Mikkelsen, T.S. *et al.* (2008) Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454, 49–55
16. Gurdon, J.B. and Wilmut, I. (2011) Nuclear transfer to eggs and oocytes. *Cold Spring Harb. Perspect. Biol.* 3, a002659
17. Gorkin, D.U. *et al.* (2014) The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* 14, 762–775
18. Zaret, K.S. and Carroll, J.S. (2011) Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 25, 2227–2241
19. Soufi, A. *et al.* (2012) Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* 151, 994–1004
20. Soufi, A. *et al.* (2015) Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* 161, 555–568
21. Aksoy, I. *et al.* (2013) Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *EMBO J.* 32, 938–953
22. Li, W. *et al.* (2013) Chemical approaches to stem cell biology and therapeutics. *Cell Stem Cell* 13, 270–283
23. Esteban, M.A. *et al.* (2010) Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* 6, 71–79
24. Wang, T. *et al.* (2011) The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. *Cell Stem Cell* 9, 575–587
25. Esteban, M.A. and Pei, D. (2012) Vitamin C improves the quality of somatic cell reprogramming. *Nat. Genet.* 44, 366–367
26. Wernig, M. *et al.* (2008) c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2, 10–12
27. Nakagawa, M. *et al.* (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* 26, 101–106
28. Cheng, S.W. *et al.* (1999) c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. *Nat. Genet.* 22, 102–105
29. Wang, Y. *et al.* (2011) Reprogramming of mouse and human somatic cells by high-performance engineered factors. *EMBO Rep.* 12, 373–378
30. Hirai, H. *et al.* (2011) Radical acceleration of nuclear reprogramming by chromatin remodeling with the transactivation domain of MyoD. *Stem Cells* 29, 1349–1361
31. Wu, Y. *et al.* (2015) Srebp-1 interacts with c-Myc to enhance somatic cell reprogramming. *Stem Cells* Published online September 21, 2015. http://dx.doi.org/10.1002/stem.2209
32. Kim, D. *et al.* (2009) Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* 4, 472–476
33. Zhou, H. *et al.* (2009) Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* 4, 381–384
34. Cho, H.J. *et al.* (2010) Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. *Blood* 116, 386–395
35. Lee, J. *et al.* (2012) Activation of innate immunity is required for efficient nuclear reprogramming. *Cell* 151, 547–558

generate specific metabolites that can impact on the cellular epigenetic state?

Is ultra-high (near 100%) efficiency reprogramming (e.g., using MBD3/NuRD-depleted cells) deterministic at both the transcriptional and epigenetic levels? Are there any other pathways to deterministic reprogramming?

Are the changes in the 3D genome aiming to mimic the ESC state completely accomplished in reprogramming?

Can changes in the 3D genome driven by individual-specific genetic variation determine epigenetic differences among human iPSCs?

Can alternative pluripotency states in addition to epiblast-like cells [167] and F-cells [168] be achieved through reprogramming? If so, what are the mechanisms?

Can reprogramming be achieved in 24–48 h and without cell division?

36. Buganim, Y. *et al.* (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150, 1209–1222
37. O'Malley, J. *et al.* (2013) High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature* 499, 88–91
38. Zhuang, Q. *et al.* (2013) Class IIa histone deacetylases and myocyte enhancer factor 2 proteins regulate the mesenchymal-to-epithelial transition of somatic cell reprogramming. *J. Biol. Chem.* 288, 12022–12031
39. Zhu, S. *et al.* (2015) Reprogramming fibroblasts toward cardiomyocytes, neural stem cells and hepatocytes by cell activation and signaling-directed lineage conversion. *Nat. Protoc.* 10, 959–973
40. Bar-Nur, O. *et al.* (2015) Lineage conversion induced by pluripotency factors involves transient passage through an iPSC stage. *Nat. Biotechnol.* 33, 761–768
41. Maza, I. *et al.* (2015) Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. *Nat. Biotechnol.* 33, 769–774
42. Shu, J. *et al.* (2013) Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* 153, 963–975
43. Montserrat, N. *et al.* (2013) Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* 13, 341–350
44. Loh, K.M. and Lim, B. (2011) A precarious balance: pluripotency factors as lineage specifiers. *Cell Stem Cell* 8, 363–369
45. Sridharan, R. *et al.* (2009) Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 136, 364–377
46. Rao, R.A. *et al.* (2015) Ezh2 mediated H3K27me3 activity facilitates somatic transition during human pluripotent reprogramming. *Sci. Rep.* 5, 8229
47. Thomas, L.R. *et al.* (2015) Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC. *Mol. Cell* 58, 440–452
48. Ang, Y.S. *et al.* (2011) Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* 145, 183–197
49. Rais, Y. *et al.* (2013) Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* 502, 65–70
50. Buganim, Y. *et al.* (2013) Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet.* 14, 427–439
51. Luo, M. *et al.* (2013) NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. *Stem Cells* 31, 1278–1286
52. Lujan, E. *et al.* (2015) Early reprogramming regulators identified by prospective isolation and mass cytometry. *Nature* 521, 352–356
53. dos Santos, R.L. *et al.* (2014) MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. *Cell Stem Cell* 15, 102–110
54. Koche, R.P. *et al.* (2011) Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* 8, 96–105
55. Cheloufi, S. *et al.* (2015) The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* 528, 218–224
56. Taberlay, P.C. *et al.* (2011) Polycomb-repressed genes have permissive enhancers that initiate reprogramming. *Cell* 147, 1283–1294
57. Doege, C.A. *et al.* (2012) Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2. *Nature* 488, 652–655
58. Zhang, H. *et al.* (2013) Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell Stem Cell* 13, 30–35
59. Wei, Z. *et al.* (2013) Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. *Cell Stem Cell* 13, 36–47
60. Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13, 484–492
61. Baumann, K. (2014) Epigenetics: enhancers under TET control. *Nat. Rev. Mol. Cell Biol.* 15, 699
62. Gao, 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
63. Di Stefano, B. *et al.* (2014) C/EBPalpha poises B cells for rapid reprogramming into induced pluripotent stem cells. *Nature* 506, 235–239
64. Luperchio, T.R. *et al.* (2014) Genome regulation at the peripheral zone: lamina associated domains in development and disease. *Curr. Opin. Genet. Dev.* 25, 50–61
65. Peric-Hupkes, D. *et al.* (2010) Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* 38, 603–613
66. Apostolou, E. *et al.* (2013) Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell* 12, 699–712
67. Denholtz, M. *et al.* (2013) Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* 13, 602–616
68. Schoenfelder, S. *et al.* (2010) Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat. Genet.* 42, 53–61
69. Kagey, M.H. *et al.* (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435
70. Zuo, B. *et al.* (2012) Influences of lamin A levels on induction of pluripotent stem cells. *Biol. Open* 1, 1118–1127
71. Li, M. *et al.* (2012) Navigating the epigenetic landscape of pluripotent stem cells. *Nat. Rev. Mol. Cell Biol.* 13, 524–535
72. Polo, J.M. *et al.* (2010) Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* 28, 848–855
73. Min, I.M. *et al.* (2011) Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.* 25, 742–754
74. Liu, L. *et al.* (2014) Transcriptional pause release is a rate-limiting step for somatic cell reprogramming. *Cell Stem Cell* 15, 574–588
75. Goodrich, J.A. and Tjian, R. (2010) Unexpected roles for core promoter recognition factors in cell-type-specific transcription and gene regulation. *Nat. Rev. Genet.* 11, 549–558
76. Pijnappel, W.W. *et al.* (2013) A central role for TFIID in the pluripotent transcription circuitry. *Nature* 495, 516–519
77. Liu, Z. *et al.* (2011) Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* 146, 720–731
78. Fuda, N.J. *et al.* (2009) Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186–192
79. Yik, J.H. *et al.* (2003) Inhibition of P-TEFb (CDK9/cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol. Cell* 12, 971–982
80. Wu, S.Y. and Chiang, C.M. (2007) The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J. Biol. Chem.* 282, 13141–13145
81. Wu, T. *et al.* (2015) The BET family member BRD4 interacts with OCT4 and regulates pluripotency gene expression. *Stem Cell Rep.* 4, 390–403
82. Rahl, P.B. *et al.* (2010) c-Myc regulates transcriptional pause release. *Cell* 141, 432–445
83. Chen, X. *et al.* (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117
84. Kim, J. *et al.* (2010) A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* 143, 313–324
85. Kim, J. *et al.* (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132, 1049–1061
86. Walz, S. *et al.* (2014) Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. *Nature* 511, 483–487
87. Sabo, A. *et al.* (2014) Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature* 511, 488–492

88. Xu, Y. *et al.* (2013) Proliferation rate of somatic cells affects reprogramming efficiency. *J. Biol. Chem.* 288, 9767–9778
89. Whyte, Warren A. *et al.* (2013) Master transcription factors and Mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319
90. Li, X. *et al.* (2015) Small-molecule-driven direct reprogramming of mouse fibroblasts into functional neurons. *Cell Stem Cell* 17, 195–203
91. Nishimura, K. *et al.* (2014) Manipulation of KLF4 expression generates iPSCs paused at successive stages of reprogramming. *Stem Cell Rep.* 3, 915–929
92. Sridharan, R. *et al.* (2013) Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency. *Nat. Cell Biol.* 15, 872–882
93. Onder, T.T. *et al.* (2012) Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 483, 598–602
94. Liang, G. *et al.* (2012) Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. *Nat. Cell Biol.* 14, 457–466
95. Liu, J. *et al.* (2015) The oncogene c-Jun impedes somatic cell reprogramming. *Nat. Cell Biol.* 17, 856–867
96. Yu, J. *et al.* (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920
97. Fatica, A. and Bozzoni, I. (2014) Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21
98. Chen, T. *et al.* (2015) m(6)A RNA methylation is regulated by microRNAs and promotes reprogramming to pluripotency. *Cell Stem Cell* 16, 289–301
99. Aguilo, F. *et al.* (2015) Coordination of mA mRNA methylation and gene transcription by ZFP217 regulates pluripotency and reprogramming. *Cell Stem Cell* 17, 689–704
100. Loewer, S. *et al.* (2010) Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* 42, 1113–1117
101. Bao, X. *et al.* (2015) The p53-induced lincRNA-p21 derails somatic cell reprogramming by sustaining H3K9me3 and CpG methylation at pluripotency gene promoters. *Cell Res.* 25, 80–92
102. Gabut, M. *et al.* (2011) An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell* 147, 132–146
103. Ohta, S. *et al.* (2013) Global splicing pattern reversion during somatic cell reprogramming. *Cell Rep.* 5, 357–366
104. Lu, Y. *et al.* (2014) Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. *Cell Stem Cell* 15, 92–101
105. Han, H. *et al.* (2013) MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* 498, 241–245
106. Hirsch, C.L. *et al.* (2015) Myc and SAGA rewire an alternative splicing network during early somatic cell reprogramming. *Genes Dev.* 29, 803–816
107. Kim, D.H. *et al.* (2015) Single-cell transcriptome analysis reveals dynamic changes in lncRNA expression during reprogramming. *Cell Stem Cell* 16, 88–101
108. Smallwood, S.A. *et al.* (2014) Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* 11, 817–820
109. Cusanovich, D.A. *et al.* (2015) Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 348, 910–914
110. Nagano, T. *et al.* (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59–64
111. Guo, S. *et al.* (2014) Nonstochastic reprogramming from a privileged somatic cell state. *Cell* 156, 649–662
112. Chen, J. *et al.* (2010) Towards an optimized culture medium for the generation of mouse induced pluripotent stem cells. *J. Biol. Chem.* 285, 31066–31072
113. Woltjen, K. *et al.* (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 458, 766–770
114. Banito, A. *et al.* (2009) Senescence impairs successful reprogramming to pluripotent stem cells. *Genes Dev.* 23, 2134–2139
115. Li, F. *et al.* (2010) Apoptotic caspases regulate induction of iPSCs from human fibroblasts. *Cell Stem Cell* 7, 508–520
116. Ruiz, S. *et al.* (2011) A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity. *Curr. Biol.* 21, 45–52
117. Zhao, W. *et al.* (2013) Jmjd3 inhibits reprogramming by upregulating expression of INK4a/Arf and targeting PHF20 for ubiquitination. *Cell* 152, 1037–1050
118. Burdon, T. *et al.* (2002) Signalling, cell cycle and pluripotency in embryonic stem cells. *Trends Cell Biol.* 12, 432–438
119. Edel, M.J. *et al.* (2010) Rem2 GTPase maintains survival of human embryonic stem cells as well as enhancing reprogramming by regulating p53 and cyclin D1. *Genes Dev.* 24, 561–573
120. Ocana, O.H. and Nieto, M.A. (2010) Epithelial plasticity, stemness and pluripotency. *Cell Res.* 20, 1086–1088
121. Ji, J. *et al.* (2014) Antioxidant supplementation reduces genomic aberrations in human induced pluripotent stem cells. *Stem Cell Rep.* 2, 44–51
122. Prigione, A. *et al.* (2014) HIF1alpha modulates cell fate reprogramming through early glycolytic shift and upregulation of PDK1-3 and PKM2. *Stem Cells* 32, 364–376
123. Mathieu, J. *et al.* (2014) Hypoxia-inducible factors have distinct and stage-specific roles during reprogramming of human cells to pluripotency. *Cell Stem Cell* 14, 592–605
124. Singhal, N. *et al.* (2010) Chromatin-remodeling components of the BAF complex facilitate reprogramming. *Cell* 141, 943–955
125. Jiang, Z. *et al.* (2015) Knockdown of Brm and Baf170, components of chromatin remodeling complex, facilitates reprogramming of somatic cells. *Stem Cells Dev.* 24, 2328–2336
126. Jiang, B.H. *et al.* (2015) CHD1L regulated PARP1-driven pluripotency and chromatin remodeling during the early-stage cell reprogramming. *Stem Cells* 33, 2961–2972
127. Lee, M.R. *et al.* (2013) Epigenetic regulation of NANOG by miR-302 cluster-MBD2 completes induced pluripotent stem cell reprogramming. *Stem Cells* 31, 666–681
128. Wang, L. *et al.* (2014) INO80 facilitates pluripotency gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. *Cell Stem Cell* 14, 575–591
129. Moon, J.H. *et al.* (2011) Reprogramming fibroblasts into induced pluripotent stem cells with Bmi1. *Cell Res.* 21, 1305–1315
130. Li, W. *et al.* (2009) Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. *Stem Cells* 27, 2992–3000
131. Chen, J. *et al.* (2013) H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* 45, 34–42
132. Shi, Y. *et al.* (2008) A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2, 525–528
133. Tran, K.A. *et al.* (2015) Collaborative rewiring of the pluripotency network by chromatin and signalling modulating pathways. *Nat. Commun.* 6, 6188
134. Qin, H. *et al.* (2014) Systematic identification of barriers to human iPSC generation. *Cell* 158, 449–461
135. Mansour, A.A. *et al.* (2012) The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* 488, 409–413
136. Kidder, B.L. *et al.* (2013) Extended self-renewal and accelerated reprogramming in the absence of Kdm5b. *Mol. Cell. Biol.* 33, 4793–4810
137. Vidal, S.E. *et al.* (2014) Combinatorial modulation of signaling pathways reveals cell-type-specific requirements for highly efficient and synchronous iPSC reprogramming. *Stem Cell Rep.* 3, 574–584
138. Chu, Z. *et al.* (2015) PRMT5 enhances generation of induced pluripotent stem cells from dairy goat embryonic fibroblasts via down-regulation of p53. *Cell Prolif.* 48, 29–38
139. Wei, T. *et al.* (2015) An HDAC2–TET1 switch at distinct chromatin regions significantly promotes the maturation of pre-iPS to iPS cells. *Nucleic Acids Res.* 43, 5409–5422

140. Mu, W.L. *et al.* (2015) Sox2 deacetylation by Sirt1 is involved in mouse somatic reprogramming. *Stem Cells* 33, 2135–2147
141. Lee, Y.L. *et al.* (2012) Sirtuin 1 facilitates generation of induced pluripotent stem cells from mouse embryonic fibroblasts through the miR-34a and p53 pathways. *PloS ONE* 7, e45633
142. Sharma, A. *et al.* (2013) The role of SIRT6 protein in aging and reprogramming of human induced pluripotent stem cells. *J. Biol. Chem.* 288, 18439–18447
143. Bhutani, N. *et al.* (2013) A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells. *FASEB J.* 27, 1107–1113
144. Bhutani, N. *et al.* (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* 463, 1042–1047
145. Kumar, R. *et al.* (2013) AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* 500, 89–92
146. Hu, X. *et al.* (2014) Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell* 14, 512–522
147. Chen, J. *et al.* (2015) The combination of Tet1 with Oct4 generates high-quality mouse-induced pluripotent stem cells. *Stem Cells* 33, 686–698
148. Costa, Y. *et al.* (2013) NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature* 495, 370–374
149. Chen, J. *et al.* (2013) Vitamin C modulates TET1 function during somatic cell reprogramming. *Nat. Genet.* 45, 1504–1509
150. Gonzalez-Munoz, E. *et al.* (2014) Histone chaperone ASF1A is required for maintenance of pluripotency and cellular reprogramming. *Science* 345, 822–825
151. Pasque, V. *et al.* (2012) Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. *J. Cell Sci.* 125, 6094–6104
152. Gaspar-Maia, A. *et al.* (2013) MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. *Nat. Commun.* 4, 1565
153. Barrero, M.J. *et al.* (2013) Macrohistone variants preserve cell identity by preventing the gain of H3K4me2 during reprogramming to pluripotency. *Cell Rep.* 3, 1005–1011
154. Shinagawa, T. *et al.* (2014) Histone variants enriched in oocytes enhance reprogramming to induced pluripotent stem cells. *Cell Stem Cell* 14, 217–227
155. Yang, P. *et al.* (2011) RCOR2 is a subunit of the LSD1 complex that regulates ESC property and substitutes for SOX2 in reprogramming somatic cells to pluripotency. *Stem Cells* 29, 791–801
156. Feng, B. *et al.* (2009) Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat. Cell Biol.* 11, 197–203
157. Heng, J.C. *et al.* (2010) The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* 6, 167–174
158. Maekawa, M. *et al.* (2011) Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* 474, 225–229
159. Picanco-Castro, V. *et al.* (2011) Pluripotent reprogramming of fibroblasts by lentiviral mediated insertion of SOX2, C-MYC, and TCL-1A. *Stem Cells Dev.* 20, 169–180
160. Redmer, T. *et al.* (2011) E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO Rep* 12, 720–726
161. Kawamura, T. *et al.* (2009) Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* 460, 1140–1144
162. Bao, X. *et al.* (2013) MicroRNAs in somatic cell reprogramming. *Curr. Opin. Cell Biol.* 25, 208–214
163. Anokye-Danso, F. *et al.* (2011) Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* 8, 376–388
164. Miyoshi, N. *et al.* (2011) Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell* 8, 633–638
165. Hou, P. *et al.* (2013) Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 341, 651–654
166. Long, Y. *et al.* (2015) Bromodeoxyuridine promotes full-chemical induction of mouse pluripotent stem cells. *Cell Res.* 25, 1171–1174
167. Han, D.W. *et al.* (2011) Direct reprogramming of fibroblasts into epiblast stem cells. *Nat. Cell Biol.* 13, 66–71
168. Tonge, P.D. *et al.* (2014) Divergent reprogramming routes lead to alternative stem-cell states. *Nature* 516, 192–197

Give your research a boost with alerts from Cell Press. You'll be glad you did.

Get first access with immediate, regular electronic table of contents (eToCs) alerts delivered directly to your desktop, free of charge, that keep you informed of breakthroughs in your field.

Exciting extra features like video abstracts, podcasts, and blog posts give you additional depth and context and you can access news and commentary, normally not accessible online, in advance of publication.

Register today
Visit cell.com/alerts

CellPress

Cell

Article

A XEN-like State Bridges Somatic Cells to Pluripotency during Chemical Reprogramming

Yang Zhao,[1,5,*] Ting Zhao,[1,2,5] Jingyang Guan,[1,5] Xu Zhang,[1,5] Yao Fu,[1,5] Junqing Ye,[1,3,5] Jialiang Zhu,[1] Gaofan Meng,[1,3] Jian Ge,[1] Susu Yang,[1] Lin Cheng,[1] Yaqin Du,[1] Chaoran Zhao,[1] Ting Wang,[1] Linlin Su,[1] Weifeng Yang,[4] and Hongkui Deng[1,3,*]

[1]Department of Cell Biology, School of Basic Medical Sciences, Peking University Stem Cell Research Center, Center for Molecular and Translational Medicine, State Key Laboratory of Natural and Biomimetic Drugs, Peking University Health Science Center, and the MOE Key Laboratory of Cell Proliferation and Differentiation, College of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100191, China

[2]Peking University-Tsinghua University-National Institute of Biological Sciences Joint Graduate Program, College of Life Sciences, Peking University, Beijing 100871, China

[3]Shenzhen Stem Cell Engineering Laboratory, Key Laboratory of Chemical Genomics, Peking University Shenzhen Graduate School, Shenzhen 518055, China

[4]Beijing Vitalstar Biotechnology Co., Ltd., Beijing 100012, China

[5]Co-first author

*Correspondence: yangzhao@pku.edu.cn (Y.Z.), hongkui_deng@pku.edu.cn (H.D.)

http://dx.doi.org/10.1016/j.cell.2015.11.017

SUMMARY

Somatic cells can be reprogrammed into pluripotent stem cells (PSCs) by using pure chemicals, providing a different paradigm to study somatic reprogramming. However, the cell fate dynamics and molecular events that occur during the chemical reprogramming process remain unclear. We now show that the chemical reprogramming process requires the early formation of extra-embryonic endoderm (XEN)-like cells and a late transition from XEN-like cells to chemically-induced (Ci)PSCs, a unique route that fundamentally differs from the pathway of transcription factor-induced reprogramming. Moreover, precise manipulation of the cell fate transition in a step-wise manner through the XEN-like state allows us to identify small-molecule boosters and establish a robust chemical reprogramming system with a yield up to 1,000-fold greater than that of the previously reported protocol. These findings demonstrate that chemical reprogramming is a promising approach to manipulate cell fates.

INTRODUCTION

Pluripotent stem cells can be induced from somatic cells by nuclear transfer into oocytes, transgene delivery, or treatment with chemical compounds (Gurdon, 1962; Hou et al., 2013; Takahashi and Yamanaka, 2006). Chemically induced reprogramming was first demonstrated by our group in 2013, when we showed that pluripotent stem cells can be generated from mouse somatic cells using a cocktail of seven small molecules (Hou et al., 2013). Chemical reprogramming provides a fundamentally new paradigm for studying pluripotency and cell fate reprogramming (Chou and Cheng, 2013; Hou et al., 2013). In addition, this chemical strategy shows promise in cell fate manipulation because small molecules can be cell permeable and easy to manipulate, they do not integrate in to chromosomes, and their effect is reversible. Therefore, chemically induced pluripotent stem cells may have many advantages in cell therapy, disease modeling, and drug discovery (Hou et al., 2013; Li et al., 2013).

To date, the molecule roadmap of the reprogramming process has been extensively studied in experiments that induce pluripotency using the transcription factors *Oct4*, *Sox2*, *Klf4*, and *c-Myc* (OSKM). As previously reported, the changes in gene expression and epigenetic states are primarily induced by two major waves during reprogramming (Hansson et al., 2012; Polo et al., 2012). The mesenchymal-to-epithelium transition (MET) mediates an early stage of reprogramming that is induced by OSKM (Li et al., 2010; Samavarchi-Tehrani et al., 2010), whereas *Sox2* initiates a deterministic, later stage of reprogramming (Buganim et al., 2012). Moreover, a primitive streak-like state has been reported as an intermediate state during the cell fate transition from somatic cells to pluripotent cells (Takahashi et al., 2014).

Because chemical reprogramming has been established more recently, comparatively little is known about the reprogramming process that is induced by small molecules. In particular, the small-molecule cocktails that are used in chemical reprogramming have been identified by phenotypic screening and do not involve the direct activation of classical reprogramming factors (Hou et al., 2013), such as *Oct4*, *Sox2*, *Klf4*, and *c-Myc*. Therefore, it is intriguing to speculate about the similarities and differences between OSKM-induced and chemically induced reprogramming. In addition, a major advantage of the chemical approach is that small molecules can be fine-tuned in terms of their concentrations, durations, structures, and combinations, providing the opportunity to manipulate the chemically induced pluripotent stem cells (CiPSC) generation process more precisely during each stage of

CrossMark

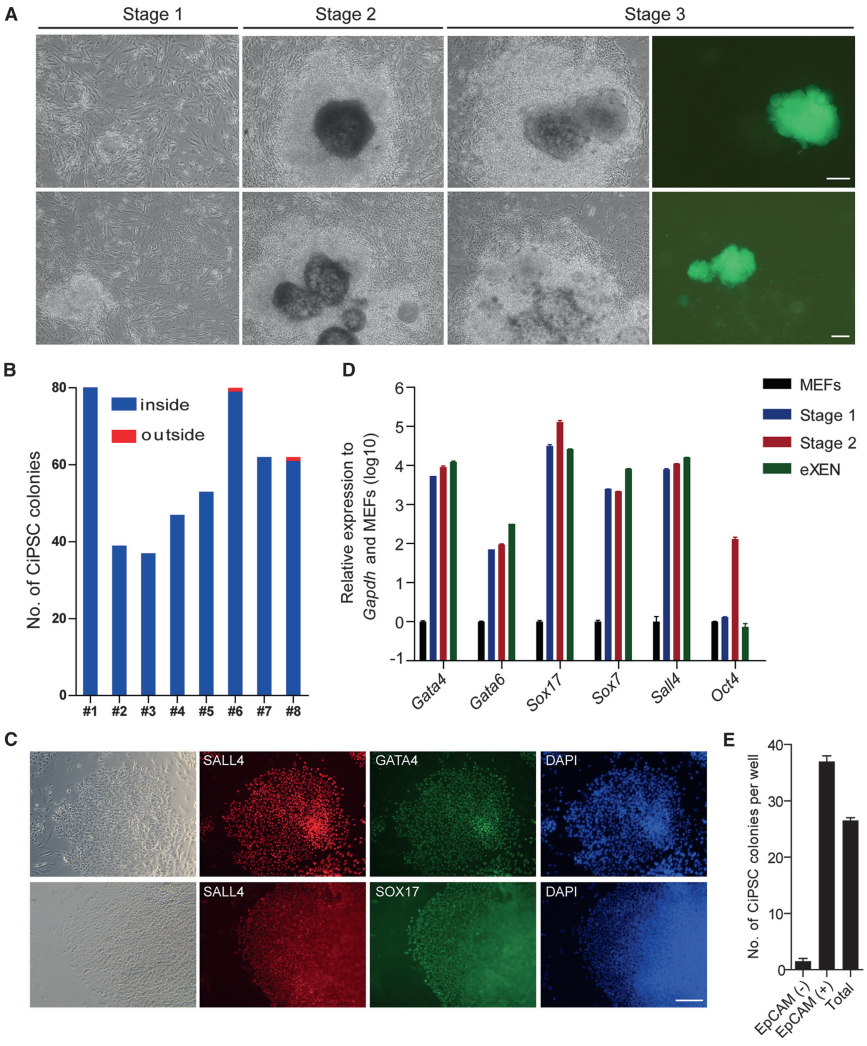


Figure 1. The Induction of Cell Colonies Expressing XEN Cell Markers Is a Cornerstone Event during Chemical Reprogramming

(A) Imaging tracing of the formation process of two CiPSC colonies (stage 3) in stage 1 and stage 2 during chemical reprogramming. Scale bars, 100 μm.

(B) Numbers of CiPSC colonies generated from the inside (blue) and outside (red) of epithelial colonies in eight batches of experiments. For experiments 5 and 7, the cell confluence of epithelial colonies was <20%.

(C) Immunofluorescence of typical epithelial colonies at the end of stage 1, with the expression of SALL4 (top, red) and GATA4 (top, green), SALL4 (bottom, red) and SOX17 (bottom, green) 4 days after replating at day 12. Scale bar, 100 μm.

(D) qRT-PCR analysis of XEN cell markers (*Gata4*, *Gata6*, *Sox17*, *Sox7*, *Sall4*) and pluripotency marker *Oct4* in MEFs, cells at the end of stage 1 (day 16) and stage 2 (day28) and eXEN (embryo-derived XEN cells).

(E) CiPSC colony numbers generated at the end of stage 3 from EpCAM-negative (–), EpCAM-positive (+) and total cell populations sorted at day 20 (in stage 2 of chemical reprogramming).

Data are represented as mean ± SD.

See also Figure S1.

reprogramming. Therefore, the identification of key molecular events and intermediate cell states would help greatly improve the efficiency of reprogramming by allowing for the fine-tuning of small molecules and culture conditions based on the markers of each stage of reprogramming.

In this study, we identified an extraembryonic endoderm (XEN)-like state as an intermediate during the early stage of chemical reprogramming. Moreover, the chemical reprogramming efficiency is greatly enhanced by small-molecule boosters and the more precise optimization of the reprogramming conditions for each step identified with the XEN-like state as an indicator.

RESULT

The Induction of Cell Colonies Expressing XEN Cell Markers Is a Cornerstone Event during Chemical Reprogramming

In our previously developed method for inducing pluripotent stem cells from fibroblasts, there are three essential stages in the chemical reprogramming process. During these stages, a cocktail of five small molecules, "VC6TF" (VPA, CHIR99021, 616452, tranylcypromine, and forskolin), was used in stage 1 for 16–20 days; another small molecule, DZNep, was added at the start of stage 2 for the next 20–24 days; and 2i-medium (with dual inhibition of ERK and GSK3 with PD0325901 and CHIR99021, respectively) was used in stage 3 for the last 12–16 days. In total, the chemical reprogramming process can take as long as 48–60 days (Hou et al., 2013).

To dissect the chemical reprogramming process, we carefully followed the change in cell morphology during chemical reprogramming in each stage. Notably, we found that at the end of stage 1, a number of epithelial colonies formed, and these epithelial cells rapidly expanded during stage 2. More importantly, by tracing the dynamic changes in cell fate during chemical reprogramming, we found that CiPSCs predominantly emerged from the inside of these epithelial cell colonies (Figures 1A, 1B, and S1A). In some experiments, 100% of the CiPSCs were generated from these colonies, even when the cells were re-plated at a lower density; the epithelial cell colonies had grown to <20% confluence. These findings indicate that CiPSCs may be generated from these epithelial cells.

To better understand these epithelial colonies as the potential intermediates of chemical reprogramming, we next used immunofluorescence and qRT-PCR to examine the gene expression pattern of these epithelial cell colonies. By immunofluorescence, we found that all epithelial colonies formed in the end of stage 1

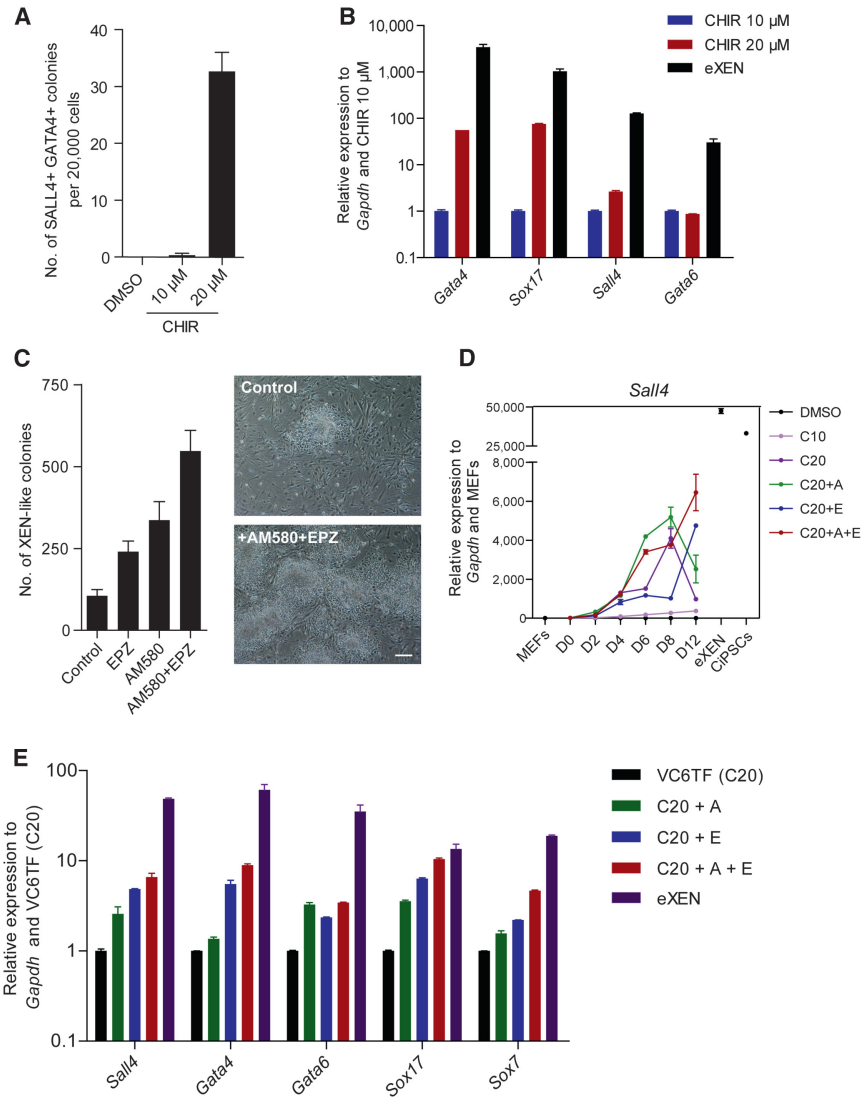


Figure 2. Identification of Small Molecules that Promote the Transition from Fibroblasts to the XEN-like Cells

(A) Numbers of SALL4 and GATA4 double-positive colonies after treatment with DMSO and VC6TF (CHIR, 10 μM and 20 μM) for 12 days.

(B) qRT-PCR analysis of XEN cell marker expression after treatment with VC6TF (CHIR, 10 μM and 20 μM) for 12 days. eXEN was set as a positive control.

(C) Numbers and phase images of XEN-like colonies after treatment with control cocktail (VC6TF with CHIR, 20 μM) and that with additional small molecule EPZ004777 (E) and AM580 (A) for 16 days. Cells were re-plated at day 12 by 1:2. Scale bar, 100 μm.

(D) qRT-PCR analysis of *Sall4* expression in cells treated with different small-molecule cocktails during stage 1.

(E) qRT-PCR analysis of XEN cell markers expression in cells treated with different small-molecule cocktails at day 12.

Data are represented as mean ± SD.

See also Figure S2.

co-expressed SALL4, GATA4, and SOX17, master genes of XENs (Lim et al., 2008) (Figure 1C). Our qRT-PCR analysis further detected the expression of other XEN marker genes, such as *Sox7* and *Gata6* in these colonies (Figure 1D). More importantly, we found expression level of these genes was comparable to that of embryo-derived XEN cells (eXENs) (Kunath et al., 2005) (Figure 1D). Then, we referred to the epithelial cells expressing XEN markers as XEN-like cells.

We next examined whether these XEN-like cells represent an intermediate state of chemical reprogramming. Using a XEN-expressing surface protein, EpCAM, we enriched XEN-like cells by fluorescence-activated cell sorting (FACS) and found that selection for EpCAM-positive cells greatly enriched the proportions of cells forming XEN-like cell colonies and subsequently generating CiPSCs by >20-fold (Figures 1E and S1B). Moreover, in another study by our group, we generated CiPSCs from neural stem cells and intestinal epithelial cells (Ye et al., 2015) and found that EpCAM also strongly enriched the cells that formed XEN-like colonies during chemical reprogramming and further generated CiPSCs from these two initial cell types (Figure S1C).

In addition, we speculated that transitional colonies that co-expressed XEN master genes and pluripotency-associated genes could be captured, if pluripotent stem cells were induced from XEN-like cells. To determine this, we detected cells that co-expressed GATA4 and OCT4 during stage 2 of reprogramming (Figure S1D), and we observed cell colonies in stage 3 expressing GATA4 in the peripheral and expressing p*Oct4*-GFP in the middle of the colonies, which could be the cell colonies under cell fate transition from XEN-like cells to pluripotent stem cells (Figure S1D). Together, these results indicate that the XEN-like cells represent an intermediate state of chemical reprogramming toward pluripotency.

Identification of Small Molecules that Promote the Transition from Fibroblasts to XEN-like Cells

The identification of an intermediate state of chemical reprogramming provides a new opportunity to optimize the reprogramming conditions and to screen novel small-molecule boosters in early reprogramming by using XEN-like colony numbers as the readout. In our previous study, we found that an enhanced concentration of CHIR99021 (from 10 μM to 20 μM) during stage 1 facilitates the generation of CiPSCs from neonatal and adult fibroblasts (Hou et al., 2013). Here, we found that the increased concentration of CHIR99021 is also beneficial for the formation of XEN-like colonies from MEFs (Figure 2A). Through qRT-PCR analysis, we found that

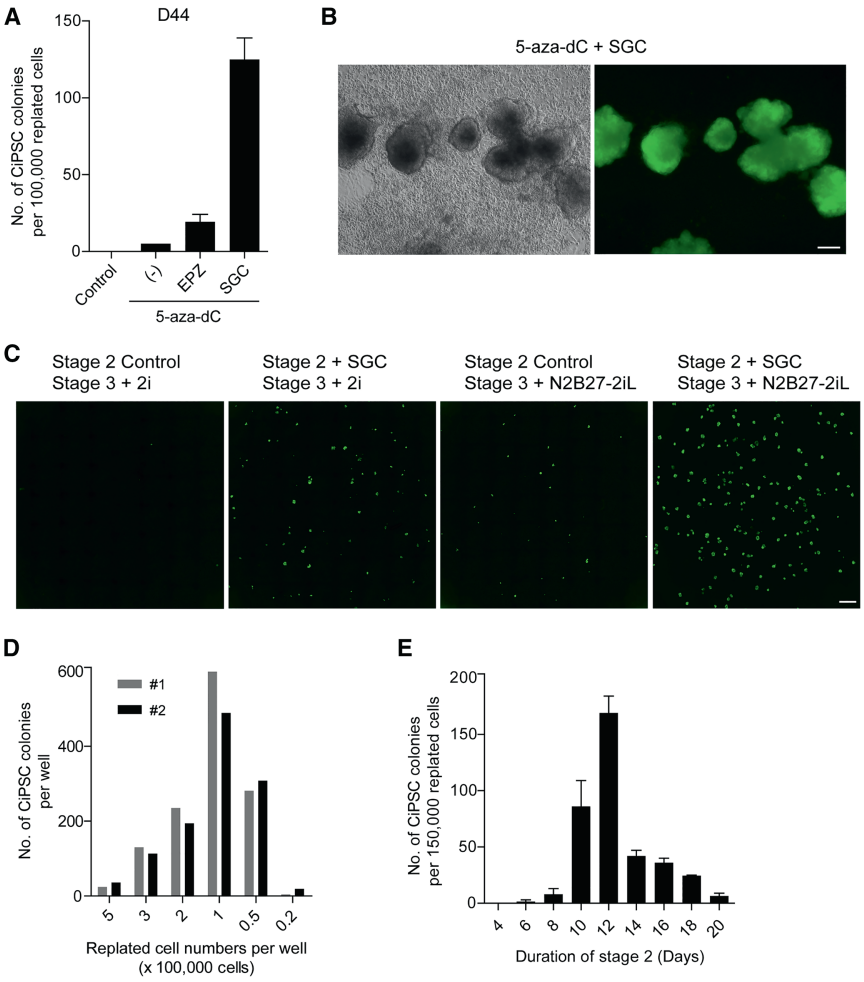


Figure 3. Identification of Small Molecules that Promotes the Transition from a XEN-like State to a Pluripotency State

(A) Numbers of CiPSC colonies at day 44 induced with a control cocktail (VC6TFAZ) and with 5-aza-dC, 5-aza-dC plus EPZ004777 (EPZ), or SGC0946 (SGC) in stage 2 of 12 days.

(B) Phase and fluorescence images of primary CiPSC colonies in stage 3 (day 40) with treatment with VC6TFAZ plus 5-aza-dC and SGC0946 in stage 2. Scale bar, 100 μm.

(C) High content screening fluorescent imaging of primary CiPSC colonies as indicated by p*Oct4*-GFP expression at day 40 using different reprogramming conditions. The views were jointed together with all 4× magnification fluorescent views in a well by Meta Xpress software. The control in stage 2 represents VC6TFAZ, 2i in stage 3 represents 2i-medium used in initial protocol. Scale bar, 2,000 μm.

(D) Numbers of CiPSC colonies at the end of reprogramming from different densities of cells in each well (6-well plate) re-plated at day 12. Numbers 1 and 2 were two independent experiments.

(E) Numbers of CiPSC colonies at the end of reprogramming, with a different time course in stage 2. Data are represented as mean ± SD.

See also Figure S3.

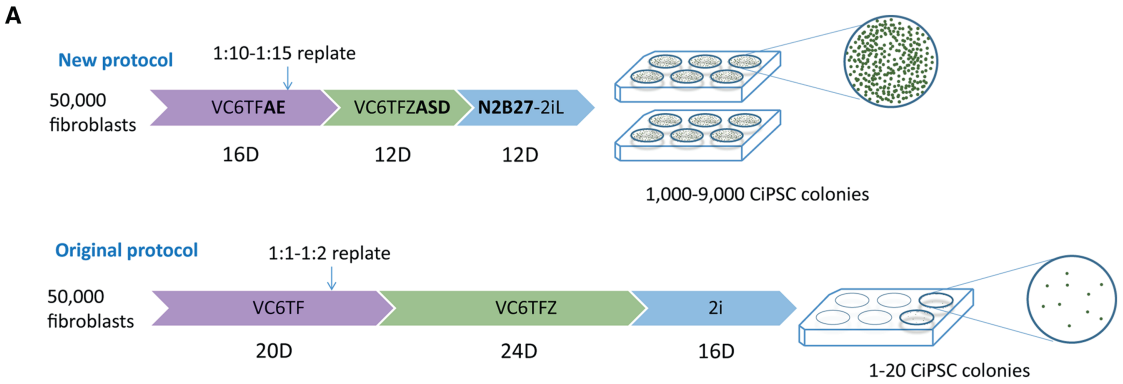
The Identification of Small Molecules that Promote the Transition from a XEN-like to a Pluripotent State

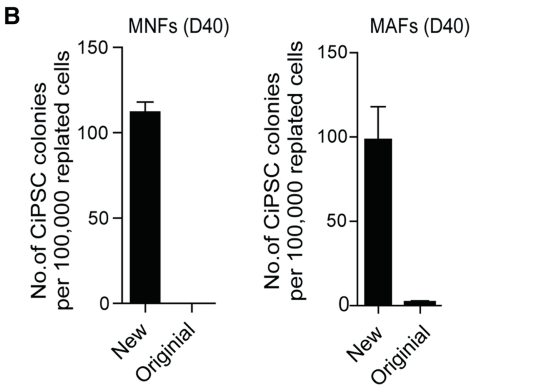
The enhanced generation of XEN-like colonies facilitated the screening of small molecules that promote the transition from XEN-like cells to CiPSCs during stages 2 and 3. Previously, the optimized duration for stage 2 was 20–24 days; few CiPSC colonies could be obtained if stage 2 was shortened to 12 days (Hou et al., 2013). Here, we conducted small-molecule screenings in the presence of small molecule cocktail VC6TFA plus DZNep (Z) for only 12 days instead of 20–24 days of stage 2. After screening 88 selected small molecules, we found that CiPSC colonies formed in stage 3 only when supplemented with 5-aza-dC during stage 2 (Figure 3A). Next, we found that 5-aza-dC (D) and EPZ004777 (E) had synergistic effects and promoted the kinetics of stage 2 (Figure 3A). By using a cocktail of eight small molecules, VC6TFAZDE, for 12 days during stage 2, we were able to obtain up to 20 CiPSC colonies from 100,000 re-plated cells at the end of the reprogramming process of 44 days (Figure 3A). In contrast, few CiPSC colonies were obtained during the same time course without 5-aza-dC and EPZ004777.

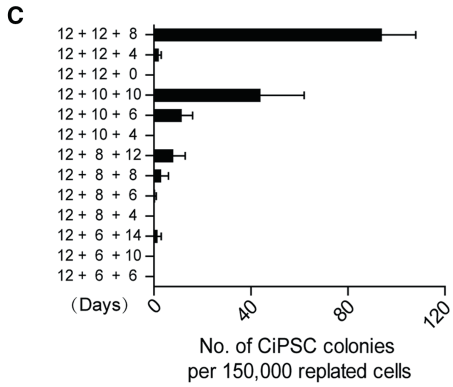
Strikingly, when we used another DOT1L inhibitor, SGC0946 (S), in place of EPZ004777 during stage 2 (Figures S3A and S3B), the reprogramming efficiency was increased further by as much as 5-fold (Figures 3A and 3B), particularly when an optimized 2i-medium (N2B27-2iL medium) was used (Figures 3C and S3B). Next, we further optimized the duration and concentration of the small molecules and the re-plating cell density

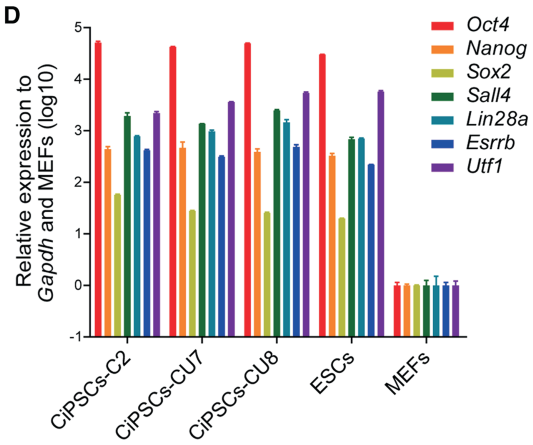
this increased concentration of CHIR99021 promotes an up to 10-fold increase in the expression of XEN master genes *Gata4* and *Sox17* and an epithelium cell marker, *EpCAM* (Figures 2B and S2A).

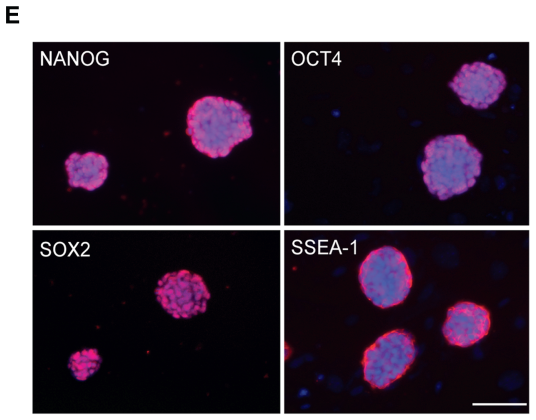
Next, we tested the effects of a selected small-molecule library of previously reported reprogramming boosters (Mikkelsen et al., 2008; Onder et al., 2012; Wang et al., 2011) in the presence of a small-molecule cocktail, VC6TF, with 20 μM CHIR99021, on the reprogramming of cell fate from fibroblasts to XEN-like cells. Among the tested small molecules, an RA agonist, AM580 (A), and a DOT1L inhibitor, EPZ004777 (E), enhanced the formation of XEN-like colonies by 2- to 3-fold (Figure 2C). When these small molecules were used together in a cocktail of seven small molecules, VC6TFAE, the number of XEN-like colonies was enhanced by >5-fold (Figure 2C). These findings were further validated by counting the numbers of SALL4 and GATA4 double-positive colonies and by detecting the expression of XEN marker genes by qRT-PCR (Figures 2D, 2E, and S2B–S2D). Together, the numbers of XEN-like colonies, an indicator of early reprogramming, could be enhanced more than 50-fold by adjusting the small molecule concentrations and adding additional small molecules.

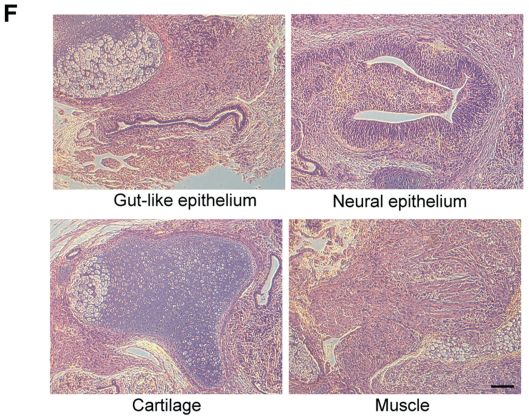


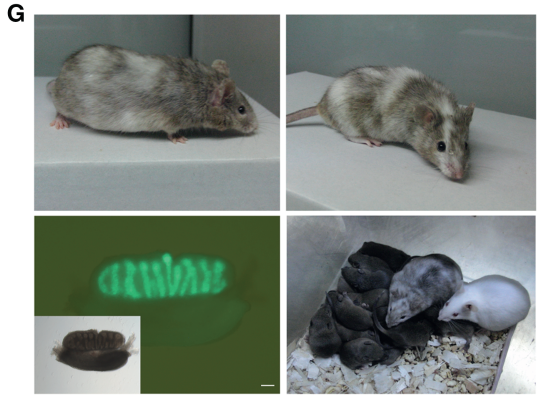












(legend on next page)

(Figures 3D, 3E, and S3C–S3E). Using the small-molecule cocktail VC6TFAZDS during stage 2 for 12 days, we were able to induce ~100–600 CiPSC colonies from 50,000 re-plated cells during the final stage of chemical reprogramming. Notably, although SGC0946 can be more effective than EPZ004777 during stage 2, SGC0946 was not able to substitute for EPZ004777 during stage 1 because cell viability was decreased if SGC0946 was used from the start of chemical reprogramming. Importantly, CiPSC colonies emerged from 24%–53% of the XEN-like colonies in five independent experiments, indicating the transition ratio of a single XEN-like cell in the start of stage 2 to CiPSCs in the end of stage 3 (Figure S3F). In addition, almost all CiPSC colonies were derived from XEN-like colonies, even though the efficiency was greatly improved (Figure S3G).

In particular, we found that the duration of the small-molecule treatment and the re-plating cell density were both highly critical during the later stages of reprogramming. Although the old protocol favors a re-plating cell density of 300,000 cells per well (Hou et al., 2013), the new protocol favors a cell density of 50,000–100,000 cells per well of a 6-well plate (Figure 3D). Furthermore, our initial protocol required 20–24 days of reprogramming during stage 2 to achieve optimal reprogramming efficiency (Hou et al., 2013), whereas our new protocol had an optimal stage 2 duration of 12 days (Figure 3E). In summary, we were able to greatly improve the cell transition from XEN to pluripotent stem cells by adding new small molecules, modifying the small-molecule structure and optimizing the re-plating density and time course of small-molecule treatment.

A Robust CiPSC Induction Protocol Was Established by Modulating the Cell Transitions through a XEN-like State

Next, we combined the optimized reprogramming conditions for stages 1, 2, and 3. Using this new protocol, a well of 50,000 initial fibroblasts was induced, and the cells were expanded to >1,000,000 or more re-plated cells (re-plated into 10–15 wells). A total of 1,000–9,000 CiPSC colonies were obtained at the end of the reprogramming period with a total induction time of 40 days (16, 12, and 12 days for stages 1, 2, and 3, respectively) (Figure 4A). Moreover, this new protocol was reproduced independently more than 30 times, and CiPSCs could also be generated from neonatal dermal fibroblasts (MNFs) and adult lung fibroblasts (MAFs) at a significantly enhanced efficiency (Figure 4B).

We further examined the minimal time course required in inducing CiPSCs by using this new small-molecule cocktail and the optimized reprogramming conditions. We found that a minimum of 12 days were required in the formation of XEN-like colonies (cells were re-plated at day 8), and at least another 14 days were required to induce CiPSCs from XEN-like cells (Figure 4C). In total, at the cost of efficiency, a minimum of 26 days are required to induce CiPSCs by using the new protocol (Figure 4C). In comparison, at least 44 days were required to generate only 0–1 CiPSC colony from 40,000 initial cells using the original protocol at day 44.

CiPSC colonies were then picked to establish CiPSC lines for further characterization. As shown by qRT-PCR and immunostaining, the CiPSCs expressed all the tested marker genes for pluripotent stem cells, such as *Oct4*, *Sox2*, and *Nanog* (Figures 4D and 4E). Using RNA sequencing (RNA-seq), we found that CiPSCs induced with this new protocol had gene expression profiles similar to those of ESCs (Figures S4A and S4B). CiPSCs were further tested for their potential for in vivo development, and we found that all six tested CiPSC lines were able to form teratoma after injection into SCID mice (Figure 4F) and generate chimeric mice after blastocyst injection (Figure 4G). Among five tested CiPSC lines, four lines showed germline integration potential in chimeric mice, and germline transmission offspring were obtained from chimeric mice (Figure 4G). Moreover, CiPSCs were maintained with normal karyotypes (Figure S4C). Together, these results indicate that we were able to establish a robust CiPSC induction protocol by manipulating the cell fate transition more precisely through the XEN-like state.

Gene Expression Dynamics during CiPSC Generation

The robust CiPSC induction method developed here allowed us to better study the molecular mechanisms underlying the reprogramming process. We then examined the expression of some typical pluripotency-associated genes during chemical reprogramming, and we revealed the sequential expression of pluripotency genes (Figures 5A and 5C). Through single cell qRT-PCR, we found that ~50% of the cells in stage 2 co-expressed XEN cell markers (Figure S5A).

Figure 4. A Robust CiPSC Induction Protocol Was Established by Modulating the Cell Transitions through a XEN-like State

(A) Schematic comparison of the new protocol in this study and our initial protocol (Hou et al., 2013). In total, 1–20 CiPSC colonies (in 2 wells) were induced from initial 50,000 fibroblasts in a 60-day induction using the initial protocol. Whereas 1,000–9,000 CiPSC colonies (in 10–15 wells) were obtained from initial 50,000 fibroblasts after 40 days of small-molecule treatment by using the new protocol. 2iL represents 2i-medium with LIF.

(B) Comparison of the original protocol and the new protocol in primary CiPSC colony numbers (day 40) induced from mouse neonatal fibroblasts (MNFs, left) and mouse adult fibroblasts (MAFs, right), respectively.

(C) Numbers of CiPSC colonies generated under shortened durations for each stage. For example, "12+10+6" represents a sequential duration of 12, 10, and 6 days for stage 1, 2, and 3, respectively. As shown, the minimal time course required for CiPSC induction was 26 days (12+8+6).

(D) Pluripotency marker expression in CiPSC colonies induced by the new protocol, analyzed by qRT-PCR.

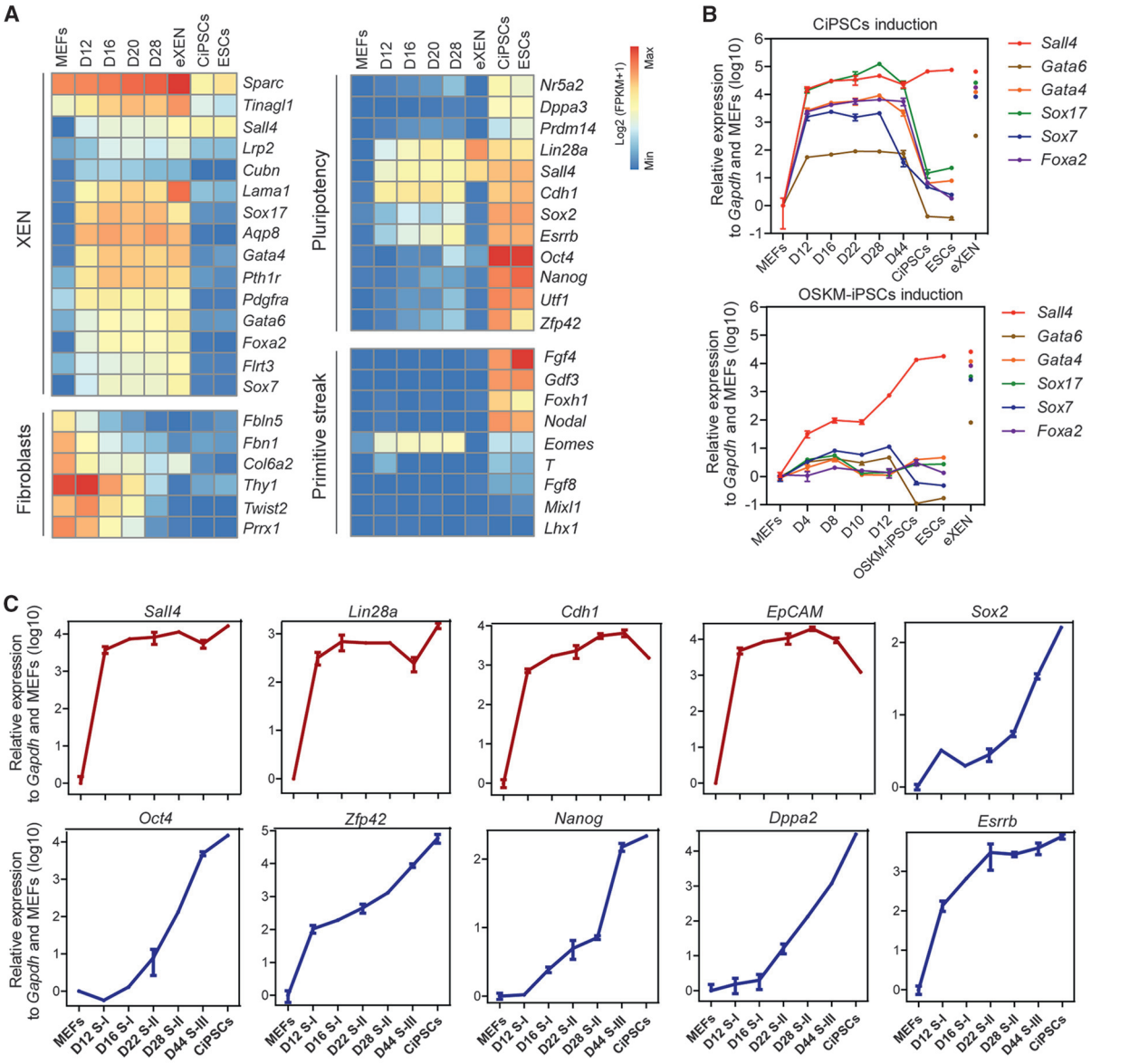
(E) Immunofluorescence of pluripotency markers in CiPSC colonies induced by the new protocol. Scale bar, 100 μm.

(F) Teratoma generated from CiPSCs induced by the new protocol. Top left: gut-like epithelium (endoderm); top right: neural epithelium (ectoderm). Bottom left: cartilage (mesoderm); bottom right: muscle (mesoderm). Scale bar, 100 μm.

(G) Chimeric mice generated from CiPSCs induced from MEFs (top left) and MNFs (top right). Germline contribution of CiPSCs in the gonad of a chimeric mouse embryo was indicated by p*Oct4*-GFP expression (bottom left). Germ-line transmission offspring generated from CiPSCs were shown (bottom right). Scale bar, 100 μm.

Data are represented as mean ± SD.

See also Figure S4.



(legend on next page)

Notably, similar to embryo-derived XEN cells, the XEN-like cells formed during chemical reprogramming expressed several pluripotency genes, such as *Sall4* and *Lin28a*, during the early stages of reprogramming (Lim et al., 2008; McDonald et al., 2014) (Figures 5A–5C). During an extended time in culture, the XEN-like cells expressed other pluripotency-associated genes on days 16–28, such as *Oct4*, *Esrrb*, and *Dppa2* (Figures 5B and 5C). During stage 3, the expression of most pluripotency marker genes, including *Nanog*, was activated in CiPSCs (Figures 5B and 5C). This finding indicates a process of sequential gene activation from XEN-like cells to pluripotent stem cells. Interestingly, *Sall4*, *Lin28a*, *Esrrb*, the major genes associated with pluripotency that are highly expressed during stages 1 and 2 of cell reprogramming, have previously been reported to be predictive markers of transcription factor-induced reprogramming and to be sufficient for inducing iPSCs with high quality when concomitantly expressed with *Nanog* (Buganim et al., 2012, 2014).

Next, we examined the role of additional small molecules. Through qRT-PCR analysis and RNA sequencing, we found that AM580 and EPZ004777 both promote the expression of XEN marker genes, such as *Sall4*, *Gata4*, and *Sox17*, during stage 1 of chemical reprogramming from fibroblasts to XEN-like cells (Figure 2E). SGC0946 and 5-aza-dC promote the expression of pluripotency genes, such as *Oct4* and *Dppa* family genes in XEN-like cells, during stage 2 of chemical reprogramming (Figure S5B). These findings indicate that stage 1 of chemical reprograming into XEN-like cells is promoted by additional small molecules that act by enhancing the expression of XEN master genes, and stage 2 can be shortened possibly due to the enhanced activation of pluripotency-associated genes by additional small molecules.

We next investigated whether the reprogramming process through a XEN-like state is a unique route toward pluripotency compared to that of the transgenic strategy, which uses OSKM (Takahashi et al., 2007, 2014; Takahashi and Yamanaka, 2006). Notably, OSKM-induced reprogramming processes do not show XEN gene profiles, analyzed by qRT-PCR (Figure 5B). This finding was also consistent with the previous gene profiling data during the reprogramming process in other reports (Golipour et al., 2012; Mikkelsen et al., 2008; Polo et al., 2012; Sridharan et al., 2009) (Figure S5C).

Moreover, we further examined whether primitive streak genes were expressed during the chemical reprogramming process, because a primitive streak state has been reported during the process of OSKM-induced reprogramming (Takahashi et al., 2014). The expression of primitive streak markers, such as *T* and *Mixl1*, was not detected during chemical reprogramming (Figure 5A), demonstrating a unique XEN-like state during the chemical reprogramming process, which differs from that of the reprogramming induced by transgenes (Figure 5D).

XEN Master Gene Expression Is Essential during Chemical Reprogramming

To further support the XEN state as an intermediate for chemical reprogramming and to understand the role of XEN master genes in chemical reprogramming, we performed knockdown and ectopic expression experiments. We found that the knockdown of any one of the key XEN genes *Sall4*, *Gata4*, *Gata6*, or *Sox17* led to a significant downregulation in the mRNA levels of the other XEN genes and decreased XEN-like colony numbers, thus resulting in less *Oct4* expression and fewer CiPSCs at the end of the reprogramming period (Figures 6A–6C, S6A, and S6B). Consistently with our previously results, the expression of XEN genes was essential to *Oct4* expression (Hou et al., 2013; Shu et al., 2013, 2015). Moreover, in another study, we found that XEN-gene expression was also enhanced in in chemical reprogramming from neural stem cells and intestinal epithelium cells (Ye et al., 2015), and the knockdown of these genes impaired XEN-like colony formation and further CiPSC induction from these two initial cell types (Figure S6C). These results further indicate that the XEN-like state is essential to the chemical reprogramming process. In contrast, these XEN master genes, such as *Gata4*, *Gata6*, and *Sox17*, were not required in OSKM-induced reprogramming (Figure S6D), which suggest different roadmaps underlying chemical reprogramming and OSKM-induced reprogramming (Figure 5D).

Furthermore, we found that the overexpression of two of the XEN master genes (*SALL4* plus *GATA4* or *SALL4* plus *GATA6*) in fibroblasts sufficed in inducing XEN-like colony formation in the absence of the three key small molecules, CHIR99021, 616452, and forskolin (Figure 6D). The resulting XEN-like cells showed gene expression pattern similar to that of the small molecule-induced XEN-like cells (Figure S6E). Moreover, *Oct4* expression was detected in the XEN-like colonies induced by the two combinations of XEN cell master transcription factors (Figures 6E and 6F). These findings suggest that XEN genes are both necessary and sufficient to initiate the expression of *Oct4*, a master gene of pluripotency.

Notably, although these XEN master gene-induced XEN-like cells expressed *Oct4*, they could not be further reprogrammed into iPSCs, even with a prolonged culture in 2i-medium. We

Figure 5. Gene Expression Dynamics during CiPSC Generation

(A) Gene expression heatmap of XEN, fibroblasts, pluripotency, and primitive streak-related genes. The gene expression patterns of MEFs, eXEN, CiPSCs, ESCs, and reprogrammed cells at day 12 (D12), day 16 (D16), day 20 (D20), and day 28 (D28) during chemical reprogramming were analyzed by RNA-seq.

(B) Comparison of XEN-related gene expression pattern at indicated time points during chemical reprogramming (top) and OSKM-induced reprogramming (bottom) measured by qRT-PCR. CiPSCs, OSKM-iPSCs, ESCs, and eXEN were set as controls.

(C) Dynamic expression change in pluripotency-associated genes and MET-related genes at the indicated time points during chemical reprogramming, examined by qRT-PCR.

(D) Schematic representations of the two routes of somatic reprogramming by using chemical approach (left) and transgenic approach (right). In the transgene-induced reprogramming process, primitive streak-like mesendoderm (PSMN) has been reported as the transient state (Polo et al., 2012; Takahashi et al., 2014), whereas in chemically induced reprogramming, as reported in this study, a unique XEN-like state bridges the transition of fibroblasts to CiPSCs.

Data are represented as mean ± SD.

See also Figure S5.

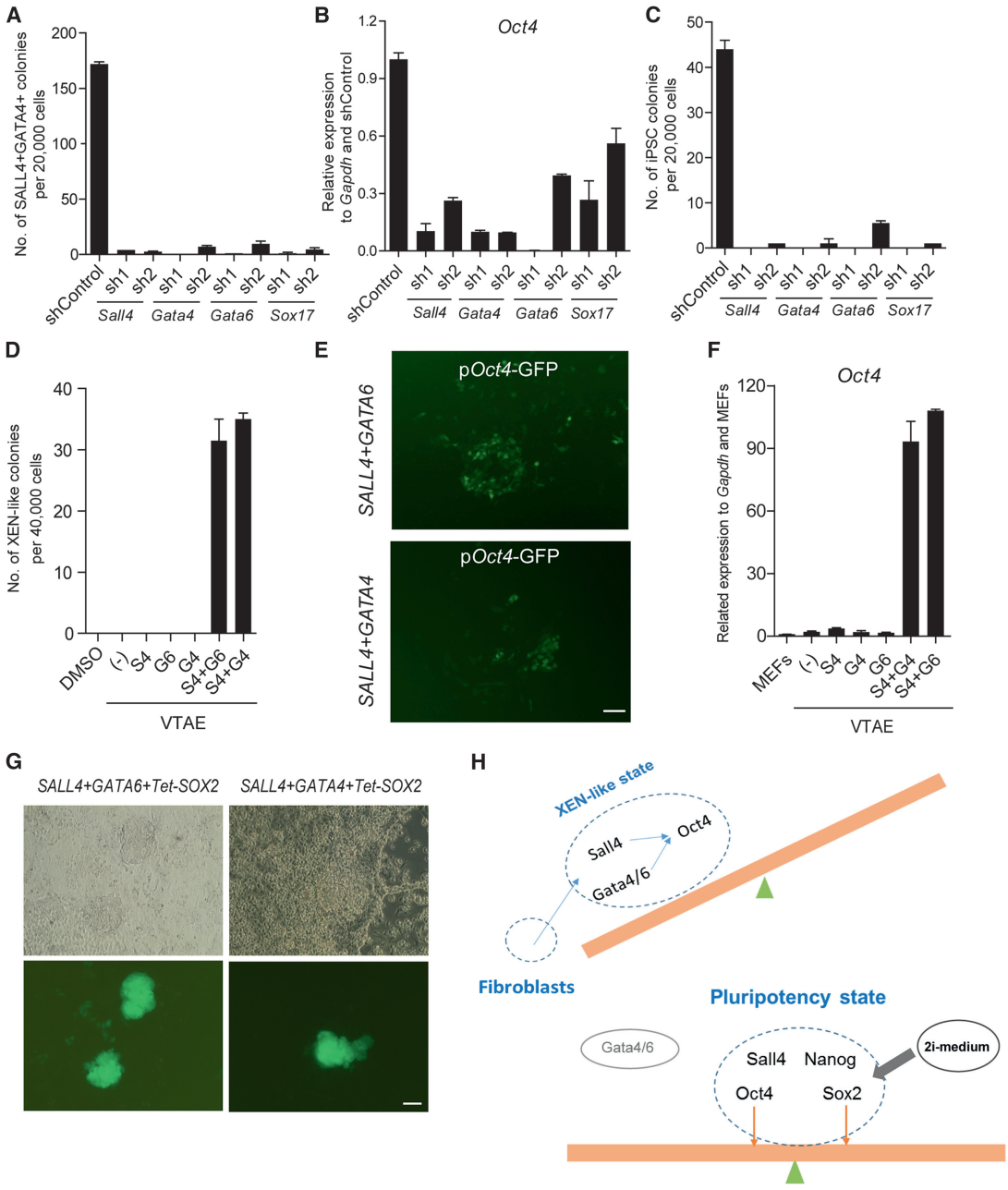


Figure 6. XEN Master Gene Expression Is Essential during Chemical Reprogramming

(A) Numbers of SALL4 and GATA4 double-positive colonies with *Sall4*, *Gata4*, *Gata6*, or *Sox17* knockdown at day 12 of chemical reprogramming. Non-targeting vector shRNA (shControl) was used as negative control. Sh1 and sh2 represent two shRNA vectors for each gene.

(B) The expression of *Oct4* on day 28 with *Sall4*, *Gata4*, *Gata6*, or *Sox17* knockdown measured by qRT-PCR relative to that treated with a non-targeting vector (shControl).

(C) Numbers of CiPSC colonies with *Sall4*, *Gata4*, *Gata6*, or *Sox17* knockdown.

(D) Numbers of XEN-like colonies by overexpression of *SALL4* (*S4*), *GATA4* (*G4*), *GATA6* (*G6*), or their combinations in the presence of small-molecule cocktail VTAE (withdrawal of C6F from VC6TFAE) treatment. (–) represents cells treated with VTAE without the overexpression of XEN-related genes. DMSO-treated cells are shown as negative controls.

(E) Fluorescence images of colonies with p*Oct4*-GFP expression induced by the overexpression of *SALL4* plus *GATA4* or *GATA6* in the presence of VTAE. Scale bar, 100 μm.

(F) qRT-PCR analysis of *Oct4* induced by overexpression of *SALL4*, *GATA4*, *GATA6*, and their combination in the presence of VTAE on day 20.

(G) Phase and fluorescence images of primary iPS colonies induced by *SALL4*, *GATA6* (or *GATA4)* plus Dox-induced *SOX2*. Scale bar, 100 μm.

(legend continued on next page)

found the exogenous XEN genes downregulated the endogenous expression of *Sox2* (Figure S6F). Accordingly, when *Sox2* was exogenously provided in an appropriate time window after XEN gene overexpression, iPSCs were obtained (Figures 6G and S6G). This finding is consistent with our previous findings of a seesaw model in regulating pluripotency establishment, in which the *Gata* family genes and *Sox2* should be in a balance to achieve pluripotency (Shu et al., 2013), whereas in chemical reprogramming such a balance could be a dynamic process rather than a steady equilibrium (Figure 6H).

The XEN-like Intermediates Resemble Embryo-Derived XEN Cells in Gene Expression Patterns, In Vivo Development Potential, and Reprogramming Potential

To better understand this XEN-like state, we further compared the chemically-induced XEN-like cells to embryo-derived XEN cells (eXENs) (Kunath et al., 2005). We found that XEN-like cells could not be maintained in the traditional XEN culture medium (Figure S7A), whereas eXEN cell lines could be maintained long-term and expanded in the stage 1 medium of chemical reprogramming for more than 27 passages, with the gene expression pattern and in vivo development potential similar to eXENs maintained in traditional XEN medium (Figures 7A–7D, S7B, and S7C). In addition, we found that by using stage 1 medium with chemical cocktail VC6TF, chemically-derived eXEN cell lines (CeXENs) could be established directly from blastocysts and expanded long-term for >32 passages, with a XEN-like gene expression pattern and in vivo integration capability into extraembryonic parietal endoderm (Figures 7A–7D, S7B, and S7C).

By qRT-PCR, we found XEN-like cells express a comparable level of XEN master genes to that of eXENs and CeXEN (Figures 1D, 7A, and 7B). Through global gene expression profiling at the end of stages 1 and 2, we found that XEN-like cells showed gene expression profiles close to eXENs and CeXENs (Figure 7C). Through a principal component analysis (PCA) analysis of gene expression profiling, we found a clear roadmap from fibroblasts toward pluripotent stem cells through such a XEN-like state close to eXENs and CeXENs (Figure 7D).

In particular, the XEN-like intermediates in chemical reprogramming were closer to CeXENs than traditional eXENs in gene expression profiles (Figures 7C and 7D). Moreover, we found that mRNA level of *EpCAM*, *Cdh1* and *Sox2* in XEN-like cells and CeXENs was notably higher than that in traditional eXENs (Figures 7A and S7D). It is possible that these differences between XEN-like cells and traditional eXENs in gene expression pattern were resulted from their different culture conditions. Interestingly, when we further checked gene profiling data of in vivo XENs in a previous report (Yan et al., 2013), we found that the authentic XEN cells in vivo express *EPCAM* and *CDH1* in a high level comparable to that of ESCs and express *SOX2* in a relatively low level, a pattern similar to that of XEN-like cells and CeXENs, but not eXENs (Figures S7D and S7E). These suggest that although XEN-like cells showed some differences to eXENs in culture conditions and gene expression patterns, they were more similar to CeXEN, another type of embryo-derived XEN cells.

We further examined the in vivo development potential of XEN-like cells during chemical reprogramming. XEN-like cells induced in different time courses of chemical reprogramming were injected into mouse blastocysts. Similarly to eXEN cells, XEN-like cells at days 11–25 of chemical reprogramming were able to integrate into the parietal endoderm of the extraembryonic tissues with a comparable efficiency of eXENs, without any integration in the embryos (Figures 7A, 7E, and S7C). In particular, XEN-like cells in day 16 of chemical reprogramming showed the highest ratio of XEN integration (Figure S7C). These findings suggest that XEN-like cells resemble embryo-derived XEN cells in terms of development potential.

Furthermore, we found that either eXENs derived by traditional XEN culture medium or CeXENs established by the stage 1 medium of chemical reprogramming, were capable of further reprogramming into CiPSCs by using the protocol of late chemical reprogramming in stages 2 and 3 (Figures 7A and 7F). Notably, 17–34 CiPSC colonies were generated from 2,000 CeXENs within 24 days, a reprogramming efficiency even higher than that of XEN-like cells. CiPSCs generated from eXENs and CeXENs were further characterized to possess an expression pattern similar to pluripotency stem cells and the potential to generate chimeric mice (Figures 7F, S7F, and S7G). These findings further support that XEN-like cells induced in the early stage of CiPSC generation, which were similar to CeXENs, were amenable to being further reprogrammed in the late stage of chemical reprogramming (Figure 7A).

DISCUSSION

Our discovery of the XEN-like state uncovered a unique route in chemical reprogramming of somatic cells toward pluripotency, but not in OSKM-induced reprogramming with a reported primitive streak-like intermediates (Takahashi et al., 2014) (Figure 5D). Similarly to XEN cells in vivo, the induced XEN-like cells have already expressed *Sall4* and *Lin28a*, two master genes of pluripotency. It is possible that the shared genes expressing in both XEN cells and pluripotent stem cells, such as *Sall4* and *Lin28a*, make the pluripotency state more accessible during the cell fate transition from the XEN-like state to pluripotency. Moreover, we demonstrated the sequential expression of many pluripotency marker genes, such as *Oct4* and *Esrrb*, in XEN-like cells, during stage 2 of chemical reprogramming, indicating that the pluripotency network is easily initiated in XEN-like cells. This finding is consistent with the recent report that in vivo XEN cells spontaneously transition into cells constituting epiblast, which are pluripotent (Xenopoulos et al., 2015). Interestingly, in stage 3 of chemical reprogramming, 2i-medium containing a FGF/ERK signaling inhibitor, PD0325901, initiated the expression of *Nanog* and *Sox2* in XEN-like cells and cell fate conversion

(H) The "seesaw" model in chemical reprogramming. Chemical reprogramming is a dynamic process with an unbalanced state inclining to extraembryonic endoderm in early reprogramming (top), which is further balanced in the end of reprogramming (bottom).
Data are represented as mean ± SD.
See also Figure S6.

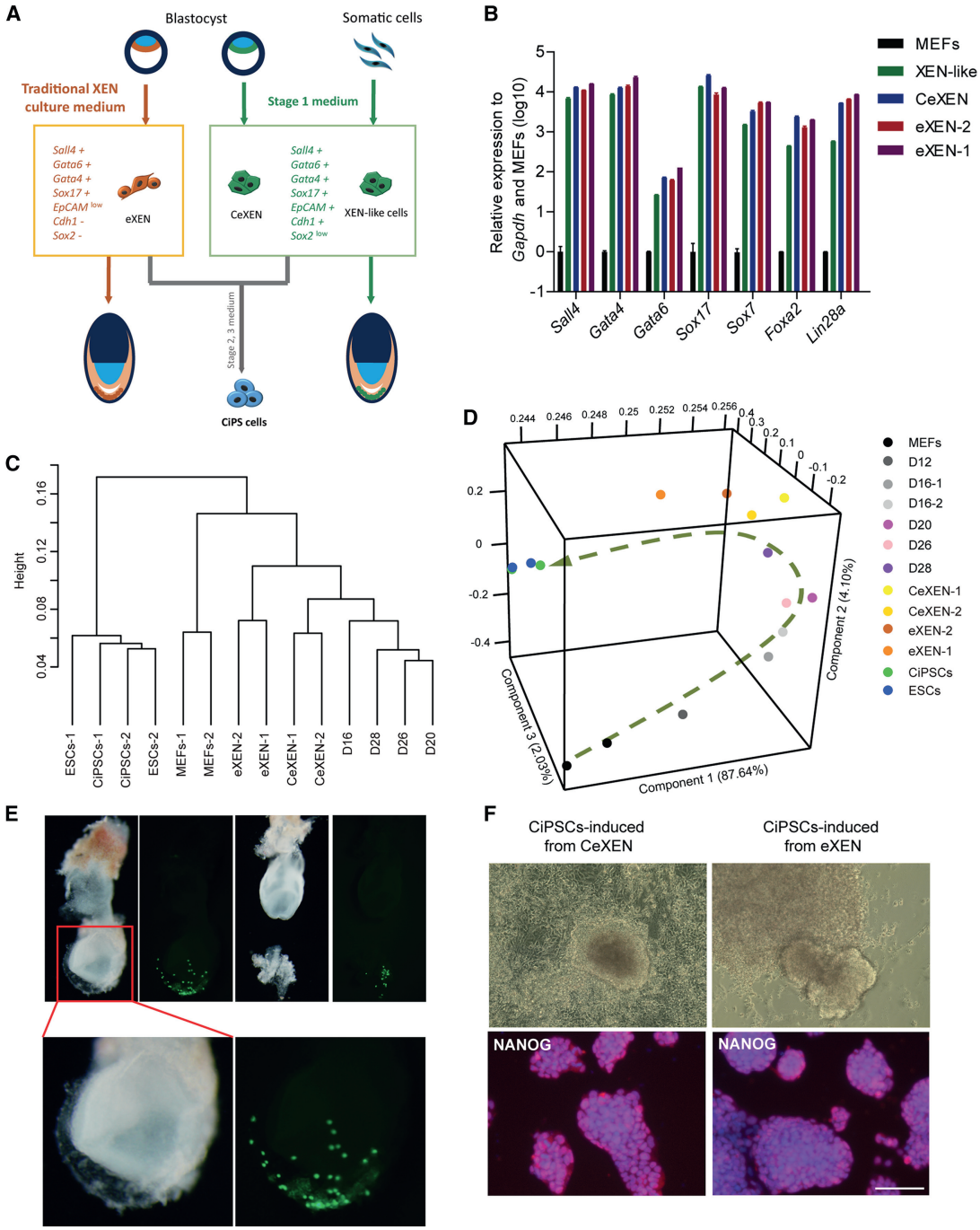


Figure 7. The XEN-like Intermediates Resemble Embryo-Derived XEN Cells in Gene Expression Patterns, In Vivo Development Potential, and Reprogramming Potential

(A) Schematic representation of different types of XEN cells mentioned in this study. All three of these types of XEN cell lines can form parietal endoderm chimera conceptus after injection to blastocyst and can be further reprogrammed to CiPSCs with stage 2 and 3 medium of chemical reprogramming. Their expression patterns were indicated as shown. +, highly expressed; low, expressed in a low level; –, no expression detected.

(B) Relative expression of XEN-related genes in different cell types as indicated measured by qRT-PCR. eXEN-1 and eXEN-2 were two sublines of eXENs, maintained in traditional XEN culture medium (Kunath et al., 2005) and stage 1 medium of chemical reprogramming, respectively.

(C) Hierarchical clustering of global gene expression profiles in different cell types. XEN-like cell samples at different time points (day 16, 20, 26, and 28) during chemical reprogramming were indicated as D16, D20, D26, and D28. Controls were two batches of MEFs (MEFs-1, MEFs-2), eXEN-1, eXEN-2, two CeXEN cell lines (CeXEN-1, CeXEN-2), two CiPS cell lines (CiPSCs-1, CiPSCs-2), and two ES cell lines (ESCs-1, ESCs-2).

(D) Principal component analyses (PCA) of global gene expression profiles of indicated cell types mentioned above.

(legend continued on next page)

into pluripotent cells, reminiscent of the cell fate determination between primitive endoderm (also known as XEN) and epiblast regulated by FGF/ERK signaling in the mouse blastocyst as previously reported (Yamanaka et al., 2010). The compatibility of gene expression in XEN and pluripotent cells and the close relations of XEN and epiblast in development make the XEN-like state an ideal bridge toward pluripotency.

Moreover, the master genes of XEN activated during chemical reprogramming directly contribute to the establishment of pluripotency. Our group has previously shown that *Gata4* and *Gata6*, two of the XEN-genes, were able to substitute for *Oct4* in transgenic reprogramming in the presence of exogenous *Sox2*, *Klf4*, and *c-Myc* (Shu et al., 2013). In addition, they were found to activate the expression of *Sall4* and subsequently *Oct4* (Shu et al., 2015). In this study, we found the overexpression of *Sall4* and *Gata4/6*, master genes of XENs, is sufficient to stimulate the expression of *Oct4*, independent of the key small molecules that used in chemical reprogramming. Together, XEN-like cells could be a unique intermediate state of somatic reprograming, which is primed for further reprogramming.

However, in the later stage of reprogramming, these XEN genes need to be silenced to stimulate the expression of additional pluripotency genes, such as *Nanog* and *Sox2*. As reported, there were several mutual antagonistic mechanisms between XEN genes and pluripotency-associated genes, such as the competition between *Sox17* and *Sox2* and the incompatibility between *Gata6* and *Nanog* (Aksoy et al., 2013; Chazaud et al., 2006; Niakan et al., 2010). In this study, we found the expression of *Sox2* was repressed by the XEN genes. We also show that the previously reported "seesaw" between XEN genes and *Sox2* (Shu et al., 2013) is still required at the end of chemical reprogramming, as the exogenously provided *Sox2* is required and sufficient to fulfil the late stage reprogramming initiated by the overexpression of *Sall4* and *Gata4/6*. These indicate the seesaw model in chemical reprogramming is more dynamic rather than a steady equilibrium, and the XEN-like state is transient in chemical reprogramming serving as a bridge of cell fate transitions from somatic cells to pluripotent stem cells, which can be "burned" when pluripotency is achieved.

Most important, the identification of the unique route mediated by the XEN-like state allows manipulating cell fate transitions more precisely in a stepwise manner. Chemical reprogramming efficiency was greatly improved by adding new small molecules to the cocktails at each stage, adjusting the concentration of CHIR99021, changing the components of the 2i-medium, modifying the small-molecule structure, and optimizing the re-plated cell density and time course of small-molecule treatment. Most notably, by slightly modifying the structure of EPZ004777 to become SGC0946, which is added in stage 2, the efficiency of chemical reprogramming was significantly increased. To date, the total yields of CiPSC colonies have been enhanced by up to 1,000-fold compared to the use of old protocol. Together, these findings show a major advantage for the small-molecule approach compared to the transgenic approach, as small molecules are easier to combine, control, optimize, adjust, and withdraw, thereby controlling gene expression and cell fate more precisely both spatially and temporally.

Overall, our discovery of the XEN-like state during chemical reprogramming in this study and our previous findings regarding the role of *Gata* family genes in inducing pluripotency together reveal a unique route toward pluripotency. In addition, small molecule-induced cell fate reprogramming can be greatly enhanced when manipulated more precisely by small molecule fine-tuning. Thus, chemical reprogramming provides a unique, promising strategy for future applications and for better understanding cell pluripotency and cell fate reprogramming.

EXPERIMENTAL PROCEDURES

CiPSC Induction from Fibroblasts

Small molecules that used in this study were described in Table S1. In an optimized protocol, MEFs were plated at 300,000 cells per 100 mm dish, or 50,000 cells per well of a 6-well plate. The next day (day 0), the culture was changed into stage 1 medium (containing 100 ng/ml bFGF, 0.5 mM VPA, 20 μM CHIR99021, 10 μM 616452, 5 μM tranylcypromine, 50 μM forskolin, 0.05 μM AM580 and 5 μM EPZ004777). On day 12, the cells were trypsinized, harvested and then re-plated at 50,000–200,000 cells per well of a 6-well plate (1:10–15). During days 12–16, concentrations of bFGF, CHIR, and forskolin were reduced to 25 ng/ml, 10 μM, and 10 μM, respectively. On day 16, XEN-like epithelial colonies were formed and the culture was changed into stage 2 medium (containing 25 ng/ml bFGF, 0.5 mM VPA, 10 μM CHIR99021, 10 μM 616452, 5 μM tranylcypromine, 10 μM forskolin, 0.05 μM AM580, 0.05 μM DZNep, 0.5 μM 5-aza-dC, and 5 μM SGC0946). On day 28, the culture was transferred into stage 3 medium (N2B27-2iL medium with 3 μM CHIR99021, 1 μM PD0325901, and 1,000 U/ml LIF). After another 8–12 days, 2i-competent, ESC-like, and GFP-positive (if using p*Oct4*-GFP reporter) CiPSC colonies emerged and were then picked up for expansion and characterization.

XEN-like Cells Chimera Assay

GFP-labeled XEN-like cells were induced from the GFP-labeled MEFs. For chimera test, GFP-labeled XEN-like cell colonies were picked and trypsinized to single cells. Approximately 10–15 XEN-like cells were injected into blastocysts and transferred to the uterus of E2.5 pseudopregnant females. Chimera conceptus between E6.5–8.5 were dissected carefully to keep the parietal yolk sac intact and observed with fluorescence stereoscopy.

CiPSC Induction from eXEN Cells

eXENs (or CeXENs) were plated at 4,000–20,000 cells per well of a 6-well plate on MEF feeders. Following the treatment with stage 1 medium for 4 days (optional), stage 2 medium for 12 days, and stage 3 medium for another 8–12 days, 2i-competent, ESC-like CiPSC colonies emerged and were then picked up for expansion and characterization.

A more detailed description of the materials and methods is provided in the Supplemental Experimental Procedures. All animal procedures were performed according to the Animal Protection Guidelines of Peking University, China.

(E) Representative images of two E7.5 chimeras generated by blastocyst injection with GFP-labeled XEN-like cells at day 16 of chemical reprogramming (the end of stage 1) (top). Magnified images focused on the distal part of one conceptus show the integration of XEN-like cells to parietal yolk sac (bottom). The separation of parietal yolk sac from one conceptus indicates the chimeric integration of the XEN-like cells was exclusively in the extraembryonic region (top right).
(F) Primary CiPSC colonies and immunostaining of NANOG in CiPSC lines induced from CeXEN cells (left) and eXEN cells (right). Scale bar, 100 μm.
Data are represented as mean ± SD.
See also Figure S7.

ACCESSION NUMBERS

The accession number for the RNA-seq data reported in this paper is NCBI GEO: GSE73631.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.11.017.

AUTHOR CONTRIBUTIONS

Y.Z., T.Z., J.G., X.Z., Y.F., J.Y., J.Z., G.M., J.G., S.Y. and L.C. performed experiments. Y.D., C.Z., T.W., L.S., and W.Y. contributed to mouse in vivo assays. Y.Z. and H.D. supervised the study and wrote the paper.

ACKNOWLEDGMENTS

We thank Dr. Janet Rossant for the gift of embryo-derived XEN cell line. We thank Shuguang Duo, Jiebin Dong, Qiong Huo, Weifeng Lai, Yin Ji, Xiaochen Li, and Hongxia Lv for technical assistance. We thank the Core Facilities at School of Life Sciences, Peking University for assistance with confocal microscopy and high content imaging. We thank Dr. Tuoping Luo and Dr. Chao Tang for helpful suggestions. This work was supported by the National Basic Research Program of China (973 Program 2012CB966401), the National Natural Science Foundation of China (91310305 and 31521004), the National Science and Technology Major Project (2013ZX10001003), Ministry of Science and Technology (2013DFG30680), Key New Drug Creation and Manufacturing Program (2011ZX09102-010-03), and Ministry of Education of China (111 Project) to H.D. This work was supported in part by a grant from BeiHao Stem Cell and Regenerative Medicine Translational Research Institute.

Received: May 11, 2015
Revised: September 24, 2015
Accepted: November 5, 2015
Published: December 10, 2015

REFERENCES

Aksoy, I., Jauch, R., Chen, J., Dyla, M., Divakar, U., Bogu, G.K., Teo, R., Leng Ng, C.K., Herath, W., Lili, S., et al. (2013). Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. EMBO J. *32*, 938–953.

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., Markoulaki, S., van Wietmarschen, N., Hoke, H., Wu, T., Ganz, K., Akhtar-Zaidi, B., He, Y., Abraham, B.J., Porubsky, D., et al. (2014). The developmental potential of iPSCs is greatly influenced by reprogramming factor selection. Cell Stem Cell *15*, 295–309.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev. Cell *10*, 615–624.

Chou, B.K., and Cheng, L. (2013). And then there were none: no need for pluripotency factors to induce reprogramming. Cell Stem Cell *13*, 261–262.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., and Wrana, J.L. (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell Stem Cell *11*, 769–782.

Gurdon, J.B. (1962). Adult frogs derived from the nuclei of single somatic cells. Dev. Biol. *4*, 256–273.

Hansson, J., Rafiee, M.R., Reiland, S., Polo, J.M., Gehring, J., Okawa, S., Huber, W., Hochedlinger, K., and Krijgsveld, J. (2012). Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. Cell Rep. *2*, 1579–1592.

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.

Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649–1661.

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.

Li, W., Li, K., Wei, W., and Ding, S. (2013). Chemical approaches to stem cell biology and therapeutics. Cell Stem Cell *13*, 270–283.

Lim, C.Y., Tam, W.L., Zhang, J., Ang, H.S., Jia, H., Lipovich, L., Ng, H.H., Wei, C.L., Sung, W.K., Robson, P., et al. (2008). Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. Cell Stem Cell *3*, 543–554.

McDonald, A.C., Biechele, S., Rossant, J., and Stanford, W.L. (2014). Sox17-mediated XEN cell conversion identifies dynamic networks controlling cell-fate decisions in embryo-derived stem cells. Cell Rep. *9*, 780–793.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Niakan, K.K., Ji, H., Maehr, R., Vokes, S.A., Rodolfa, K.T., Sherwood, R.I., Yamaki, M., Dimos, J.T., Chen, A.E., Melton, D.A., et al. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. Genes Dev. *24*, 312–326.

Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., Cahan, P., Marcarci, B.O., Unternaehrer, J., Gupta, P.B., et al. (2012). Chromatin-modifying enzymes as modulators of reprogramming. Nature *483*, 598–602.

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.

Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., Tang, X., Yang, H., Shen, L., Zuo, X., et al. (2013). Induction of pluripotency in mouse somatic cells with lineage specifiers. Cell *153*, 963–975.

Shu, J., Zhang, K., Zhang, M., Yao, A., Shao, S., Du, F., Yang, C., Chen, W., Wu, C., Yang, W., et al. (2015). GATA family members as inducers for cellular reprogramming to pluripotency. Cell Res. *25*, 169–180.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364–377.

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.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Sasaki, A., Yamamoto, M., Nakamura, M., Sutou, K., Osafune, K., and Yamanaka, S. (2014). Induction of

pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. Nat. Commun. *5*, 3678.

Wang, W., Yang, J., Liu, H., Lu, D., Chen, X., Zenonos, Z., Campos, L.S., Rad, R., Guo, G., Zhang, S., et al. (2011). Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. Proc. Natl. Acad. Sci. USA *108*, 18283–18288.

Xenopoulos, P., Kang, M., Puliafito, A., Di Talia, S., and Hadjantonakis, A.K. (2015). Heterogeneities in Nanog expression drive stable commitment to pluripotency in the mouse blastocyst. Cell Rep. *10*, 1508–1520.

Yamanaka, Y., Lanner, F., and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development *137*, 715–724.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1131–1139.

Ye, J., Ge, J., Zhang, X., Cheng, L., Zhang, Z., He, S., Wang, Y., Lin, H., Yang, W., Liu, J., et al. (2015). Pluripotent stem cells induced from mouse neural stem cells and small intestinal epithelial cells by small molecule compounds. Cell Res. http://dx.doi.org/10.1038/cr.2015.142.

CellPress

Cell Stem Cell
Article

Reprogrammed Stomach Tissue as a Renewable Source of Functional β Cells for Blood Glucose Regulation

Chaiyaboot Ariyachet,[1,2] Alessio Tovaglieri,[3] Guanjue Xiang,[4] Jiaqi Lu,[5] Manasvi S. Shah,[3,6] Camilla A. Richmond,[6,7] Catia Verbeke,[8] Douglas A. Melton,[1,2] Ben Z. Stanger,[9] David Mooney,[8] Ramesh A. Shivdasani,[2,10] Shaun Mahony,[4] Qing Xia,[5] David T. Breault,[2,3,6] and Qiao Zhou[1,2,*]

[1]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
[2]Harvard Stem Cell Institute, Cambridge, MA 02138, USA
[3]Division of Endocrinology, Boston Children's Hospital, Boston, MA 02115, USA
[4]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA
[5]State Key Laboratory of Natural and Biomimetic Drugs, Department of Chemical Biology, Peking University School of Pharmaceutical Sciences, Beijing 100191, China
[6]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA
[7]Division of Gastroenterology, Boston Children's Hospital, Boston, MA 02115, USA
[8]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA
[9]Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, PA 19104, USA
[10]Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA
*Correspondence: qiao_zhou@harvard.edu
http://dx.doi.org/10.1016/j.stem.2016.01.003

SUMMARY

The gastrointestinal (GI) epithelium is a highly regenerative tissue with the potential to provide a renewable source of insulin$^+$ cells after undergoing cellular reprogramming. Here, we show that cells of the antral stomach have a previously unappreciated propensity for conversion into functional insulin-secreting cells. Native antral endocrine cells share a surprising degree of transcriptional similarity with pancreatic β cells, and expression of β cell reprogramming factors in vivo converts antral cells efficiently into insulin$^+$ cells with close molecular and functional similarity to β cells. Induced GI insulin$^+$ cells can suppress hyperglycemia in a diabetic mouse model for at least 6 months and regenerate rapidly after ablation. Reprogramming of antral stomach cells assembled into bioengineered mini-organs in vitro yielded transplantable units that also suppressed hyperglycemia in diabetic mice, highlighting the potential for development of engineered stomach tissues as a renewable source of functional β cells for glycemic control.

INTRODUCTION

Major progress has been made in recent years to produce functional insulin$^+$ cells for cell replacement therapies to treat diabetes. These regenerative technologies include directed differentiation of embryonic stem cells and direct conversion from non-β cells such as liver cells, acinar cells, and others (Hebrok, 2012; Johannesson et al., 2015; Nostro and Keller, 2012; Schiesser and Wells, 2014; Zhou and Melton, 2008). However, because ongoing pathological conditions in diabetes inflict continued damage to native and transplanted β cells (Azzi et al., 2010; Butler et al., 2003; Lakey et al., 2006; Rahier et al., 2008), it is desirable to develop a regenerative system where β cells can be produced in a renewable fashion to counteract β cell loss. The gastrointestinal (GI) tissues are potential sources for such continued generation of β cells. The stomach and intestine are unique among endodermal organs in that they harbor large numbers of adult stem/progenitor cells that constantly produce epithelial cells, including hormone-secreting enteroendocrine cells (Barker et al., 2007, 2010; May and Kaestner, 2010; Schonhoff et al., 2004a). Both organs are developmentally related to the pancreas, arising in adjacent embryonic domains (Offield et al., 1996). Development of gut enteroendocrine and pancreatic endocrine cells also depends on common critical factors, such as Ngn3 (also known as *Neurog3*) (Gu et al., 2002; Jenny et al., 2002; Lee et al., 2002). Recent studies showed that intestinal cells could be converted into insulin$^+$ cells with either endocrine-specific deletion of *FoxO1* or ubiquitous expression of NPM reprogramming factors (*Ngn3*, *Pdx1*, and *Mafa*) (Bouchi et al., 2014; Chen et al., 2014; Talchai et al., 2012a). Although these studies revealed the feasibility of deriving β-like cells from the intestine, critical barriers remain in developing these approaches into future regenerative therapies. *FoxO1* plays a critical role in protecting β cells from cellular stress (Kitamura et al., 2005; Talchai et al., 2012b), and deletion or suppression of *FoxO1* in pancreatic β cells could result in β cell failure (Talchai et al., 2012b; Talchai and Accili, 2015). Moreover, although NPM factors induce insulin$^+$ cells in the intestine, the induced cells appear to lack certain important β cell genes such as Nkx6.1 and exhibit reduced glucose responsiveness compared with pancreatic β cells (Chen et al., 2014).

We sought to devise improved strategies to derive functional insulin-secreting (insulin$^+$) cells from GI tissues and to harness

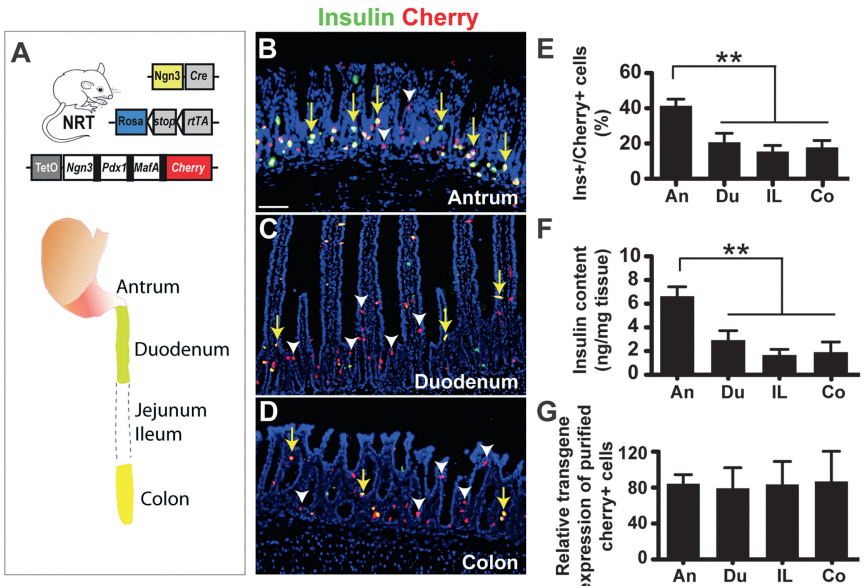


Figure 1. NPM Factors Efficiently Reprogram Gastrointestinal Endocrine Cells to Insulin$^+$ Cells with the Highest Induction Efficiency in Antral Stomach

(A) Diagram of the triple-cross transgenic mouse line, referred to as NRT (Ngn3-Cre; Rosa-floxed-rtTA; Teto-NMPcherry). Ngn3-cre is used to target inducible expression of the NPM factors (Ngn3, Pdx1, and Mafa) into the enteroendocrine cells of the antral stomach and the intestine. Black bars in Teto-NPMcherry indicate 2A peptides used to mediate polycistronic expression.

(B–G) Doxycycline treatment of NRT animals yielded many insulin$^+$cherry$^+$ cells from the antral stomach (B), the duodenum (C), and the colon (D), among other GI regions. Quantitation showed a higher induction efficiency of insulin$^+$ cells in antrum compared with duodenum (du), ileum (IL), and colon (Co) (E, n = 3 animals, p = 0.0026). Antrum tissue also has higher insulin content (F, n = 3 animals, p = 0.0046). Using FACS-purified cherry$^+$ cells, the expression level of transgenes in the endocrine population was found to be comparable (G), n = 3 animals). Scale bar, 100 μm. Yellow arrows indicate insulin$^+$cherry$^+$ cells; white arrowheads indicate insulin$^-$cherry$^+$ cells.

See also Figure S1.

the regenerative capacity of these tissues as a renewable source of β cells. We report the surprising finding that NPM factors reprogram enteroendocrine cells from the antral stomach more efficiently into functional insulin$^+$ cells compared with enteroendocrine cells from the intestine. Induced antral insulin$^+$ cells also express key β cell factors, including Nkx6.1 and Prohormone convertase 2 (PC2), which intestinal insulin$^+$ cells lack. Our data reveal that native antral enteroendocrine cells share a surprising level of transcriptional similarity with pancreatic β cells. Further, the intestine-specific *Cdx2* gene can block efficient β cell reprogramming. Thus, intrinsic molecular differences between antral stomach and intestinal enteroendocrine cells could contribute to the differential reprogramming outcomes. To explore the therapeutic potential of gastric tissue as a source of inducible β cells, we created bioengineered stomach mini-organs; upon transplantation and sphere formation, these structures produced renewable insulin$^+$ cells that reverse hyperglycemia in vivo. Our studies reveal antral stomach tissue as a previously unrecognized source that is highly amenable to reprogramming toward functional insulin$^+$ cells. We also provide proof of principle evidence that bioengineered gastric tissue could serve as a renewable source of β cells for glycemic control.

RESULTS

NPM Factors Efficiently Reprogram GI Enteroendocrine Cells to Insulin$^+$ Cells, with Antral Stomach Showing the Highest Induction Efficiency

Previous studies of reprogramming GI tissues to insulin$^+$ cells have used either deletion of *FoxO1* or expression of NPM factors (*Ngn3*, *Pdx1*, and *Mafa*). Surprisingly, no insulin$^+$ cells were reported from stomach with either approach (Chen et al., 2014; Talchai et al., 2012a). To revisit this important question, we performed additional reprogramming experiments in the GI tract. Using adenoviral infection of cultured mouse antral stomach organoids, we observed that the NPM factors are highly effective at inducing insulin expression whereas the other reprogramming factors tested, including Pax4, Insm1, Nkx6.1, and Mafa, are not effective (Figure S1). Based on this observation, we constructed new transgenic mouse lines (TetO-NPMcherry) in which the inducible TetO promoter drives polycistronic expression of NPM factors and the red fluorescent protein Cherry (Figure 1A). Global expression of NPM factors leads to rapid animal death due to hypoglycemia (unpublished observations). To enable long-term observation and comparison of induced insulin$^+$ cells from different GI regions, we targeted NPM factors to the GI enteroendocrine lineage, which shares molecular and developmental similarity with pancreatic endocrine cells (Habib et al., 2012; May and Kaestner, 2010; Schonhoff et al., 2004a), making it an excellent target for β cell conversion. We crossed the TetO-NPMcherry line with the bacterial artificial chromosome (BAC)-transgenic Ngn3-Cre line (Schonhoff et al., 2004b) and the knockin Rosa-floxed-rtTA line (Jackson Laboratory) to derive a triple-transgenic line we call NRT (Figure 1A). The well-described Ngn3-Cre line labels all enteroendocrine cells in the intestine and the majority of antral stomach enteroendocrine cells (Schonhoff et al., 2004b) (Figure S1).

After doxycycline (Dox) treatment of NRT animals for 10 days, we observed numerous insulin$^+$ cells in the antral stomach and along the entire length of the intestine (Figures 1B–1D, yellow arrows). The fundus region of the stomach contains relatively few Ngn3$^+$ endocrine cells, and very few of these expressed insulin, suggesting that fundal cells resist NPM-mediated conversion (Figure S1). Quantitative analysis showed significantly higher reprogramming efficiency in the antrum (41.5% ± 8.5%, mean ± SD) than in the proximal (duodenum, 21.4% ± 6.7%), middle (ileum, 14.6% ± 3.3%), or distal (colon, 15.5% ± 3.4%) intestine (Figure 1E). The antral stomach also contains substantially higher

levels of insulin protein compared with the intestine (Figure 1F), even though levels of reprogramming factor expression in fluorescence-activated cell sorting (FACS)-purified cherry$^+$ cells from the antrum and different intestinal regions are comparable (Figure 1G).

Enteroendocrine cells in the stomach and intestine include multiple subtypes based on hormone expression (Habib et al., 2012; May and Kaestner, 2010; Schonhoff et al., 2004a). To evaluate whether insulin$^+$ cells are preferentially induced in certain subtypes, we quantified seven major enteroendocrine subtypes before and after induction of insulin$^+$ cells (Figure S1). All endocrine subtypes we examined were reduced upon doxycycline treatment, with the exception of serotonin$^+$ cells, which do not originate from the Ngn3$^+$ lineage (Schonhoff et al., 2004b) (Figure S1). These data indicate that insulin$^+$ cells in both stomach and duodenum arise from multiple endocrine subtypes and/or their common progenitors. We also found the vast majority of induced GI insulin$^+$ cells to be mono-hormonal (Figure S1). These data collectively show that NPM factors can robustly reprogram GI endocrine cells into insulin$^+$ cells, with the highest reprogramming efficiency in the antral stomach.

Induced GI Insulin$^+$ Cells Can Reverse Hyperglycemia Long-Term and Regenerate Rapidly upon Ablation

To test whether the induced GI insulin$^+$ cells can secrete insulin and reverse hyperglycemia, we ablated pancreatic β cells in NRT mice with streptozotocin (STZ), which renders the animals hyperglycemic. Upon Dox treatment and induction of insulin$^+$ cells in the GI tract, hyperglycemia was rapidly reversed and blood glucose levels remained normal for as long as we tracked them (Figure 2A, up to 6 months). In contrast to control animals, which died with hyperglycemia within 8 weeks, nearly every Dox-treated animal was rescued (Figure 2B). Consistent with this effect, intraperitoneal glucose tolerance test (IPGTT) showed substantial improvement after doxycycline induction (Figure 2C) and near-normal blood insulin level in STZ-ablated and Dox-induced animals (Figure 2D).

To confirm that rescue from hyperglycemia results from induction of insulin$^+$ cells in the GI tract, we surveyed insulin expression in other Ngn3-expressing tissues including the brain, testis, and pancreas. No insulin$^+$ cells were found in the brain or testis (data not shown). In the pancreas of NRT animals, STZ treatment led to near complete ablation of endogenous β cells (Figure S2), but Dox treatment induced insulin in glucagon$^+$ cells, which comprise the majority of islet cells after β cell ablation (Figure S2). These glucagon$^+$insulin$^+$ cells do not, however, express other β cell factors such as Glut2 and Nkx6.1, and their insulin expression level is significantly lower than in native β cells (Figure S2). To assess the possibility that these glucagon$^+$insulin$^+$ cells may nevertheless contribute to reversal of hyperglycemia after Dox induction in NRT animals, we resected ~80% of the pancreas and thus most glucagon$^+$insulin$^+$ cells. No significant changes in blood glucose level followed (Figure 2A). The remnant 20% pancreas showed 0.15 ± 0.03 μg of total insulin (mean ± SD), significantly below the insulin content of antrum (1.89 ± 0.36 μg) or duodenum (1.20 ± 0.63 μg; Figure S4). In comparison, a normal mouse pancreas contains ~10 μg insulin, although only a fraction of the β cell mass is required to maintain normoglycemia (Bonner-Weir, 2000). These data collectively indicate that induced insulin$^+$ cells from the GI tract are the main source of secreted insulin that led to long-term reversal of hyperglycemia.

The GI tract is a highly regenerative organ, with resident glandular stem cells continuously producing new epithelial cells (Barker et al., 2010; Barker et al., 2007). To evaluate the capacity of GI β cell regeneration from the stem cell compartment, we conducted a second round of STZ treatment (Figures 2A and 2E). Similar to pancreatic β cells, induced insulin$^+$ cells from the antrum and intestine were sensitive to the toxin and disappeared, leading to hyperglycemia (Figures 2A and 2E). However, the diabetic state was again rapidly reversed, concomitant with the reappearance of GI insulin$^+$ cells (Figure 2E). These data illustrate the high regenerative capacity of the genetically engineered GI tissues and their ability to sustain injuries and maintain suppression of hyperglycemia.

We also evaluated the lifespan of antral and intestinal insulin$^+$ cells and their relative contributions toward glycemic control (Figure S3). In a pulse-chase experiment, GI insulin$^+$ cells were first induced by Dox treatment, followed by Dox withdrawal. Intestinal insulin$^+$ cells disappeared within 7 days, whereas stomach insulin$^+$ cells persisted for more than 20 days, consistent with estimated turnover rates of the native intestinal and antral epithelia (Karam and Leblond, 1993; Lehy and Willems, 1976; Messier and Leblond, 1960; Thompson et al., 1990). Antral insulin$^+$ cells continued to suppress hyperglycemia after intestinal insulin$^+$ cells had disappeared (Figure S3). Thus, antral insulin$^+$ cells have a longer lifespan than their intestinal counterparts and can suppress hyperglycemia independently.

Antral Insulin$^+$ Cells Bear Close Molecular and Functional Resemblance to Pancreatic β Cells

Immunohistochemistry revealed that induced insulin$^+$ cells from the antral stomach and the proximal and distal intestine all express β cell factors such as c-peptide, glucose transporter 2 (Glut2, or Slc2a2), prohormone convertase 1/3 (PC1/3), and Pax6 (Figure 3A, quantification shown in Figure S4). However, other key β cell genes, including Nkx6.1, Nkx2.2, and prohormone convertase 2 (PC2), are expressed exclusively or predominantly in antral insulin$^+$ cells (Figure 3A, quantitation shown in Figure S4). qPCR analysis further confirmed that many β cell factors are expressed at substantially higher levels in antral insulin$^+$ cells than in duodenal or colonic insulin$^+$ cells (Figure S4). Endogenous Pdx1, but not endogenous Mafa, is expressed in the native duodenum and antrum (Figure S4), as previously reported (Habib et al., 2012; Offield et al., 1996). Endogenous Mafa is activated strongly in antral insulin$^+$ cells, but only weakly in duodenal and colonic insulin$^+$ cells (Figure S4), whereas endogenous Pdx1 is induced in both antral and intestinal insulin$^+$ cells (Figure S4). In contrast, endogenous Ngn3 is not induced (Figure S4). We observed continued expression of FoxO1 expression in both antral and intestinal insulin$^+$ cells (Figure S4).

To assess functional properties of induced insulin$^+$ cells from the stomach and the intestine, we harvested whole epithelial tissues from the antrum, duodenum, and colon of NRT animals after 10 days of Dox treatment. In vitro glucose-stimulated insulin secretion assays were performed with each sample at low-glucose (1.7 mM) and high-glucose (20.2 mM) conditions. Our data showed that although all GI insulin$^+$ cells can respond to high glucose (Figure 3B), the responsiveness of antral insulin$^+$

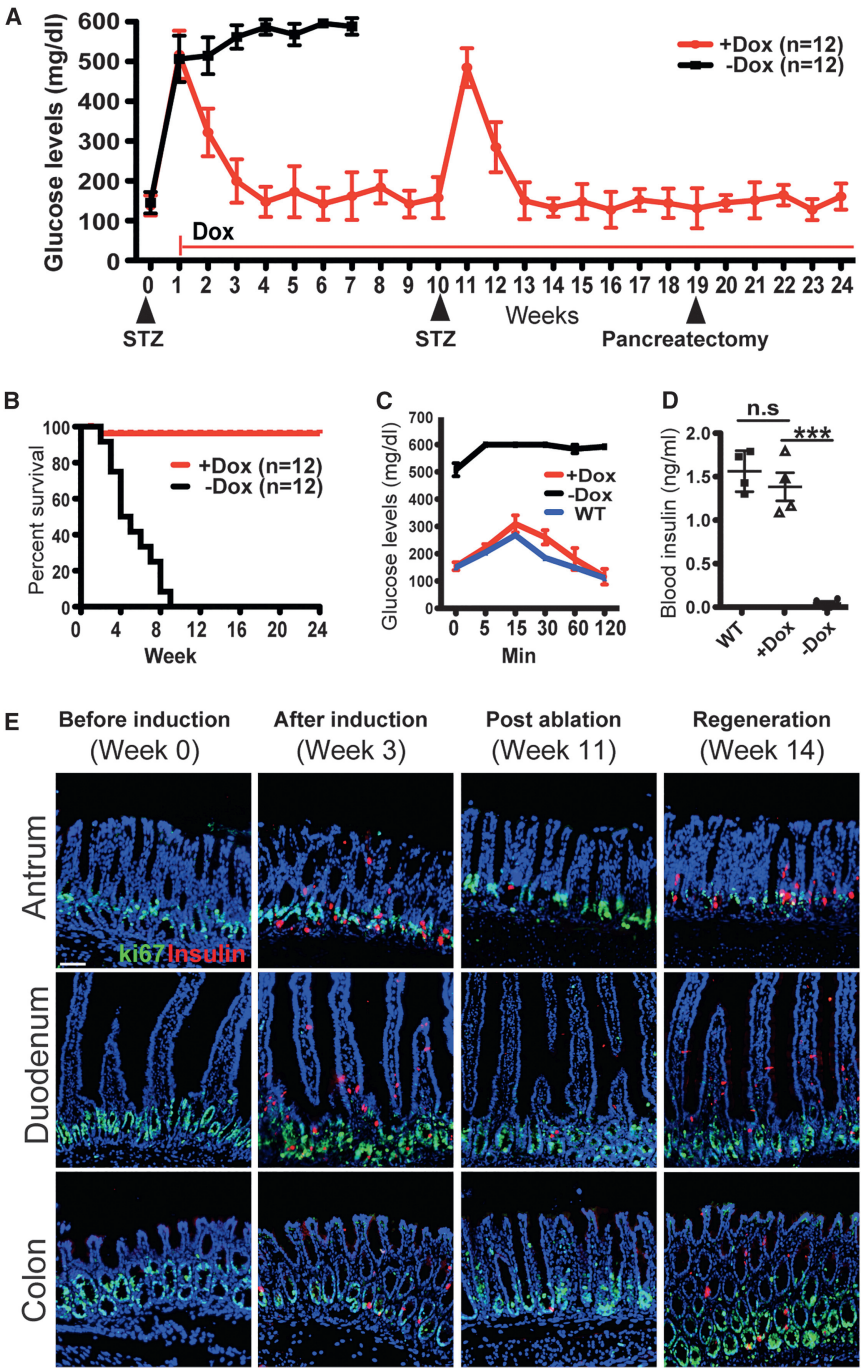


Figure 2. Induced Insulin$^+$ Cells from the GI Tract Can Reverse Hyperglycemia Long-Term and Regenerate Rapidly

(A) Glucose monitoring of hyperglycemic NRT animals over 6 months. Streptozotocin (STZ) was used to ablate endogenous pancreatic β cells and create hyperglycemia. Doxycycline (Dox) was administered continuously from week 1 onward (red line). Compared with persistent hyperglycemia and death of control animals (−Dox group, black squares), Dox treatment led to long-term suppression of hyperglycemia (+Dox group, red circles). A second round of STZ ablation was conducted at week 10 to evaluate the regenerative capacity of this experimental system. The ensuing hyperglycemia was suppressed again by week 13. Pancreatectomy was performed on week 19 to remove ~80% of the pancreas. No significant effect on blood glucose levels was observed.

(B–D) Dox treatment and induction of insulin$^+$ cells led to significant improvement in the survival of hyperglycemic NRT animals (B, n = 12 animals in each group). Glucose tolerance tests showed near-normal responses for Dox-treated animals (C, n = 4 animals in each group). The blood insulin levels of the induced animals are comparable with that of wild-type animals and significantly higher than non-induced animals (D, n = 4 animals in each group, $p < 0.001$).

(E) Immunohistochemistry showed before and after induction of insulin$^+$ cells (first and second panel, respectively). STZ treatment was used at week 11 to ablate the induced insulin$^+$ cells from the GI tract (third panel). Insulin$^+$ cells were regenerated rapidly 3 weeks later (last panel). Ki67 staining labels the proliferating stem/progenitor cell compartment at the base of the glands. Scale bar, 100 μm.

All quantitative data presented as mean ± SD. Statistical significance was evaluated with the Student's t test (***$p < 0.001$). See also Figures S2 and S3.

cells is significantly higher than that of duodenal and colonic insulin$^+$ cells (Figure 3B; data standardized as high-glucose versus low-glucose response ratio: 3.02 ± 0.55 for antrum, 1.65 ± 0.37 for duodenum, and 1.61 ± 0.46 for colon).

To further evaluate the function of induced GI insulin$^+$ cells, we tested their physiological response to glibenclamide (Glib), an anti-diabetic drug that binds to Sur1 and inhibits the ATP-sensitive potassium channel in β cells. Glib treatment led to insulin release from antral, but not from duodenal or colonic, insulin$^+$ cells (Figure 3C). Conversely, treatment with Diazoxide (Dzx), a potassium channel activator, suppressed insulin release from antral insulin$^+$ cells, whereas duodenal and colonic cells showed no response (Figure 3C). Moreover, antral insulin$^+$ cells responded to exendin-4, an antidiabetic drug that activates glucagon-like-peptide receptor (Glp1R), leading to increased insulin release at high glucose concentrations (Figure 3D), whereas duodenal and colonic insulin$^+$ did not respond (Figure 3D). Consistent with these physiological data, antral insulin$^+$ cells express significantly higher levels of Sur1 and Glp1R, compared with duodenal and colonic insulin$^+$ cells (Figure S4).

Thus, molecular and physiological studies together indicate that antral endocrine cells can be reprogrammed efficiently into insulin$^+$ cells that resemble pancreatic β cells, whereas conversion from intestinal endocrine cells is comparatively incomplete.

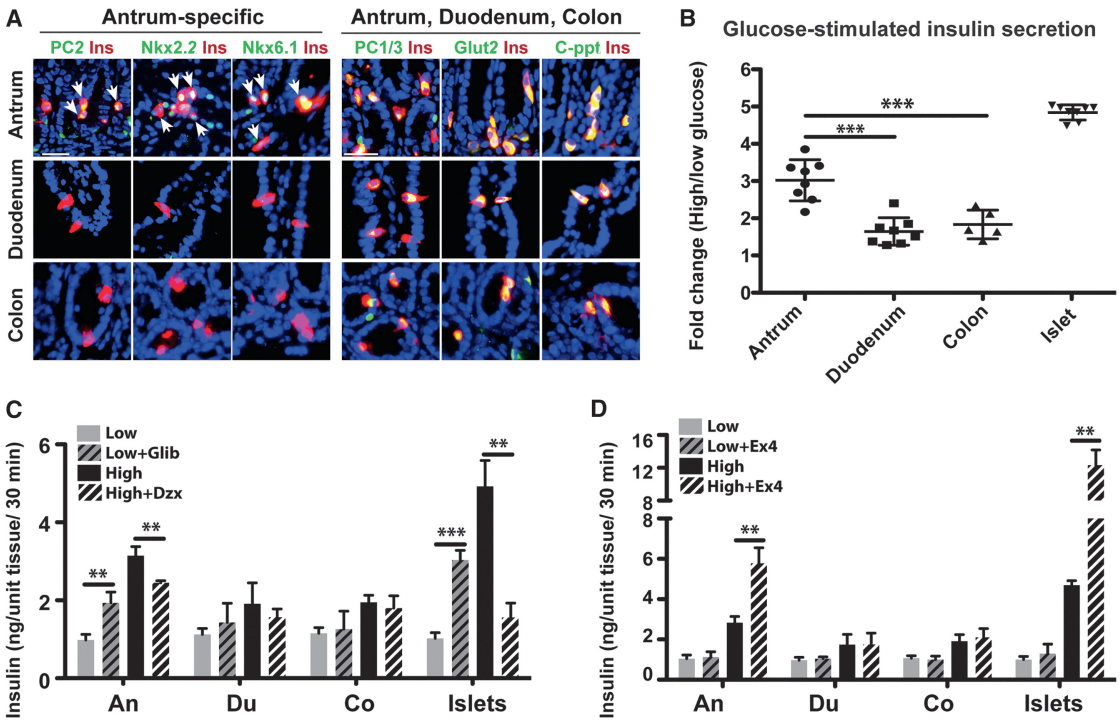


Figure 3. Induced Insulin$^+$ Cells from the Antral Stomach More Closely Resemble β Cells Molecularly and Functionally

(A) Immunohistochemistry showed that induced insulin$^+$ cells from the antrum express β cell genes Nkx6.1, Nkx2.2, and Prohormone convertase 2 (PC2), which are largely absent from duodenum and colon insulin$^+$ cells. In contrast, Prohormone convertase 1/3 (PC1/3), glucose transporter 2 (Glut2), and c-peptide (c-ppt) are expressed commonly in antral, duodenal, and colonic insulin$^+$ cells. Arrows indicate antral insulin$^+$ that are PC2$^+$, Nkx2.2$^+$, and Nkx6.1$^+$. Scale bar, 50 μm.

(B) Glucose stimulated insulin secretion (GSIS) in vitro. Antral tissues have significantly higher glucose responsiveness, defined as fold increase of insulin release at high versus low glucose conditions, compared with duodenal and colonic tissues (n = 8, $p < 0.001$).

(C and D) The antidiabetic drug Glibenclamide (Glib) stimulated insulin release from the antral insulin$^+$ cells whereas Diazoxide (Dzx), a suppressor of insulin release, reduced antral insulin secretion (C, n = 4). In contrast, duodenal and colonic insulin$^+$ cells do not respond to Glib or Dzx (C). Antral insulin$^+$ cells also respond to Exendin-4 (Ex4) with enhanced insulin secretion at high glucose levels whereas duodenal and colonic cells do not respond to Ex4 (D, n = 4).

All quantitative data presented as mean ± SD. Statistical significance was evaluated with the Student's t test (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$). See also Figure S4.

Native Antral Endocrine Cells Share Substantial Transcriptional Similarity with Pancreatic β Cells

What mechanisms might underlie the significant difference? One long-standing hypothesis postulates that the more transcriptional and epigenetic similarities two cells share, the easier it is to interconvert them (Graf and Enver, 2009; Gurdon and Melton, 2008). Transcriptional studies of specific intestinal endocrine populations have been reported (Egerod et al., 2012; Habib et al., 2012), but transcriptomes of antral endocrine cells remain uncharacterized. We therefore profiled the transcriptomes of enteroendocrine cells from the antrum, duodenum, and colon and assessed their similarity to pancreatic β cells. We used Ngn3-GFP reporter mice to isolate enteroendocrine cells from the different GI regions (Lee et al., 2002); Ngn3 expression in the gut is transient and restricted to endocrine progenitors (Jenny et al., 2002; Lee et al., 2002). Ngn3-GFP labels a mixture of chromogranin$^-$ (Chga$^-$) and chromogranin$^+$ (Chga$^+$) cells, representing immature and mature endocrine cells, respectively (Lee et al., 2002) (Figure 4A). Our quantitation showed that the relative proportions of GFP$^+$Chga$^-$ and GFP$^+$Chga$^+$ cells are comparable in antral stomach, duodenum, and colon (Figure S5). GFP$^+$ cells purified by FACS from the different GI regions (Figure 4B) constitute ~1%–2% of the total cell population (Figure 4B), consistent with the estimated prevalence of gut endocrine cells (Schonhoff et al., 2004a).

We generated global transcriptome data from the purified cells with Illumina arrays. Comparative analyses showed that endocrine cells from the proximal and distal intestine are more similar to each other and less similar to antral endocrine cells (Figure S5, Spearman correlation coefficients: 0.91 [duodenum versus colon], 0.82 [antrum versus duodenum], and 0.80 [antrum versus colon]). The overall similarity of proximal and distal intestine endocrine cells is high and consistent with published studies (Egerod et al., 2012; Habib et al., 2012) (1,470 differentially expressed genes listed in Table S3). We performed pairwise comparison of the three GI endocrine populations with our published transcriptome data of β cells, which was obtained by FACS purification from the islets of MIP-GFP animals (Li et al., 2014b). This analysis showed overall higher transcriptional similarity between antral and β cells than between intestinal and β cells (Figures 4C and S5, Spearman correlation coefficients: 0.72 [antrum versus β], 0.57 [duodenum versus β], and 0.57 [colon versus β]; Steiger's *Z*-test for dependent correlations: $p = 6.5 \times 10^{-185}$). Thus, although enteroendocrine cells from the

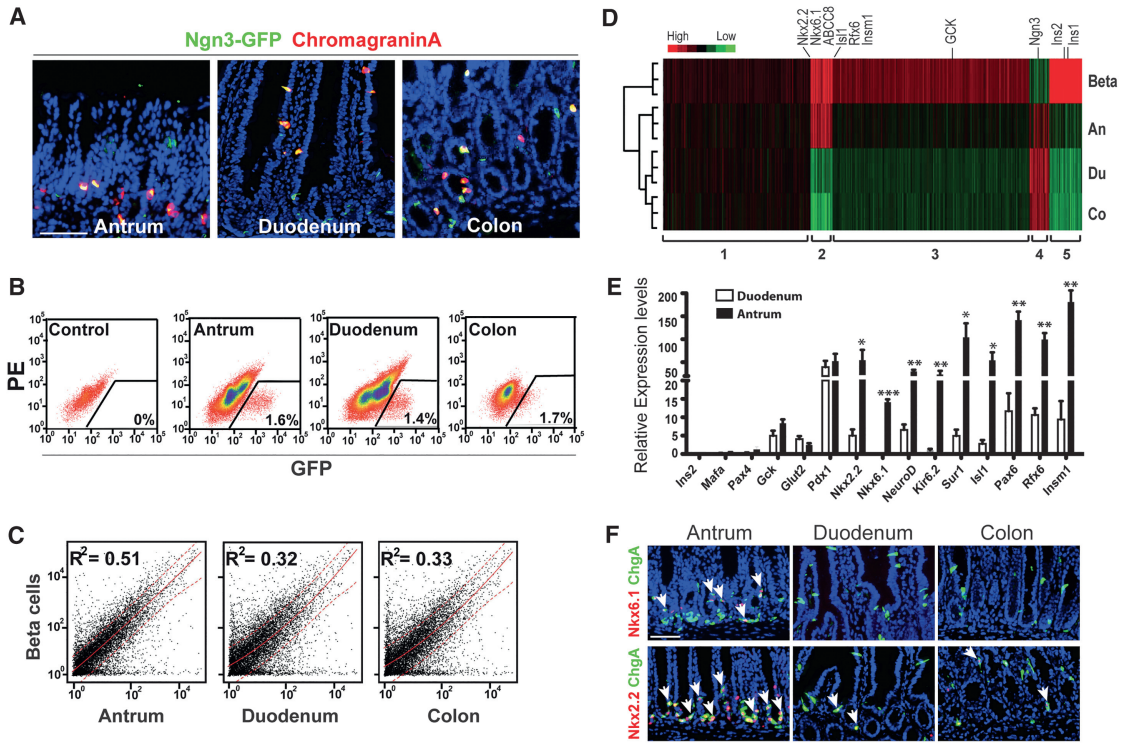


Figure 4. Enteroendocrine Cells of the Antral Stomach Share Substantial Transcriptional Similarity with Pancreatic β Cells

(A) Immunohistochemistry showing distribution of GFP$^+$ in the GI tract of the Ngn3-GFP mouse line. The GFP$^+$ cells include both relatively immature (GFP$^+$ Chromogranin$^-$) and more mature enteroendocrine cells (GFP$^+$Chromogranin$^+$). Scale bar, 50 μm.

(B and C) Ngn3-GFP$^+$ cells were purified by FACS from antrum, duodenum, and colon (B). Scatterplots of transcriptome comparisons between pancreatic β cells and the GI enteroendocrine populations (C). Antral enteroendocrine cells show a greater similarity with β cells.

(D and E) Analysis of 2,398 β cell-enriched genes showed a general trend of elevated expression in antral enteroendocrine cells compared with duodenal and colonic enteroendocrine cells (D). In particular, antral enteroendocrine cells share a group of genes (group 2) with β cells (D) that are enriched for factors important in β cell development and function (E). Quantitative data presented as mean ± SD. Statistical significance was evaluated with the Student's t test (*$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

(F) Immunohistochemistry showed that Nkx6.1 is present in a population of Chga$^+$ enteroendocrine cells in the antrum, but not expressed in duodenum or colon (top, arrows). Nkx2.2 is expressed in a majority of Chga$^+$ enteroendocrine cells in the antrum and a minority of Chga$^+$ cells in the duodenum and colon (F, bottom, arrows). Scale bar, 50 μm.

See also Figure S5.

antrum, duodenum, and colon are more similar to each other than they are to pancreatic β cells, β cells appear to share more transcriptional similarity with antral enteroendocrine cells than intestinal enteroendocrine cells.

To evaluate the expression of β cell-enriched genes in GI enteroendocrine cells, we focused analysis on a collection of 2,398 genes that show higher expression in β cells than in acinar cells (Li et al., 2014b). Antral enteroendocrine cells showed higher expression of many β cell-enriched genes (Figure 4D, group 2 and 3 genes; Table S2) compared with intestinal enteroendocrine cells. In particular, many genes critical for β cell development and function, such as Nkx6.1, Nkx2.2, NeuroD1, Isl1, Rfx6, Insm1, Sur1 (ABCC8), and Glucokinase (GCK), are enriched in antral, compared with duodenal or colonic, enteroendocrine cells (Figures 4E and S5). Immunohistochemistry showed Nkx6.1 expression in a subset of antral GFP$^+$ cells (24.9% ± 3.5%, mean ± SD), but not in the duodenum or colon (Figure 4F). The vast majority of antral Nkx6.1$^+$ cells also express Chga (94.7% ± 3.1%) (Figures 4F and S5). Nkx2.2 is expressed in a majority of GFP$^+$ cells in the antrum (57.2% ± 4.7%), but only in a minority of duodenal (18.7% ± 2.7%) or colonic (23.4% ± 4.3%) Chga$^+$ cells (Figure 4F). Most Nkx2.2$^+$ cells express Chga (69.4% ± 3.4%, 64.0% ± 5.1%, and 65.5% ± 8.3% in antrum, duodenum, and colon, respectively) (Figure 4F; Figure S5). Gene Ontology analyses show that whereas enteroendocrine cells from all GI regions are enriched for pathways involved in regulation of hormone secretion, G-protein-coupled receptor signaling, and vesicle-mediated transport, antral enteroendocrine cells are enriched specifically for the "glucose homeostasis" module (Figure S5; Table S2). Together, these studies reveal a surprising intrinsic difference between endocrine cell populations from the antral stomach and intestine, which likely contributes to their differential capacity for β cell reprogramming.

The Intestine-Specific Gene *Cdx2* Can Inhibit β Cell Conversion

In a prior study of acinar to β cell conversion, we showed that persistent expression of acinar cell fate regulators Ptf1a and

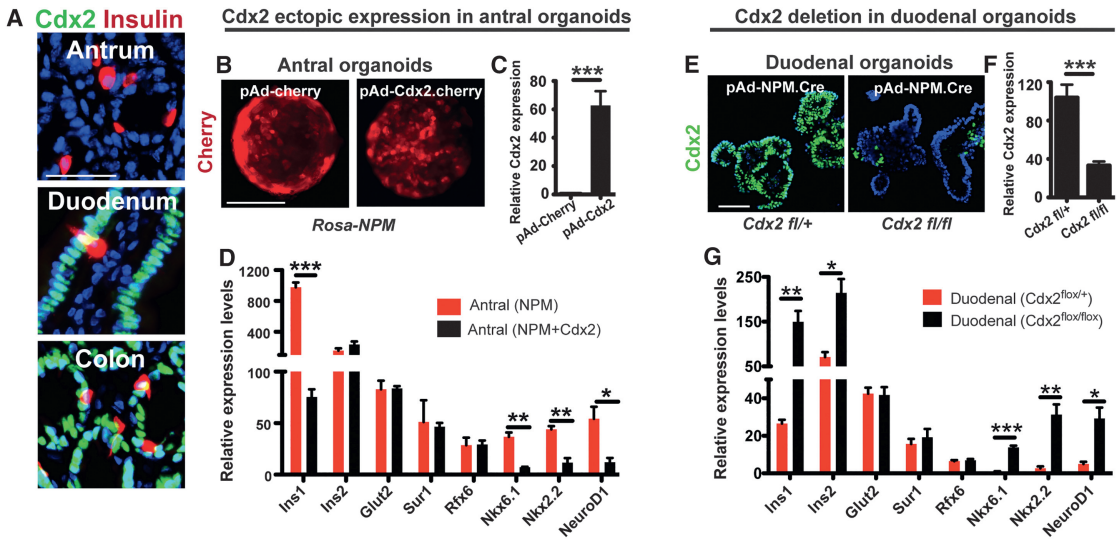


Figure 5. The Intestine-Specific Cell Fate Regulator *Cdx2* Can Inhibit β Cell Conversion

(A) Duodenal and colonic insulin$^+$ cells express Cdx2, the master regulator of intestine cell fate whereas antral stomach cells do not express Cdx2 before or after induction in NRT animals. Scale bar, 50 μm.

(B–D) Epithelial organoids were established from antral tissues of double-transgenic Rosa-rtTA;Teto-NPMcherry (Rosa-NPM) animals and infected with either control adenovirus expressing Cherry (pAd-cherry) or adenovirus expressing Cdx2 and Cherry (pAd-Cdx2.cherry) (B and C). Dox treatment was subsequently used to induce β cell conversion in these antral organoids. qPCR analysis showed that ectopic Cdx2 suppressed the expression of multiple β cell genes (D, n = 3). Scale bar, 100 μm.

(E–G) Duodenal organoids were established from Cdx2$^{fl/+}$ and Cdx2$^{fl/fl}$ animals (E, left and right, respectively). Infection with an adenovirus co-expressing both NPM factors and the Cre recombinase led to simultaneous deletion of floxed Cdx2 allele and expression of NPM factors (E). Complete removal of Cdx2 was observed in majority of Cdx2$^{fl/fl}$ duodenal cells by immunohistochemistry and qPCR analysis (E and F) and led to enhanced expression of multiple β cell genes from the duodenal organoids (G, n = 3). Scale bar, 100 μm.

Quantitative data presented as mean ± SD. Statistical significance was evaluated with the Student's t test (*p < 0.05, **p < 0.01, and ***p < 0.001). See also Figure S6.

Nr5a2 blocks acquisition of β cell fate (Li et al., 2014c). *Cdx2* is an intestine-specific master regulator gene (Gao et al., 2009), and its persistent expression in intestinal insulin$^+$ cells (Figure 5A) raises the question of whether Cdx2 might block intestinal cells from adopting more complete β cell features. To test this hypothesis, we generated epithelial organoids from the antrum and duodenum of double transgenic Rosa-rtTA;TetO-NPMcherry (Rosa-NPM) animals and treated them with Dox in culture. Similar to our observations in vivo, antral organoids produced more C-peptide$^+$ cells with higher levels of β cell factors compared with intestinal organoids (Figure S6). Next, we expressed either the control cherry gene or *Cdx2* using adenoviral infection in the double-transgenic antral organoids (Figures 5B and 5C), followed by treatment with Dox to activate β cell conversion. Cdx2 significantly suppressed expression of multiple β cell genes, including NeuroD1, Nkx2.2, and Nkx6.1 (Figure 5D).

To further evaluate the role of Cdx2 in intestine reprogramming, we deleted *Cdx2* from duodenal organoids. We established duodenal organoids from animals where a single allele or both alleles of the *Cdx2* gene are floxed (Figure 5E, *Cdx2*$^{fl/+}$ and *Cdx2*$^{fl/fl}$). Infection with an polycistronic adenovirus expressing NPM factors and the Cre recombinase (pAd-NPM.Cre) led to simultaneous removal of the floxed *Cdx2* allele(s) and expression of NPM factors (Figure 5E). Immunohistochemistry and qPCR confirmed complete removal of *Cdx2* from the majority of *Cdx2*$^{fl/fl}$ duodenal cells (Figures 5E and 5F). Cdx2 deletion significantly enhanced expression of several β cells genes, including Insulin1, insulin2, Nkx6.1, and NeuroD. These data together suggest that Cdx2 acts as a molecular barrier to β cell conversion; thus, failure to downregulate Cdx2 in intestinal insulin$^+$ cells likely contributes to their incomplete acquisition of β cell properties.

Constructing Bioengineered Stomach and Intestine Mini-organs to Produce Insulin$^+$ Cells

Among GI tissues, antral stomach is a superior source of functional β cells by NPM-mediated conversion, and antral insulin$^+$ cells are rapidly replenished from the native stem cell compartment. However, inducing β cells from the native GI tract in situ may have limitations in therapy, because the native endocrine populations regulate many physiological processes (Field et al., 2010; May and Kaestner, 2010; Schonhoff et al., 2004a), and diverting them into β cells may disrupt normal endocrine homeostasis. Moreover, induced β cells positioned along the native GI epithelium may inadvertently respond to dietary as well as blood glucose. To circumvent these potential barriers to therapeutic application, we studied the feasibility of constructing "stomach mini-organs" that contain genetically engineered antral tissues as a reservoir of new β cells.

Following published protocols on bioengineering stomach (Maemura et al., 2004; Speer et al., 2011), we embedded gastric gland units from the antrum of CAGrtTA::TetO-NPMcherry (CAG-NPM) animals in Matrigel, loaded them onto

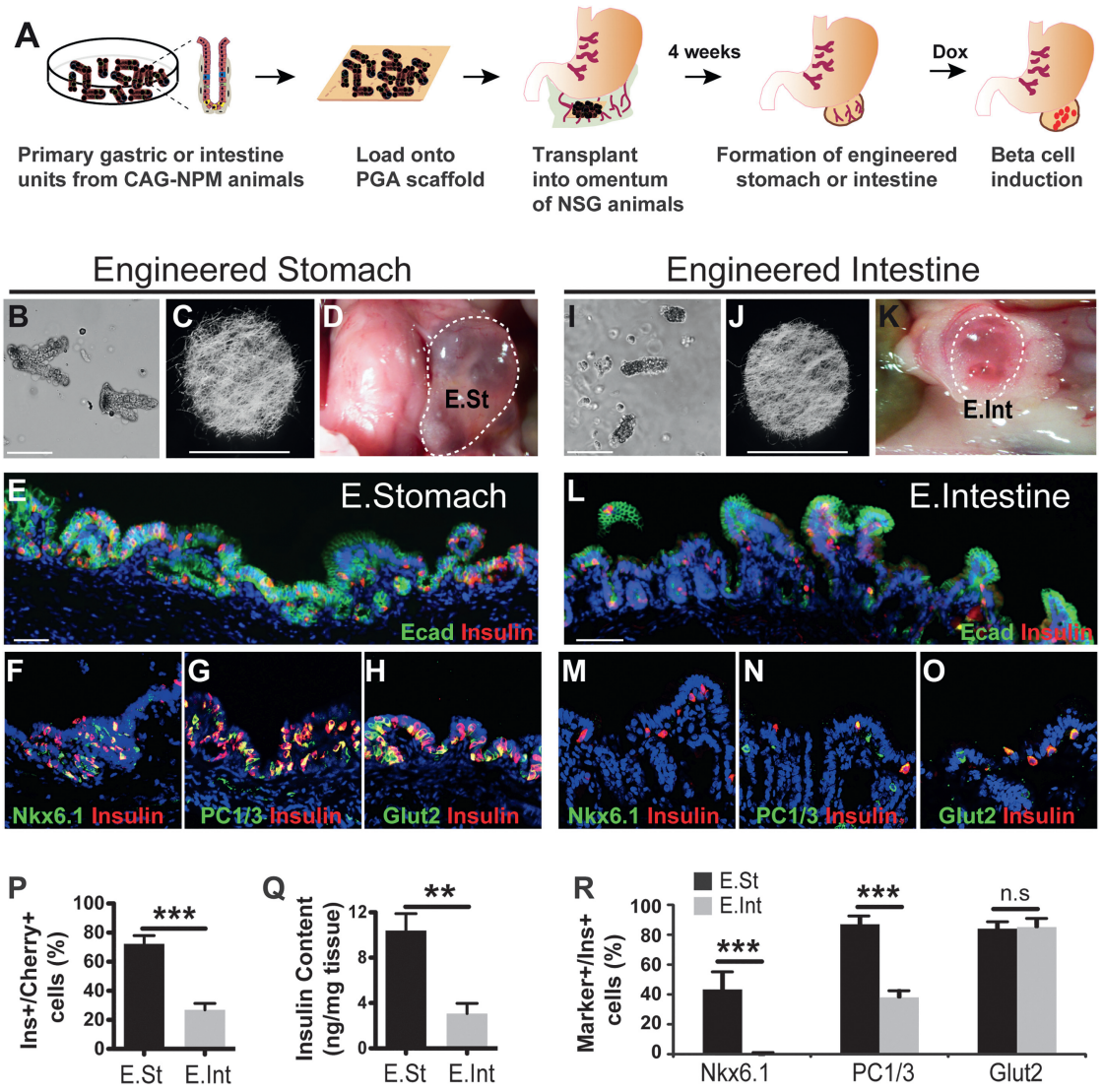


Figure 6. Construction of Bioengineered Stomach and Intestine Mini-organs to Produce Insulin$^+$ Cells

(A–K) Schematic diagram of engineering stomach and intestine mini-organs (A). Gastric or intestinal units were isolated from the antrum or duodenum of CAG-NPM (Cag-rtTA::TetO-NPMcherry) animals (B and I) and loaded onto polyglycolic acid scaffolds (C and J). The scaffolds were placed inside the omental flap of recipient immune-deficient NSG animals. 4 weeks later, an engineered stomach (E. St) or intestine (E. Int) sphere formed (D and K, circled tissue). Scale bars represent 400 μm (B and I) and 6 mm (C and J).

(E–R) In engineered stomach and intestine spheres where reconstitution of epithelium was successful, Dox treatment led to induction of many insulin$^+$ cells (E and L). The induction efficiency is higher for stomach tissues (P, n = 3). Stomach tissues also have higher insulin content (Q, n = 3). The majority of insulin$^+$ cells from engineered stomach express Nkx6.1, PC1/3, and Glut2 (F, G, H, and R) whereas insulin$^+$ cells from engineered intestine lack Nkx6.1 and have reduced PC1/3 expression (M, N, O, and R). Quantitation presented as mean ± SD. Statistical significance was evaluated with the Student's t test (**$p < 0.01$ and ***$p < 0.001$). See also Figure S7.

poly(glycolic acid) (PGA) scaffolds, and transplanted the material into the omental flap of immunodeficient NSG recipient animals (Figures 6A–6C). Four weeks later, bioengineered stomach spheres measuring 0.5 to 1 cm in diameter formed outside the native gut (Figure 6D). By histology, 5 out of 15 such spheres showed robust epithelial reconstitution, while the others showed little or no epithelium (Figure S7). Antral glands in the native stomach are composed largely of mucous and endocrine cells and lack acid-secreting parietal cells. The engineered stomachs also showed a simple organization, with one or several layers of Ecadherin$^+$ cells surrounded by connective tissue (Figure S7). The epithelial component contained Sox9$^+$ stem/progenitor cells (Furuyama et al., 2011), Mucin5$^+$ secretory cells, and Chga$^+$ endocrine cells (Figure S7). In parallel, we used a similar bioengineering approach to construct "intestine mini-organs" using duodenal gland units. The success rate for epithelial reconstitution was lower in intestinal spheres (3 out of 15), which contained Muc2$^+$ secretory cells and Chga$^+$ endocrine cells, similar to the native duodenal epithelium. Our observations are consistent with other published studies on

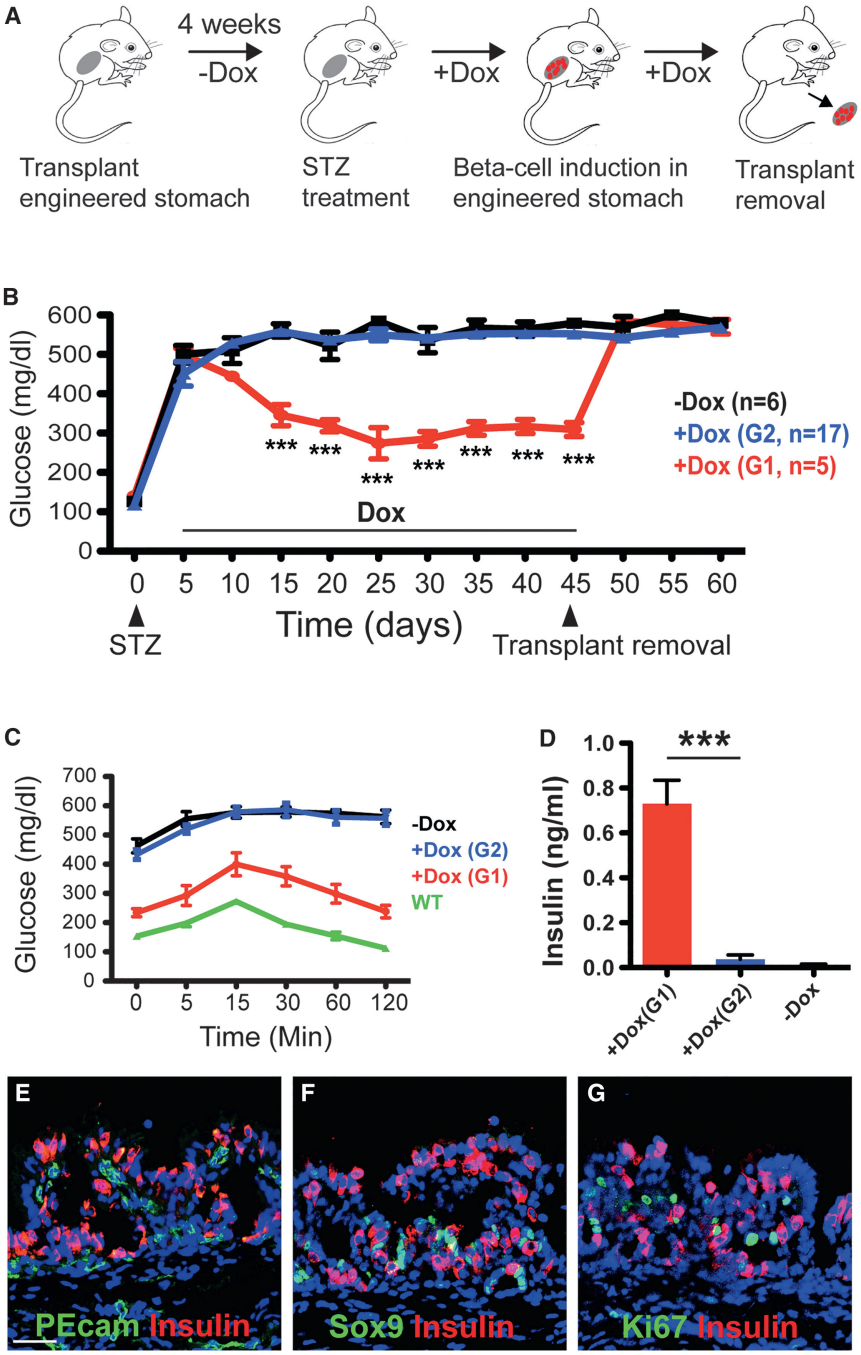


Figure 7. Transplanted Stomach Mini-organs Can Reverse Hyperglycemia in Diabetic Mice

(A) Diagram of the experimental design. STZ treatment was used to ablate endogenous β cells in NSG animals transplanted with 4-week-old stomach spheres, followed by continuous Dox treatment of induce insulin$^+$ cells. At the end of the experiment, the engineered stomachs were removed surgically.

(B–D) STZ treatment led to rapid hyperglycemia that persists in the absence of treatment (−Dox group, n = 6, black squares) (B). After Dox treatment, a group of five animals showed prolonged suppression of hyperglycemia (G1 animals, n = 5, red squares), whereas another group of animals remained hyperglycemic (G2 animals, n = 17, blue squares) (B). After 6 weeks, the engineered stomach spheres were removed from the G1 animals, which led to their reversal back to hyperglycemia (B). G1 animals showed improved response in glucose tolerance test (C, n = 4) and substantially higher blood insulin levels (D, n = 4) compared with G2 animals or control STZ-treated animals without Dox induction. Wilde-type control animals in (C) (green squares) are non-STZ-treated animals with intact pancreatic β cell mass. Quantitative data presented as mean ± SEM. Statistical significance was evaluated with the Student's t test (***p < 0.001).

(E–G) Sox9$^+$ and Ki67$^+$ cells are present in the engineered stomach after 4-week Dox treatment (F and G), indicating persistence of stem/progenitor cells. PEcam$^+$ blood vessels are closely associated with insulin$^+$ cells inside the engineered stomach sphere (E). Scale bars, 100 μm. Blue channel, DAPI.

See also Figure S7.

bioengineered stomach and intestine (Maemura et al., 2004; Speer et al., 2011).

To evaluate induction of insulin$^+$ cell in the engineered stomach and intestine spheres, we administered Dox for 2 weeks. Many insulin$^+$ cells appeared in the epithelial layer of stomach as well as intestinal spheres. The stomach spheres had significantly more insulin$^+$ cells, higher reprogramming efficiency, and higher insulin content per milligram of tissue (Figures 6E, 6L, 6P, and 6Q). The majority of stomach insulin$^+$ cells express Nkx6.1, Glut2, and PC1/3, whereas intestine insulin$^+$ cells lack Nkx6.1 and have reduced PC1/3 expression (Figures 6F–6H, 6M–6O, and 6R).

Transplanted Stomach Mini-organs Can Control Hyperglycemia in Diabetic Mice

To assess if β cells induced in the engineered stomachs could release functional insulin, we ablated pancreatic β cells in transplanted animals using STZ and then induced insulin$^+$ cells in the engineered stomach spheres by administering Dox (Figure 7A). Of the 22 treated animals, 5 showed sustained decreases in blood glucose levels after Dox treatment (group 1), whereas the others remained hyperglycemic (group 2) (Figure 7B). We monitored animals for 6 weeks and subsequently removed the grafted stomach spheres from G1 mice, which restored hyperglycemia (Figure 7B). Engineered stomach spheres from G1 animals showed good epithelial structures containing many insulin$^+$ cells (Figure S7), whereas spheres from the G2 groups showed limited epithelial structures with few insulin$^+$ cells (Figure S7). Consistent with the glucose monitoring data and histology, G1 animals showed improved responses to

intraperitoneal glucose challenge (Figure 7C). Blood insulin levels in G1 animals also were substantially higher than in G2 animals (Figure 7D).

Immunohistochemistry revealed PEcam$^+$ blood vessels closely associated with insulin$^+$ cells in engineered stomach spheres (Figure 7E), consistent with previous observations that induced β cells, similar to endogenous β cells, can secrete VEGF and remodel local vasculature (Zhou et al., 2008). Moreover, large numbers of Sox9$^+$ stem/progenitor cells and Ki67$^+$ proliferating epithelial cells are present in the engineered stomach spheres before and after Dox treatment, indicating persistence of a stem/progenitor compartment (Figures 7F and 7G). These studies collectively indicate that induced insulin$^+$ cells from the bioengineered stomach spheres can release insulin into the circulation and regulate blood glucose levels.

DISCUSSION

The GI tract is a highly regenerative endodermal organ. We sought to harness this regenerative capacity to create a renewable source of functional insulin$^+$ cells by NPM-mediated reprogramming. Our data show that antral stomach enteroendocrine cells are converted to insulin$^+$ cells more efficiently than intestinal enteroendocrine cells and possess molecular and functional hallmarks of pancreatic β cells. Thus, the antral stomach is a surprisingly good source for reprogrammed insulin$^+$ cells, and we demonstrate the application of bioengineered stomach spheres to control blood glucose levels.

Expression of NPM factors previously led to formation of insulin$^+$ cells in the intestine (Chen et al., 2014). Our experimental system is similar to this previous report and confirms induction of insulin$^+$ cells in the intestine with incomplete β cell conversion. In contrast, antral stomach endocrine cells are more fully reprogrammed, with robust expression of key β cell genes and substantially improved glucose responsiveness. Our studies suggest that the difference can be attributed, at least in part, to intrinsic molecular differences between antral and intestinal enteroendocrine cells. Higher levels of β cell fate regulators in antral enteroendocrine cells may facilitate their conversion, whereas *Cdx2*, which is specifically expressed in all intestinal, but not stomach, cells inhibits conversion. It is notable that Cdx2 expression persists in induced insulin$^+$ intestinal cells. Prior studies have shown that ectopic Cdx2 expression in stomach promotes an intestine fate (Silberg et al., 2002; Verzi et al., 2013), whereas Cdx2 loss in cultured intestinal organoids activates antral differentiation (Simmini et al., 2014). Continued expression of Cdx2 in intestinal insulin$^+$ cells may thus present a molecular barrier for complete reprogramming.

Compared with the gastric antrum, the gastric corpus contains few Ngn3-derived enteroendocrine cells, and few such cells expressed insulin after NPM induction. Global expression of NPM factors also induced few insulin$^+$ cells in the fundus (Figure S1). Thus, gastric corpus endocrine cells, which are distinct from those in the antrum or intestine (Choi et al., 2014; Li et al., 2014a) and mainly derive from Ngn3-independent lineages (Li et al., 2014a; Schonhoff et al., 2004b), are not amenable to NPM-mediated β cell conversion. What might account for this resistance? The antral stomach shares a close developmental origin with the pancreas, with both organs arising from a common Pdx1$^+$ endodermal domain during embryogenesis (Wells and Melton, 1999). Therefore, we speculate that the epigenetic landscape of endocrine cells from the fundus is more distinct than those from the antrum, making them harder to convert into β cells. Future studies will be necessary to understand these regional distinctions.

FoxO1 deletion also leads to formation of insulin$^+$ cells in the intestine, suggesting a therapeutic path toward inducing insulin$^+$ cells in situ (Bouchi et al., 2014; Talchai et al., 2012a). Our approach offers several advantages. First, with our method, induced insulin$^+$ cells preserve FoxO1 function, which is known to protect β cells from physiologic stress (Kitamura et al., 2005; Talchai et al., 2012b). Second, with bioengineered stomach spheres, native endocrine cell populations in the gut remain undisturbed, and their functions in physiology are preserved. Third, by separating engineered stomachs from the native organ, induced β cells can be positioned to respond only to changes in blood and not luminal glucose levels.

In summary, our study offers a new approach to harness the intrinsic regenerative capacity of the stomach epithelium to replenishing β cell mass in vivo. Given ongoing pathological insults that continuously erode native or transplanted β cells in diabetes, long-term treatment may require repeated transplants. The regenerative system we propose could eliminate that need, and the number and size of transplanted stomach spheres could be manipulated to control β cell numbers. Coupled with recent progress in genome engineering and the ready access to human gastric epithelium from biopsies and differentiated induced pluripotent stem cells or embryonic stem cells (McCracken et al., 2014), the therapeutic applications of this approach are considerable.

EXPERIMENTAL PROCEDURES

Mouse Strains

R26-floxed-rtTA, R26-floxed-GFP, R26-rtTA, CAG-rtTA, Ngn3-Cre, and NSG mouse strains were obtained from The Jackson Laboratory. Ngn3-GFP (Gu et al., 2002; Jenny et al., 2002; Lee et al., 2002) and Cdx2$^{fl/fl}$ (Silberg et al., 2002; Verzi et al., 2013) mice have been described previously. The TetO-NPMcherry mouse lines were made by standard pronuclear injection at the Harvard Genomic Modification Facility (Cambridge, MA). All animal experiments are approved by the Harvard Institutional Animal Care and Use Committee.

Antibodies and Immunofluorescence

Tissues were processed as previously described (Li et al., 2014b). Primary antibodies are listed in Supplemental Experimental Procedures. Pictures were taken with a Zeiss LSM 510 META confocal microscope. For quantification of marker$^+$ cells such as insulin$^+$ cells, a total of at least 1,000 Marker$^+$ were analyzed from tissues harvested from three different animals. Typically, at least ten randomly selected sections were counted per animal.

FACS Isolation of GFP$^+$ Cells and Gene Profiling

To dissociate tissues into single cells for FACS purification, we modified published protocols (Habib et al., 2012; Reimann et al., 2008; Talchai et al., 2012a). See Supplemental Experimental Procedures for details. GFP$^+$ cells were isolated by FACS with FACSaria (BD Bioscience). Transcriptome data were generated with Illumina microarrays (Sentrix BeadChip MouseRef-8 v2 Arrays) that contain probes for ~19,000 genes. Data analysis, including statistical methods used, is described in Supplemental Experimental Procedures.

Glucose Stimulated Insulin Secretion Assay and Drug Treatments

Isolation of gastrointestinal epithelial cells was carried out as previously described with slight modifications (Habib et al., 2012; Reimann et al., 2008;

Talchai et al., 2012a). Briefly, antrum tissue was cut into small pieces and incubated in 10 mM EDTA for 30 min, followed by mechanical dissociation to extract cell clusters. To collect duodenal samples, the duodenum tissue was first cut into small pieces, incubated in 10 mM EDTA for 5 min, and mechanically stripped to release the villi. Each cell fraction was incubated with Krebs Ringer buffer supplemented with 1.7 mM or 20.2 mM glucose, glibenclamide (10 nM; Tocris), diazoxide (0.5 mM; Sigma), or exendin-4 (100 nM; Sigma). Released insulin amounts were determined by ELISA (Alpco) at the Joslin Specialized Assay core (Joslin Diabetes Center, Boston). Data from different tissue sources were standardized according to basal insulin release level. Each standardized sample is equivalent to 1/15 of antrum or 5 mm of duodenum, or 10 mm of colon in a single animal.

Physiological Studies

Diabetic animals were produced with intraperitoneal injection of STZ (150–170 mg/kg) in 6- to 8-week-old animals. Animals that displayed >400 mg dl^{-1} blood glucose levels for 2 consecutive days after STZ administration were used for experiments. Blood glucose was measured with an Ascensia Elite glucometer (Bayer). For blood glucose monitoring, a short 2-hr fasting precedes glucose measurements. Glucose tolerance test was performed with intraperitoneal injection of 1 mg/g body weight of glucose, preceded by 4-hr fasting. Blood insulin was collected from tail vein blood sampling and measured by ELISA. Tissue insulin was extracted with acid-ethanol solutions and measured by ELISA.

Organoid Culture and Adenoviral Infection

Antral stomach and duodenal organoids were derived from young adult mice (1–2 months) and cultured using standard growth media, essentially as described (Barker et al., 2010). To test β cell reprogramming factors, 4- to 7-day cultures of antral or duodenal organoids were recovered from the Matrigel with the recover solution (Corning), infected with purified adenovirus at 4×10^7 plaque-forming units (PFU) in 100 μl medium for 1 hr at 37°C and re-embedded in Matrigel. qPCR analysis was performed 6–8 days after infection.

Generation of Bioengineered Stomach and Intestine

Generation of bioengineered stomach and stomach spheres was performed essentially as described previously (Maemura et al., 2004; Speer et al., 2011). See Supplemental Experimental Procedures for further details.

ACCESSION NUMBERS

The accession number for the microarray data reported in this paper is GEO: GSE76686.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2016.01.003.

AUTHOR CONTRIBUTIONS

C.A. and Q.Z. conceived and designed the experiments and analyzed the results; C.A. performed most of the experiments; A.T., J.L., M.S.S., C.A.R., and D.T.B. helped design and perform the organoid experiments; G.X. and S.M. analyzed the microarray data and performed statistical analysis; C.V. and D.M. helped design and perform the bioengineered stomach experiments; R.A.S helped design the Cdx2 experiments; D.A.M., B.Z.S., D.M., R.A.S, Q.X, and D.T.B. contributed reagents/materials/analysis; and Q.Z. and C.A. wrote the paper. All authors discussed the manuscript.

ACKNOWLEDGMENTS

We thank Boston Children's Hospital core facility for Illumina array; Joslin Specialized Assay Core for insulin measurement; Juliana Brown for expert help with pancreatectomy; members of the Zhou lab for advice and feedback; and Gordon Weir, Susan Bonner-Weir, Andrew Leiter, Konrad Hochedlinger, and Jay Rajagopal for discussion and reading of the manuscript. This study was supported by awards from the National Institute of Health and the Harvard Stem Cell Institute (U01 DK089536 and DP-0144-14-00) to Q.Z., awards (R01DK084056, the Timothy Murphy Fund, the IDDRC P30HD18655, and the HDDC P30DK034854) to D.T.B., and R01DK082889 to R.A.S. C.A. is a recipient of HSCI-Sternlicht Fellowship.

Received: May 1, 2015
Revised: October 5, 2015
Accepted: January 8, 2016
Published: February 18, 2016

REFERENCES

Azzi, J., Geara, A.S., El-Sayegh, S., and Abdi, R. (2010). Immunological aspects of pancreatic islet cell transplantation. Expert Rev. Clin. Immunol. *6*, 111–124.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., and Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003–1007.

Barker, N., Huch, M., Kujala, P., van de Wetering, M., Snippert, H.J., van Es, J.H., Sato, T., Stange, D.E., Begthel, H., van den Born, M., et al. (2010). Lgr5(+ve) stem cells drive self-renewal in the stomach and build long-lived gastric units in vitro. Cell Stem Cell *6*, 25–36.

Bonner-Weir, S. (2000). Life and death of the pancreatic beta cells. Trends Endocrinol. Metab. *11*, 375–378.

Bouchi, R., Foo, K.S., Hua, H., Tsuchiya, K., Ohmura, Y., Sandoval, P.R., Ratner, L.E., Egli, D., Leibel, R.L., and Accili, D. (2014). FOXO1 inhibition yields functional insulin-producing cells in human gut organoid cultures. Nat. Commun. *5*, 4242.

Butler, A.E., Janson, J., Bonner-Weir, S., Ritzel, R., Rizza, R.A., and Butler, P.C. (2003). Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. Diabetes *52*, 102–110.

Chen, Y.J., Finkbeiner, S.R., Weinblatt, D., Emmett, M.J., Tameire, F., Yousefi, M., Yang, C., Maehr, R., Zhou, Q., Shemer, R., et al. (2014). De novo formation of insulin-producing "neo-β cell islets" from intestinal crypts. Cell Rep. *6*, 1046–1058.

Choi, E., Roland, J.T., Barlow, B.J., O'Neal, R., Rich, A.E., Nam, K.T., Shi, C., and Goldenring, J.R. (2014). Cell lineage distribution atlas of the human stomach reveals heterogeneous gland populations in the gastric antrum. Gut *63*, 1711–1720.

Egerod, K.L., Engelstoft, M.S., Grunddal, K.V., Nøhr, M.K., Secher, A., Sakata, I., Pedersen, J., Windeløv, J.A., Füchtbauer, E.M., Olsen, J., et al. (2012). A major lineage of enteroendocrine cells coexpress CCK, secretin, GIP, GLP-1, PYY, and neurotensin but not somatostatin. Endocrinology *153*, 5782–5795.

Field, B.C., Chaudhri, O.B., and Bloom, S.R. (2010). Bowels control brain: gut hormones and obesity. Nat. Rev. Endocrinol. *6*, 444–453.

Furuyama, K., Kawaguchi, Y., Akiyama, H., Horiguchi, M., Kodama, S., Kuhara, T., Hosokawa, S., Elbahrawy, A., Soeda, T., Koizumi, M., et al. (2011). Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. Nat. Genet. *43*, 34–41.

Gao, N., White, P., and Kaestner, K.H. (2009). Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. Dev. Cell *16*, 588–599.

Graf, T., and Enver, T. (2009). Forcing cells to change lineages. Nature *462*, 587–594.

Gu, G., Dubauskaite, J., and Melton, D.A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. Development *129*, 2447–2457.

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

Habib, A.M., Richards, P., Cairns, L.S., Rogers, G.J., Bannon, C.A., Parker, H.E., Morley, T.C., Yeo, G.S., Reimann, F., and Gribble, F.M. (2012). Overlap of endocrine hormone expression in the mouse intestine revealed by transcriptional profiling and flow cytometry. Endocrinology *153*, 3054–3065.

Hebrok, M. (2012). Generating β cells from stem cells-the story so far. Cold Spring Harb. Perspect. Med. *2*, a007674.

Jenny, M., Uhl, C., Roche, C., Duluc, I., Guillermin, V., Guillemot, F., Jensen, J., Kedinger, M., and Gradwohl, G. (2002). Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. EMBO J. *21*, 6338–6347.

Johannesson, B., Sui, L., Freytes, D.O., Creusot, R.J., and Egli, D. (2015). Toward beta cell replacement for diabetes. EMBO J. *34*, 841–855.

Karam, S.M., and Leblond, C.P. (1993). Dynamics of epithelial cells in the corpus of the mouse stomach. V. Behavior of entero-endocrine and caveolated cells: general conclusions on cell kinetics in the oxyntic epithelium. Anat. Rec. *236*, 333–340.

Kitamura, Y.I., Kitamura, T., Kruse, J.P., Raum, J.C., Stein, R., Gu, W., and Accili, D. (2005). FoxO1 protects against pancreatic beta cell failure through NeuroD and MafA induction. Cell Metab. *2*, 153–163.

Lakey, J.R., Mirbolooki, M., and Shapiro, A.M. (2006). Current status of clinical islet cell transplantation. Methods Mol. Biol. *333*, 47–104.

Lee, C.S., Perreault, N., Brestelli, J.E., and Kaestner, K.H. (2002). Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. Genes Dev. *16*, 1488–1497.

Lehy, T., and Willems, G. (1976). Population kinetics of antral gastrin cells in the mouse. Gastroenterology *71*, 614–619.

Li, H.J., Johnston, B., Aiello, D., Caffrey, D.R., Giel-Moloney, M., Rindi, G., and Leiter, A.B. (2014a). Distinct cellular origins for serotonin-expressing and enterochromaffin-like cells in the gastric corpus. Gastroenterology *146*, 754–764.

Li, W., Cavelti-Weder, C., Zhang, Y., Clement, K., Donovan, S., Gonzalez, G., Zhu, J., Stemann, M., Xu, K., Hashimoto, T., et al. (2014b). Long-term persistence and development of induced pancreatic beta cells generated by lineage conversion of acinar cells. Nat. Biotechnol. *32*, 1223–1230.

Li, W., Nakanishi, M., Zumsteg, A., Shear, M., Wright, C., Melton, D.A., and Zhou, Q. (2014c). In vivo reprogramming of pancreatic acinar cells to three islet endocrine subtypes. eLife *3*, e01846.

Maemura, T., Ogawa, K., Shin, M., Mochizuki, H., and Vacanti, J.P. (2004). Assessment of tissue-engineered stomach derived from isolated epithelium organoid units. Transplant. Proc. *36*, 1595–1599.

May, C.L., and Kaestner, K.H. (2010). Gut endocrine cell development. Mol. Cell. Endocrinol. *323*, 70–75.

McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.H., Mayhew, C.N., Spence, J.R., Zavros, Y., and Wells, J.M. (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. Nature *516*, 400–404.

Messier, B., and Leblond, C.P. (1960). Cell proliferation and migration as revealed by radioautography after injection of thymidine-H3 into male rats and mice. Am. J. Anat. *106*, 247–285.

Nostro, M.C., and Keller, G. (2012). Generation of beta cells from human pluripotent stem cells: Potential for regenerative medicine. Semin. Cell Dev. Biol. *23*, 701–710.

Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., Hogan, B.L., and Wright, C.V. (1996). PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development *122*, 983–995.

Rahier, J., Guiot, Y., Goebbels, R.M., Sempoux, C., and Henquin, J.C. (2008). Pancreatic beta-cell mass in European subjects with type 2 diabetes. Diabetes Obes. Metab. *10* (suppl 4), 32–42.

Reimann, F., Habib, A.M., Tolhurst, G., Parker, H.E., Rogers, G.J., and Gribble, F.M. (2008). Glucose sensing in L cells: a primary cell study. Cell Metab. *8*, 532–539.

Schiesser, J.V., and Wells, J.M. (2014). Generation of β cells from human pluripotent stem cells: are we there yet? Ann. N Y Acad. Sci. *1311*, 124–137.

Schonhoff, S.E., Giel-Moloney, M., and Leiter, A.B. (2004a). Minireview: Development and differentiation of gut endocrine cells. Endocrinology *145*, 2639–2644.

Schonhoff, S.E., Giel-Moloney, M., and Leiter, A.B. (2004b). Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. Dev. Biol. *270*, 443–454.

Silberg, D.G., Sullivan, J., Kang, E., Swain, G.P., Moffett, J., Sund, N.J., Sackett, S.D., and Kaestner, K.H. (2002). Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. Gastroenterology *122*, 689–696.

Simmini, S., Bialecka, M., Huch, M., Kester, L., van de Wetering, M., Sato, T., Beck, F., van Oudenaarden, A., Clevers, H., and Deschamps, J. (2014). Transformation of intestinal stem cells into gastric stem cells on loss of transcription factor Cdx2. Nat. Commun. *5*, 5728.

Speer, A.L., Sala, F.G., Matthews, J.A., and Grikscheit, T.C. (2011). Murine tissue-engineered stomach demonstrates epithelial differentiation. J. Surg. Res. *171*, 6–14.

Talchai, S.C., and Accili, D. (2015). Legacy Effect of Foxo1 in Pancreatic Endocrine Progenitors on Adult β-Cell Mass and Function. Diabetes *64*, 2868–2879.

Talchai, C., Xuan, S., Kitamura, T., DePinho, R.A., and Accili, D. (2012a). Generation of functional insulin-producing cells in the gut by Foxo1 ablation. Nat. Genet. *44*, 406–412.

Talchai, C., Xuan, S., Lin, H.V., Sussel, L., and Accili, D. (2012b). Pancreatic β cell dedifferentiation as a mechanism of diabetic β cell failure. Cell *150*, 1223–1234.

Thompson, E.M., Price, Y.E., and Wright, N.A. (1990). Kinetics of enteroendocrine cells with implications for their origin: a study of the cholecystokinin and gastrin subpopulations combining tritiated thymidine labelling with immunocytochemistry in the mouse. Gut *31*, 406–411.

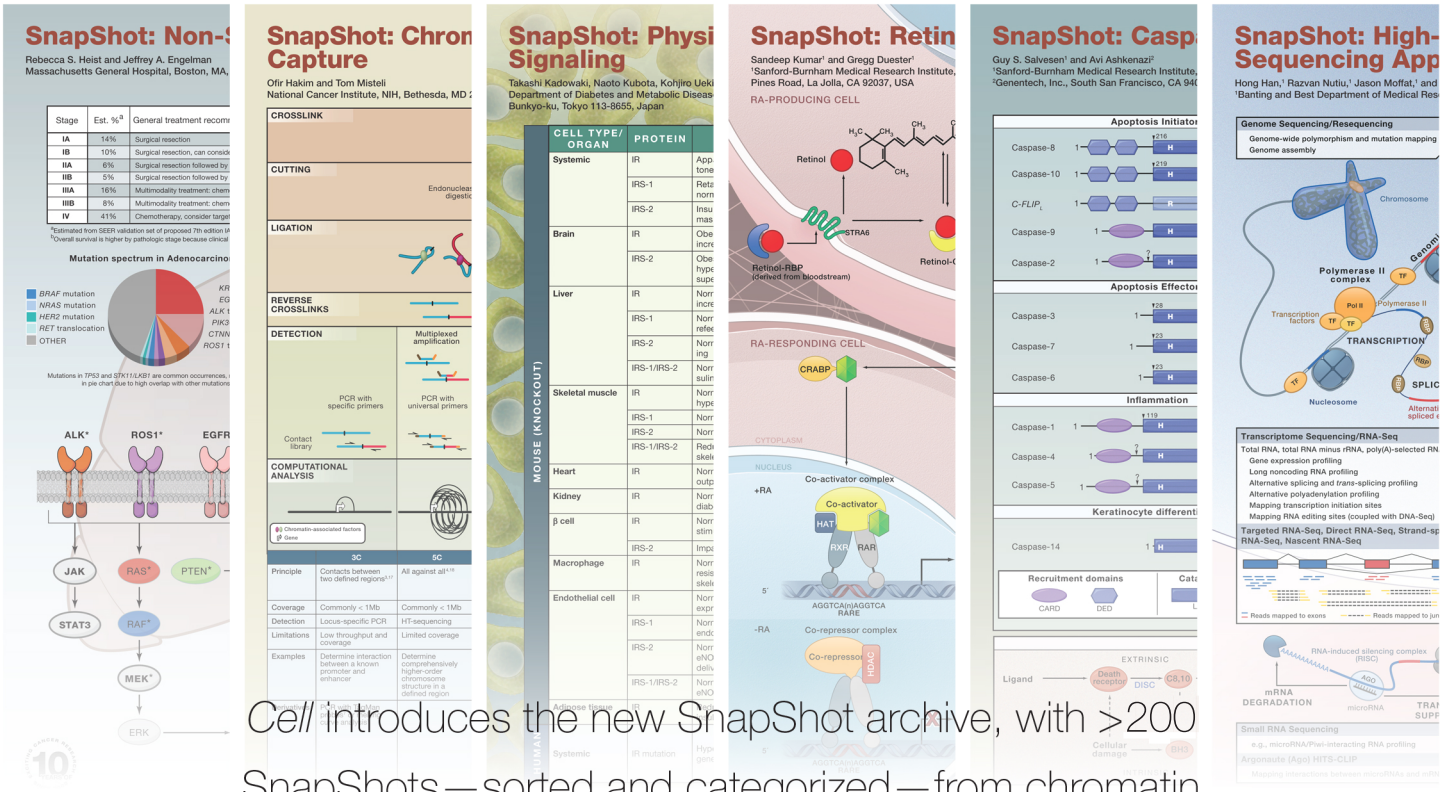
Verzi, M.P., Shin, H., San Roman, A.K., Liu, X.S., and Shivdasani, R.A. (2013). Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. Mol. Cell. Biol. *33*, 281–292.

Wells, J.M., and Melton, D.A. (1999). Vertebrate endoderm development. Annu. Rev. Cell Dev. Biol. *15*, 393–410.

Zhou, Q., and Melton, D.A. (2008). Pathways to new beta cells. Cold Spring Harb. Symp. Quant. Biol. *73*, 175–181.

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.

Cell SnapShot archive



Cell introduces the new SnapShot archive, with >200 SnapShots—sorted and categorized—from chromatin remodelers and autophagy to cancer and autism.

All SnapShots published from a year ago or more are open access and freely available.

Sponsored By

R&D SYSTEMS®
a bio-techne® brand

view the archive

cell.com/snapshots

Cell Stem Cell
Short Article

A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells

Courtney S. Young,[1,2,3,4] Michael R. Hicks,[2,3,5] Natalia V. Ermolova,[2,4] Haruko Nakano,[3,7] Majib Jan,[2,5,6] Shahab Younesi,[2,5,6] Saravanan Karumbayaram,[3,5] Chino Kumagai-Cresse,[2,4] Derek Wang,[2,5] Jerome A. Zack,[1,3,5] Donald B. Kohn,[1,2,3,5] Atsushi Nakano,[1,2,3,7] Stanley F. Nelson,[1,2,3,8] M. Carrie Miceli,[1,2,3,5] Melissa J. Spencer,[1,2,3,4,*] and April D. Pyle[1,2,3,5,*]

[1]Molecular Biology Institute, University of California, Los Angeles, CA 90095, USA
[2]Center for Duchenne Muscular Dystrophy, University of California, Los Angeles, CA 90095, USA
[3]Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, CA 90095, USA
[4]Department of Neurology, University of California, Los Angeles, CA 90095, USA
[5]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA 90095, USA
[6]CIRM Bridges Program, California State University, Northridge, CA 91330, USA
[7]Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, CA 90095, USA
[8]Department of Human Genetics, University of California, Los Angeles, CA 90095, USA
*Correspondence: mspencer@mednet.ucla.edu (M.J.S.), apyle@mednet.ucla.edu (A.D.P.)
http://dx.doi.org/10.1016/j.stem.2016.01.021

SUMMARY

Mutations in *DMD* disrupt the reading frame, prevent dystrophin translation, and cause Duchenne muscular dystrophy (DMD). Here we describe a CRISPR/Cas9 platform applicable to 60% of DMD patient mutations. We applied the platform to DMD-derived hiPSCs where successful deletion and non-homologous end joining of up to 725 kb reframed the *DMD* gene. This is the largest CRISPR/Cas9-mediated deletion shown to date in *DMD*. Use of hiPSCs allowed evaluation of dystrophin in disease-relevant cell types. Cardiomyocytes and skeletal muscle myotubes derived from reframed hiPSC clonal lines had restored dystrophin protein. The internally deleted dystrophin was functional as demonstrated by improved membrane integrity and restoration of the dystrophin glycoprotein complex in vitro and in vivo. Furthermore, miR31 was reduced upon reframing, similar to observations in Becker muscular dystrophy. This work demonstrates the feasibility of using a single CRISPR pair to correct the reading frame for the majority of DMD patients.

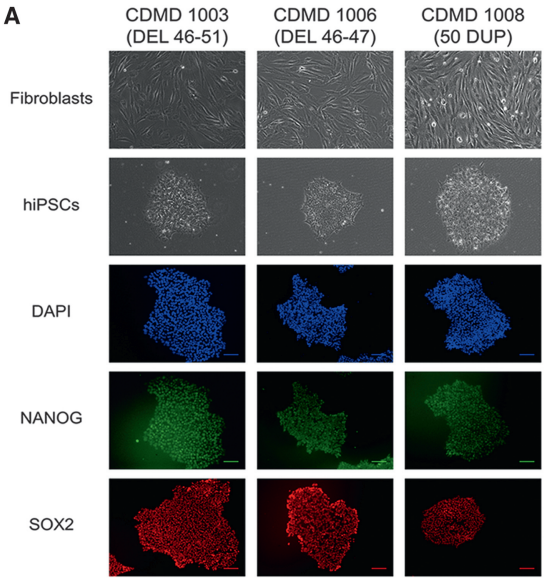
INTRODUCTION

Duchenne muscular dystrophy (DMD) is the most common fatal genetic disease of childhood, affecting ~1 in 3,500–5,000 boys. In DMD, progressive muscle degeneration generally leads to death in the twenties, and there are currently no highly effective therapies. DMD is often caused by frameshifting exonic deletions in *DMD*, which encodes dystrophin. Dystrophin stabilizes the dystrophin glycoprotein complex (DGC) at the sarcolemma; loss of functional dystrophin leads to the degradation of DGC components, which results in muscle membrane fragility and leakage of creatine kinase (CK) (Pearce et al., 1964). Approximately 60% of mutations causing DMD occur between *DMD* exons 45–55 (Béroud et al., 2007). Multiple independent clinical reports in patients and dystrophic mice have revealed that in-frame deletions of exons 45–55 produce an internally deleted dystrophin protein and are associated with a very mild Becker muscular dystrophy (BMD) disease course, with some patients still asymptomatic in their sixties (Béroud et al., 2007; Echigoya et al., 2015; Nakamura et al., 2008; Taglia et al., 2015). Thus, genetic manipulation to create a large deletion of exons 45–55 is a therapeutic strategy to restore the reading frame for 60% of DMD patients with mutations in this region.

One promising approach to induce genetic correction of *DMD* is through the use of the bacterially acquired immune surveillance system known as clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated nuclease (Cas) 9. In this system a short guide RNA (gRNA), which is complimentary to a specific site in the genome, is used to target the Cas9 nuclease and induce double-stranded breaks (DSBs). The DSBs can be repaired through non-homologous end joining (NHEJ) or homology-directed repair.

Previous work has shown that CRISPR/Cas9 components can modify the *DMD* gene (Li et al., 2015; Long et al., 2014, 2016; Nelson et al., 2016; Ousterout et al., 2015; Tabebordbar et al., 2016; Wojtal et al., 2016; Xu et al., 2015). In this investigation, we describe a therapeutically relevant CRISPR/Cas9 platform that we designed to modify *DMD*. Our platform involves excision of exons 45–55 and NHEJ to reframe dystrophin through creation of an internally deleted protein that is stable and functional. The internally deleted protein mimics the naturally occurring exon 45–55 deletion observed in mild BMD patients and encompasses 60% of DMD patient mutations.

For the first time, we demonstrate CRISPR/Cas9-mediated deletion and NHEJ of up to 725 kb of the *DMD* gene in human induced pluripotent stem cell (hiPSC) lines. We show that CRISPR/Cas9 reframed, hiPSC-derived skeletal and cardiac



D

CELL IDENTIFIER	MUTATION	FOR NHEJ # EXONS	FOR NHEJ DISTANCE
CDMD 1003	DEL 46-51	5 EXONS (45 + 52-55)	530 kb
CDMD 1006	DEL 46-47	9 EXONS (45 + 48-55)	670 kb
CDMD 1008	50 DUP	12 EXONS (45-55)	725 kb

Figure 1. CDMD hiPSCs Are Pluripotent and Genetically Stable

(A) CDMD hiPSCs were generated from DMD fibroblasts. Brightfield images depict fibroblasts before and after reprogramming to hiPSCs. Immunocytochemical staining reveals that cells express pluripotency markers NANOG (green) and SOX2 (red). Scale bar, 100 μm.

(B) Karyotyping of all lines is shown.

(C) CDMD hiPSCs were injected into mice to test teratoma formation in vivo. Representative H&E stainings of the three germ layers (endoderm, mesoderm, and ectoderm) are shown.

(D) Patient mutations for each CDMD hiPSC line are shown. In addition, the number of exons and the approximate distance necessary for successful NHEJ is indicated, based on comparative genomic hybridization data for the patient's underlying mutation size.

muscle cells express stable dystrophin that improves membrane stability and restores a DGC member, β-dystroglycan. We also demonstrate reduced microRNA 31 (miR31) levels after the reading frame is restored, consistent with the observations made in BMD patients (Cacchiarelli et al., 2011). Furthermore, we show restoration of dystrophin and β-dystroglycan in vivo after engraftment of reframed hiPSC-derived skeletal muscle cells into a mouse model of DMD. This work sets the stage for use of reframed DMD hiPSC-derived cells or in vivo correction strategies using CRISPR/Cas9 for direct translation to patients with DMD.

RESULTS

DMD hiPSC Lines Are Pluripotent and Genetically Stable

We have developed several xenobiotic-free hiPSC lines derived from wild-type and DMD patient fibroblasts using current good manufacturing practice protocols. Each DMD hiPSC line harbors a unique frameshifting *DMD* mutation within the exon 45–55 hotspot region. All hiPSC lines (Center for Duchenne Muscular Dystrophy [CDMD] 1003, 1006, and 1008) express pluripotency markers (NANOG and SOX2) and are karyotypically normal (Figures 1A and 1B). CDMD hiPSCs maintain pluripotency, as they form teratomas in vivo that represent all three germ layers (Figure 1C), and each harbor unique mutations (Figure 1D).

CRISPR/Cas9-Mediated Deletion and NHEJ of up to 725 kb in the *DMD* Gene

In order to delete exons 45–55 of *DMD*, gRNAs were designed to target introns 44 and 55. gRNA sites were chosen to only retain ~500 bp of the intron next to each of the flanking exons (44 and 56). The rationale for this design is to develop gRNAs applicable to as many patient mutations as possible and to ensure that a small functional chimeric intron is generated. During NHEJ, the 3′ end of intron 44 and the 5′ end of intron 55 join to create a ~1 kb chimeric intron (Figure 2A). We expect that introns generated in this manner are functional and splice correctly to create an in-frame transcript, with exon 44 joined with exon 56.

Since hiPSCs are challenging to genetically manipulate, human embryonic kidney (HEK) 293FT cells were used to screen five gRNAs at each intronic region. All gRNAs demonstrated individual cutting activity on Surveyor assay up to 34% (Figures S1A and S1B). Using multiplex PCR, gRNAs transfected in pairs were shown to effectively delete the entire 708 kb region encompassing exons 45–55 (Figures S1C and S1D).

In order to assess the feasibility of an exon 45–55 deletion across different patient mutations, we applied our gRNAs to three DMD hiPSC lines. The lines (CDMD 1003, 1006, and 1008) require ~530 kb, 670 kb, or 725 kb, respectively, for successful deletion and NHEJ of *DMD*. The gRNAs used were shown to be active in all three lines and effectively deleted exons 45–55 (Figures S2 and S3). Transient puromycin selection of cells nucleofected with the CRISPR plasmids improved the efficiency of deletion in CDMD 1003 and 1006 hiPSCs (Figure S3D).

Clonal Reframed DMD hiPSC Lines Contain No Off-Target Activity at Candidate Sites

Stably deleted DMD hiPSC lines were generated from CDMD 1003 and 1006 by clonal selection after nucleofection with the gRNA pair 44C4 and 55C3 (Figures 2B and 2C) and are pluripotent (Figures 2C and S4B). All reframed lines were karyotypically normal except for one clone (CDMD 1003-81), which was found to contain a 1q32 amplification confirmed via FISH analysis

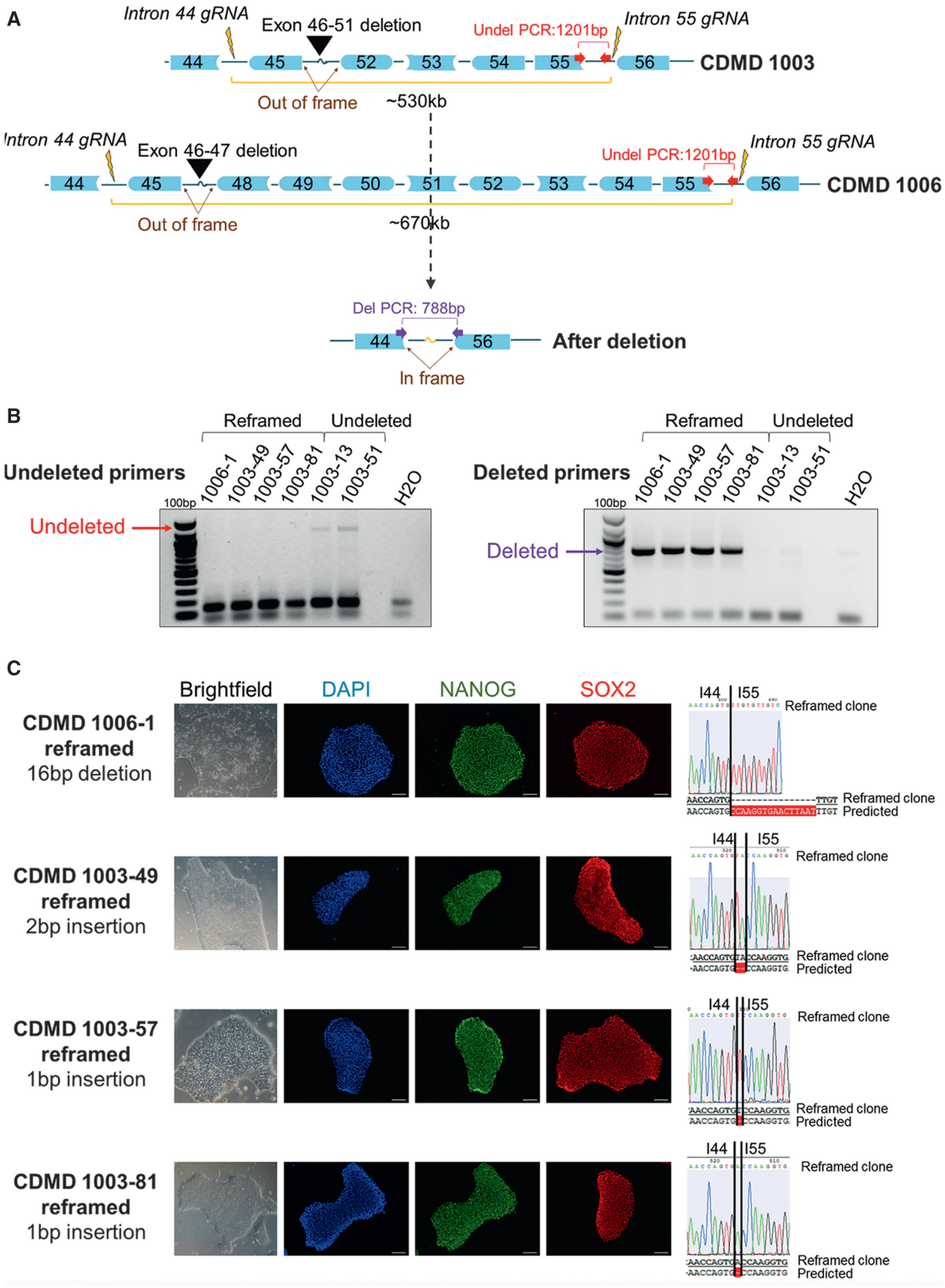


Figure 2. Generation of Stable, Pluripotent CDMD hiPSC Lines with an Exon 45–55 Deletion

(A) Shown is a cartoon (not to scale) of the region of *DMD* targeted for CRISPR/Cas9-mediated deletion using gRNAs specific to introns 44 and 55 (lightning bolts). Successful NHEJ deletes exons 45–55 and restores the reading frame for mutations within this region. Different deletion sizes are required depending on the patient's underlying mutation (black arrow heads).

(B) PCR genotyping of 117 and 109 single-cell clones from parental lines CDMD 1006 and 1003, respectively, was carried out on cells nucleofected with gRNAs 44C4 and 55C3. One clone from CDMD 1006 (CDMD 1006-1) and three from CDMD 1003 (CDMD 1003-49, 1003-57, and 1003-81) were identified as stably

(legend continued on next page)

(Figure S4A), also observed in the original parental line and in all daughter clones after post hoc analysis. The 1q32 amplification is common in hPSCs after extended propagation in culture (Dekel-Naftali et al., 2012), and thus was not a result of CRISPR-mediated off-target activity. To determine off-target activity of our gRNAs, the top ten homologous sites per guide were determined by COSMID (Cradick et al., 2014) and sequenced in all clonal and parental lines. No off-target mutations were observed at any site (Table S2). All variants, besides a heterozygous SNP in chromosome 11, were detected in less than 1% of reads, which is consistent with error in the sequencing method.

Dystrophin (DYS$^{\Delta 45–55}$) Expression Is Restored in Reframed DMD hiPSC-Derived Cardiomyocytes and Skeletal Myotubes

CRISPR/Cas9-mediated deletion of *DMD* should result in an internally deleted dystrophin protein lacking exons 45−55 (hereafter referred to as DYS$^{\Delta 45–55}$). As hiPSCs do not express dystrophin, we differentiated the reframed DMD hiPSC clonal lines to two disease-relevant cell types, cardiomyocytes and skeletal muscle myotubes, using directed differentiation or overexpression (OE) of MyoD to evaluate rescue of DYS$^{\Delta 45–55}$. PCR and sequencing of the exon 44/56 boundary in cDNA from the reframed cardiomyocyte clones demonstrated correct splicing of the dystrophin transcript (Figures S4C and S4D). Additionally, both the reframed cardiac and skeletal muscle cell lines restored dystrophin expression as assayed by immunocytochemistry and western blot (Figures 3A–3C). Compared to wild-type CDMD 1002 or human skeletal muscle myotubes (HSMM), the band was truncated by ~66 kDa as expected.

DYS$^{\Delta 45–55}$ Protein Restores Membrane Functionality to Cardiomyocytes and Skeletal Myotubes In Vitro

Cardiomyocytes or skeletal myotubes lacking dystrophin demonstrate membrane fragility in vitro and respond to osmotic stress by releasing elevated levels of CK (Guan et al., 2014; Menke and Jockusch, 1995), as is seen in human patients (Pearce et al., 1964). To determine whether DYS$^{\Delta 45–55}$ could restore stability to dystrophic plasma membranes, we subjected differentiated cardiomyocytes and skeletal muscle myotubes derived from reframed and out-of-frame hiPSCs to hypo-osmotic conditions. Cells were stressed by incubation in hypo-osmolar solutions (66–240 mosmol) and CK release into the supernatant was measured to show functional improvement after dystrophin restoration. Both the reframed CDMD 1003-49 cardiomyocytes and skeletal muscle cells demonstrated reduced CK release, similar to wild-type (CDMD 1002), versus the out-of-frame CDMD 1003 cells, indicating that DYS$^{\Delta 45–55}$ was capable of reducing membrane fragility (Figure 4A). The same trend was also observed with CDMD 1006/1006-1 cardiomyocytes (Figure S4E). After normalizing and pooling all experiments, we observed that significantly less CK was released at 93, 135, and 240 mosmol in the reframed and wild-type cells compared to out-of-frame (Figure S4F).

CRISPR/Cas9 Reframing Correlates with miR31 Levels in Skeletal Myotubes In Vitro

Elevated levels of miR31 have been observed in DMD patient biopsies compared to wild-type or BMD (Cacchiarelli et al., 2011). We measured levels of miR31 using droplet digital PCR (ddPCR) after differentiation of out-of-frame and reframed CDMD hiPSCs to skeletal myotubes. Reframing *DMD* reduced levels of miR31 (similar to wild-type cells) compared to out-of-frame *DMD,* as is observed in human dystrophinopathies (Figure 4B). Thus, reframing the *DMD* gene normalizes miR31 levels similar to BMD, demonstrating functional rescue of the dystrophic phenotype to a BMD phenotype.

DYS$^{\Delta 45–55}$ Protein Restores the DGC In Vitro and In Vivo

As a third assay of DYS$^{\Delta 45–55}$ functionality, we evaluated its ability to restore the DGC in vitro and in vivo. The DGC member β-dystroglycan was restored and detected at the membrane of reframed hiPSCs, but not out-of-frame hiPSCs, after directed differentiation to skeletal muscle in vitro by immunostaining and western blot (Figures 4C and 4D). Additionally, skeletal muscle cells derived from a wild-type (CDMD 1002), out-of-frame (CDMD 1003), or reframed (CDMD 1003-49) hiPSC line were injected into the tibialis anterior (TA) of NOD *scid* IL2Rgamma (NSG)-mdx mice. Correctly localized dystrophin and β-dystroglycan was only observed in engrafted human cells (demarked by human lamin A/C and spectrin) from the reframed or wild-type lines (Figures 4E and 4F). These studies taken together with the hypo-osmotic stress assays demonstrate the ability of DYS$^{\Delta 45–55}$ to functionally reassemble the DGC and restore membrane stability in vitro and in vivo.

DISCUSSION

Using CRISPR/Cas9 gene editing, we have induced the largest deletion accomplished to date in DMD hiPSCs and restored a functional dystrophin protein. Deletion of *DMD* exons 45–55 has the potential to be therapeutically relevant to 60% of DMD patients. Since this internal deletion has been associated with a very mild disease course in multiple independent patients, a therapy utilizing this approach should create a highly functional dystrophin. We showed successful deletion of exons 45–55 using a single gRNA pair and did not identify any off-target activity at the top ten homologous sites; however, a more comprehensive and unbiased approach should be undertaken such as whole-genome sequencing. Importantly, removal of exons 45–55 resulted in stable dystrophin protein (DYS$^{\Delta 45–55}$) in both cardiomyocytes and skeletal myotubes in vitro. Functionality of

deleted. Deletion PCR genotyping results for six hiPSC clonal lines is shown. One pair of primers (red arrows in A) was located internal to the deletion and only produced a 1,201 bp band in the undeleted clones CDMD 1003-13 and 1003-51. Another primer set (purple arrows in A) flanked the deletion region and produced a 788 bp band only when the deletion and NHEJ occurred successfully, as in the reframed clones CDMD 1006-1, 1003-49, 1003-57, and 1003-81.

(C) Each clonal line maintained normal morphology (brightfield) and expressed NANOG (green) and SOX2 (red) by immunocytochemistry. Scale bar, 100 μm. Shown to the right is the sequence of the gDNA at the rejoining site between introns 44 (I44) and 55 (I55). Sequencing revealed a 16 bp deletion in CDMD 1006-1, a 2 bp insertion in CDMD 1003-49, and 1 bp insertions in CDMD 1003-57 and CDMD 1003-81.

See also Figures S1, S2, S3, S4A, and S4B.

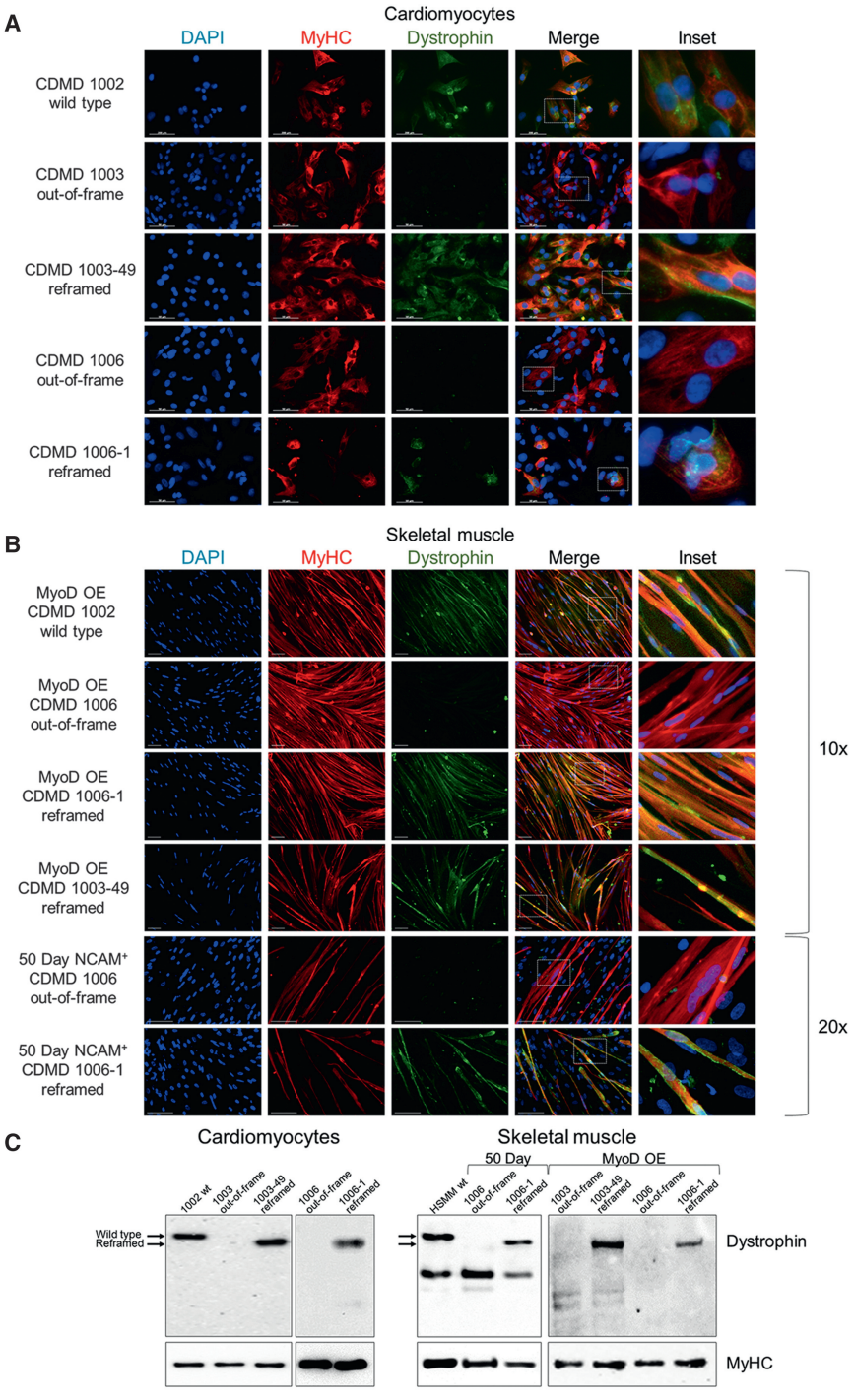


Figure 3. Reframed CDMD hiPSC-Derived Skeletal Muscle and Cardiomyocytes Restore Dystrophin Expression

(A) Immunocytochemical staining of human myosin heavy chain (MyHC, red) and dystrophin (green) of wild-type (CDMD 1002), out-of-frame (CDMD 1003 or 1006) or reframed (CDMD 1003-49 or 1006-1) cardiomyocytes derived from hiPSCs by directed differentiation. Inset depicts zoomed in region defined by the white box. Scale bar, 50 μm.

(B) Immunocytochemical staining of MyHC (red) and dystrophin (green) of wild-type (CDMD 1002), out-of-frame (CDMD 1006) or reframed (CDMD 1006-1 or 1003-49) skeletal muscle myotubes derived from hiPSCs. Myotubes were fused after MyoD OE or from sorted NCAM$^+$ cells after an adapted directed differentiation 50-day protocol was used. Inset depicts zoomed-in region defined by the white box. Scale bar, 100 μm.

(C) Western blots of cell extracts probed with anti-dystrophin. Extracts were from out-of-frame and reframed cardiomyocytes (left) and skeletal muscle myotubes (right), derived from CDMD hiPSCs. Wild-type (wt) hiPSCs (CDMD 1002) or human skeletal muscle myotubes (HSMM) were used as a control for dystrophin. The molecular weight shift caused by the exon 45–55 deletion (1779 bp, ~66 kDa) is evident in reframed versus wild-type dystrophin (arrows). A non-specific band around 220 kDa was seen in some samples. Samples were also probed with anti-MyHC as a loading control (bottom panels).

See also Figures S4C and S4D.

DYS$^{\Delta 45-55}$ was tested in cardiomyocytes and skeletal muscle derived from reframed DMD hiPSCs and demonstrated improved membrane stability by a physiologically relevant measure of CK release, similar to wild-type. The ability to evaluate cardiomyocyte functionality is an advantage of using hiPSCs, as some current preclinical and clinical studies for DMD therapies do not efficiently target the heart (e.g., exon skipping; Arechavala-Gomeza et al., 2012). Additionally, we demonstrated a normalization in miR31 levels, a microRNA that inhibits dystrophin, after reading frame restoration, similar to what is observed in human BMD patients (Cacchiarelli et al., 2011). Finally, we show restored DGC localization in vitro and in vivo, which further validates the functionality of DYS$^{\Delta 45-55}$.

Previous work by Ousterout et al. (2015) demonstrated that multiplexed gRNAs can restore the *DMD* reading frame in primary myoblasts. However, myoblasts do not provide a renewable source of stem cells, which is a requirement for long-term therapeutic efficacy (Partridge, 2002). In contrast, we used hiPSCs, which offer the opportunity to evaluate the internally deleted dystrophin protein in multiple cell types that are affected in DMD, and in future studies, they may provide a renewal source of corrected progenitor cells. Our work is further distinguished from previous studies as we are the only group to show restoration of dystrophin function on membrane integrity, miR31 expression, and the DGC in cardiac and skeletal muscle cells following CRISPR-mediated gene editing.

An advantage of our CRISPR platform is the therapeutic potential of a single pair of gRNAs to treat the majority of DMD patients. By designing gRNAs that accomplish a

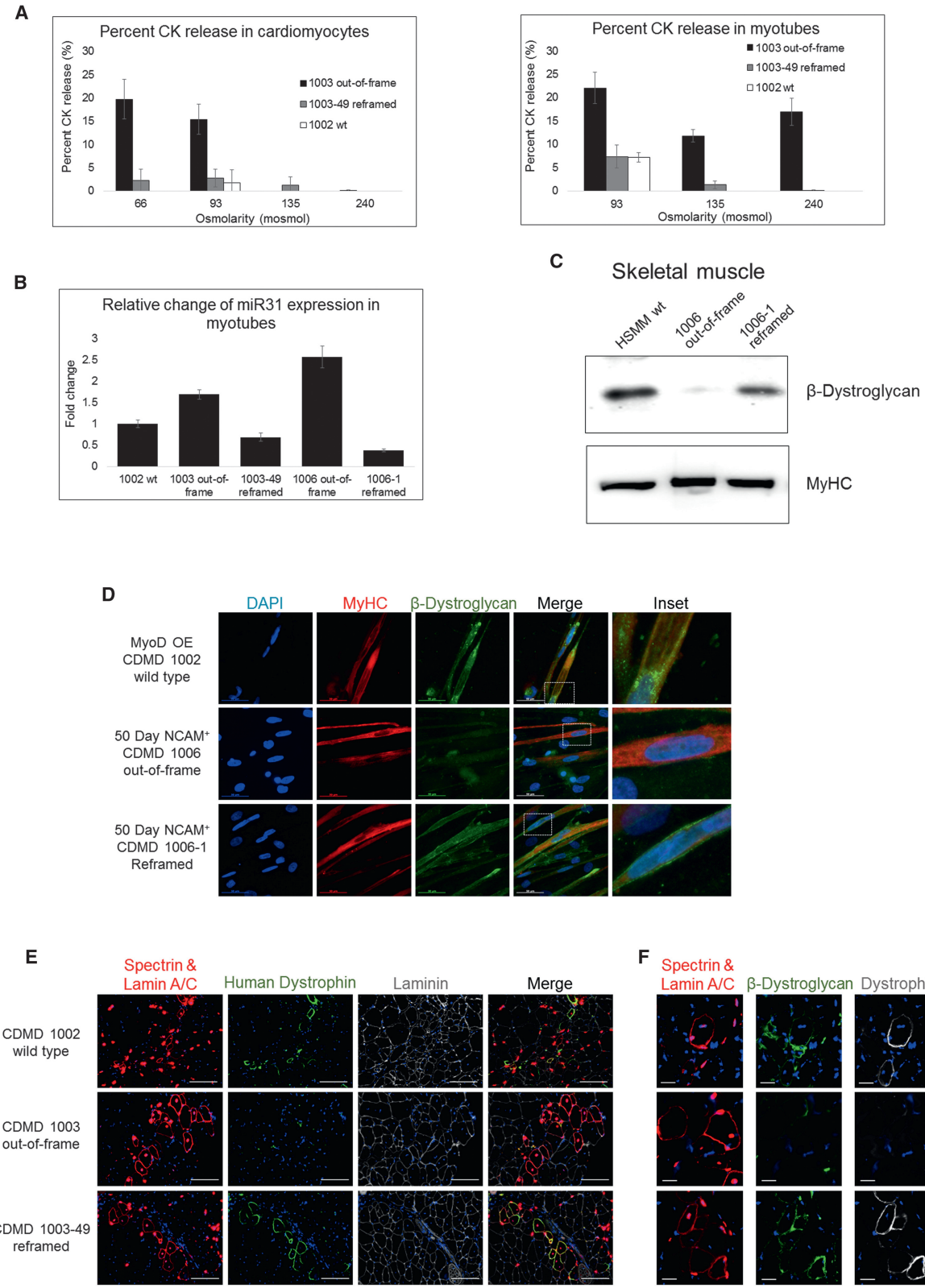


Figure 4. Reframed hiPSC-Derived Cardiomyocytes and Skeletal Muscle Cells Demonstrate Restored Function In Vitro and In Vivo

(A) Representative graphs of CK release assays from cells exposed to hypo-osmotic conditions. Cardiomyocytes and skeletal muscle myotubes derived from hiPSCs were subjected to a range of osmolarities below 240 mosmol, and CK release to the supernatant was measured as an indication of membrane fragility. Data are presented as average ± SE.

(legend continued on next page)

deletion that encompasses the majority of DMD mutations, this approach is optimized for future clinical studies. It would be unreasonable to design, validate, and evaluate off targets for every new CRISPR pair tailored for each individual patient. Additionally, CRISPR/Cas9 is advantageous over exon skipping, as it results in permanent restoration of the reading frame as opposed to transient effects on RNA splicing. Previously, Li et al. (2015) used CRISPR/Cas9 to induce exon skipping, frameshifting, or exon knockin to restore dystrophin in a DMD hiPSC line with an exon 44 deletion; however, their platform is only applicable to 3%–9% of DMD patients (Bladen et al., 2015), and two of their strategies relied on the creation of indels, which would be difficult to apply consistently to each patient. While Ousterout et al. deleted exons 45–55, they removed significantly less of the intervening region (336 kb) and thus their approach would cover fewer patient mutations within the hotspot region. This is because many mutations extend into the intronic region; thus, by designing gRNAs that encompass more of the intron, our platform is applicable to more patients.

Another benefit of using this platform to delete a large portion of *DMD*, as opposed to single exons, is the known correlation of $DYS^{\Delta 45–55}$ with a mild BMD phenotype. Large deletions in the rod domain of dystrophin often produce a more functional (more like wild-type) protein, than even very small deletions (Harper et al., 2002). Larger deletions, which remove hinge III (exons 50–51), are believed to lead to a milder BMD phenotype than smaller deletions, or those that retain hinge III (Carsana et al., 2005). Thus, in many cases larger deletions are more therapeutically beneficial than smaller ones, due to the way they affect the secondary structure of the protein.

In summary, we have developed a potentially therapeutic gene editing platform for DMD to permanently restore the dystrophin reading frame in multiple patient-derived hiPSCs. Our approach using CRISPR/Cas9 and NHEJ deletes up to 725 kb of *DMD* encompassing exons 45–55 and restores dystrophin protein function in both cardiomyocytes and skeletal muscle cells derived from reframed hiPSCs. A current limitation of this platform is that clinical protocols still need to be developed that allow rapid clonal line derivation and the utilization of hiPSC-derived cardiac and skeletal muscle progenitors combined with gene correction. Alternatively, CRISPR/Cas9 to restore the reading frame in DMD mouse models has been delivered directly in vivo (Long et al., 2016; Nelson et al., 2016; Tabebordbar et al., 2016). Thus, applications of this platform in the future will allow for the development of an in situ gene strategy or ex vivo gene correction followed by autologous cell transplantation, either of which offers tremendous potential for DMD.

EXPERIMENTAL PROCEDURES

Differentiation of hiPSCs to Skeletal Muscle Cells and Cardiomyocytes

Skeletal muscle differentiation from hiPSCs was induced using OE of a tamoxifen inducible MyoD-ERT lentivirus or an adapted 50 day directed differentiation protocol where $NCAM^+$ $HNK1^-$ cells underwent fluorescence-activated cell sorting at day 50. Cardiomyocytes were derived through aggregates over 30 days. See Supplemental Experimental Procedures.

Engraftment into Immunodeficient Mice

NSG immunodeficient mice (Jackson Laboratory) were crossed to mdx *scid* mice (Jackson Laboratory) to generate NSG-mdx mice (see Supplemental Experimental Procedures). Five- to seven-week-old NSG-mdx mice were pretreated with 50 μl of 10 μM cardiotoxin (Sigma-Aldrich) injected into the right TA 24 hr prior to engraftment. For MyoD OE cells, 100 μl of 5 mg/ml tamoxifen (Sigma-Aldrich) was i.p. injected for 5 days beginning on the day prior to engraftment. 1×10^6 cells in HBSS were injected intramuscularly and the TA was harvested after 30 days. See Supplemental Experimental Procedures.

Hypo-osmotic Stress CK Release Assay

Terminally differentiated skeletal muscle cells and cardiomyocytes plated in duplicate were stressed by incubation in hypo-osomolar solutions ranging from 66 to 240 mosmol (see Supplemental Experimental Procedures) for 20 min at 37°C. CK was measured in triplicate from the supernatant and cell lysate with the Creatine Kinase-SL kit (Sekisui Diagnostics) according to the manufacturer's instructions.

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes four figures, two tables and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2016.01.021.

AUTHOR CONTRIBUTIONS

Conceptualization, Methodology, Writing–Original Draft, Visualization, Project Administration, C.S.Y., M.J.S., and A.D.P.; Validation, C.S.Y., M.R.H., and N.V.E.; Formal Analysis, C.S.Y.; Investigation, C.S.Y., M.R.H., N.V.E., H.N., M.J., S.Y., S.K., C.K.-C., and D.W.; Resources, A.N., S.F.N., M.C.M., M.J.S., and A.D.P.; Writing–Review and Editing, C.S.Y., M.R.H., N.E.V., S.K., C.K.-C., D.B.K., A.N., S.F.N., M.C.M., M.J.S., and A.D.P.; Supervision, J.A.Z., D.B.K., A.N., M.C.M., M.J.S., and A.D.P.; Funding Acquisition, M.J.S. and A.D.P.

(B) Fold change in expression of miR31 measured by ddPCR in myotubes derived from out-of-frame or reframed hiPSCs by MyoD OE, normalized to wild-type (CDMD 1002). Data are presented as average ± SD.
(C) Western blots of cell extracts probed with anti-β-dystroglycan. Extracts were from out-of-frame and reframed skeletal muscle myotubes derived by MyoD OE. HSMM was used as a positive control. Samples were also probed with anti-MyHC as a loading control (bottom panel).
(D) Immunocytochemical staining of MyHC (red) and β-dystroglycan (green), a component of the DGC, in wild-type (CDMD 1002), out-of-frame (CDMD 1006), or reframed (CDMD 1006-1) skeletal muscle myotubes. Inset depicts zoomed-in region defined by the white box. Scale bar, 50 μm.
(E) Assessment of human dystrophin restoration in wild-type (CDMD 1002), out-of-frame (CDMD 1003), and reframed (CDMD 1003-49) MyoD OE cells engrafted into the TA of NSG-mdx mice. Engrafted human cells were identified by co-immunostaining for human spectrin and lamin A/C (shown in red). Positive staining for human dystrophin is shown in green and all fibers are shown using laminin (gray). All sections were stained with DAPI (blue) to identify nuclei. Scale bar, 100 μm.
(F) Assessment of β-dystroglycan restoration in human fibers from wild-type (CDMD 1002), out-of-frame (CDMD 1003), and reframed (CDMD 1003-49) MyoD OE cells engrafted into the TA of NSG-mdx mice. Engrafted human cells were identified by co-immunostaining for human spectrin and lamin A/C (shown in red). Positive staining for dystrophin is shown in gray and β-dystroglycan is shown in green. All sections were stained with DAPI (blue) to identify nuclei. Cell order is the same as noted in (E). Scale bar, 20 μm.
See also Figures S4E and S4F.

ACKNOWLEDGMENTS

We would like to thank Doug Black, Richard Wang, and Haibin Xi for suggestions and Diana Becerra and Jane Wen for technical assistance. The following cores were utilized: CDMD Muscle Phenotyping and Imaging Core, HTS and Cell Repository Core, and Bioinformatics and Genomics Core (supported by NIAMS-P30 AR057230); the UCLA Broad Stem Cell Research Center (BSCRC) Flow Cytometry Core Resource, the UCLA Jonsson Comprehensive Cancer Center (JCCC), and Center for AIDS Research (CFAR) Flow Cytometry Core Facility (supported by NIH P30 CA016042, 5P30 AI028697); and the UCLA Integrated Molecular Technologies Core (supported by CURE/P30 DK041301), the UCLA Humanized Mouse Core (supported by the JCCC, BSCRC, CFAR, and NIH/NIAID AI028697), and the UCLA GenoSeq core. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1144087 (C.S.Y.). Funding was provided by NIAMS of the NIH (P30 AR057230 to M.J.S., A.D.P., and A.N. and R01AR064327 to A.D.P.), the Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research at UCLA (M.J.S. and A.D.P.), and a Rose Hills Foundation Research Award (A.D.P.).

Received: September 9, 2015
Revised: December 18, 2015
Accepted: January 22, 2016
Published: February 11, 2016

REFERENCES

Arechavala-Gomeza, V., Anthony, K., Morgan, J., and Muntoni, F. (2012). Antisense oligonucleotide-mediated exon skipping for Duchenne muscular dystrophy: progress and challenges. Curr. Gene Ther. *12*, 152–160.

Béroud, C., Tuffery-Giraud, S., Matsuo, M., Hamroun, D., Humbertclaude, V., Monnier, N., Moizard, M.P., Voelckel, M.A., Calemard, L.M., Boisseau, P., et al. (2007). Multiexon skipping leading to an artificial DMD protein lacking amino acids from exons 45 through 55 could rescue up to 63% of patients with Duchenne muscular dystrophy. Hum. Mutat. *28*, 196–202.

Bladen, C.L., Salgado, D., Monges, S., Foncuberta, M.E., Kekou, K., Kosma, K., Dawkins, H., Lamont, L., Roy, A.J., Chamova, T., et al. (2015). The TREAT-NMD DMD Global Database: analysis of more than 7,000 Duchenne muscular dystrophy mutations. Hum. Mutat. *36*, 395–402.

Cacchiarelli, D., Incitti, T., Martone, J., Cesana, M., Cazzella, V., Santini, T., Sthandier, O., and Bozzoni, I. (2011). miR-31 modulates dystrophin expression: new implications for Duchenne muscular dystrophy therapy. EMBO Rep. *12*, 136–141.

Carsana, A., Frisso, G., Tremolaterra, M.R., Lanzillo, R., Vitale, D.F., Santoro, L., and Salvatore, F. (2005). Analysis of dystrophin gene deletions indicates that the hinge III region of the protein correlates with disease severity. Ann. Hum. Genet. *69*, 253–259.

Cradick, T.J., Qiu, P., Lee, C.M., Fine, E.J., and Bao, G. (2014). COSMID: A web-based tool for identifying and validating CRISPR/Cas off-target sites. Mol. Ther. Nucleic Acids *3*, e214.

Dekel-Naftali, M., Aviram-Goldring, A., Litmanovitch, T., Shamash, J., Reznik-Wolf, H., Laevsky, I., Amit, M., Itskovitz-Eldor, J., Yung, Y., Hourvitz, A., et al. (2012). Screening of human pluripotent stem cells using CGH and FISH reveals low-grade mosaic aneuploidy and a recurrent amplification of chromosome 1q. Eur. J. Hum. Genet. *20*, 1248–1255.

Echigoya, Y., Aoki, Y., Miskew, B., Panesar, D., Touznik, A., Nagata, T., Tanihata, J., Nakamura, A., Nagaraju, K., and Yokota, T. (2015). Long-term efficacy of systemic multiexon skipping targeting dystrophin exons 45-55 with a cocktail of vivo-morpholinos in mdx52 mice. Mol. Ther. Nucleic Acids *4*, e225.

Guan, X., Mack, D.L., Moreno, C.M., Strande, J.L., Mathieu, J., Shi, Y., Markert, C.D., Wang, Z., Liu, G., Lawlor, M.W., et al. (2014). Dystrophin-deficient cardiomyocytes derived from human urine: new biologic reagents for drug discovery. Stem Cell Res. (Amst.) *12*, 467–480.

Harper, S.Q., Hauser, M.A., DelloRusso, C., Duan, D., Crawford, R.W., Phelps, S.F., Harper, H.A., Robinson, A.S., Engelhardt, J.F., Brooks, S.V., and Chamberlain, J.S. (2002). Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. Nat. Med. *8*, 253–261.

Li, H.L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A., Sakurai, H., et al. (2015). Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. Stem Cell Reports *4*, 143–154.

Long, C., McAnally, J.R., Shelton, J.M., Mireault, A.A., Bassel-Duby, R., and Olson, E.N. (2014). Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science *345*, 1184–1188.

Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science *351*, 400–403.

Menke, A., and Jockusch, H. (1995). Extent of shock-induced membrane leakage in human and mouse myotubes depends on dystrophin. J. Cell Sci. *108*, 727–733.

Nakamura, A., Yoshida, K., Fukushima, K., Ueda, H., Urasawa, N., Koyama, J., Yazaki, Y., Yazaki, M., Sakai, T., Haruta, S., et al. (2008). Follow-up of three patients with a large in-frame deletion of exons 45-55 in the Duchenne muscular dystrophy (DMD) gene. J. Clin. Neurosci. *15*, 757–763.

Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Rivera, R.M.C., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., et al. (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science *351*, 403–407.

Ousterout, D.G., Kabadi, A.M., Thakore, P.I., Majoros, W.H., Reddy, T.E., and Gersbach, C.A. (2015). Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. Nat. Commun. *6*, 6244.

Partridge, T. (2002). Myoblast transplantation. Neuromuscul. Disord. *12* (*Suppl 1*), S3–S6.

Pearce, J.M.S., Pennington, R.J.T., and Walton, J.N. (1964). Serum enzyme studies in muscle disease: Part III Serum creatine kinase activity in relatives of patients with the Duchenne type of muscular dystrophy. J. Neurol. Neurosurg. Psychiatry *27*, 181–185.

Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., et al. (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407–411.

Taglia, A., Petillo, R., D'Ambrosio, P., Picillo, E., Torella, A., Orsini, C., Ergoli, M., Scutifero, M., Passamano, L., Palladino, A., et al. (2015). Clinical features of patients with dystrophinopathy sharing the 45-55 exon deletion of DMD gene. Acta Myol. *34*, 9–13.

Wojtal, D., Kemaladewi, D.U., Malam, Z., Abdullah, S., Wong, T.W.Y., Hyatt, E., Baghestani, Z., Pereira, S., Stavropoulos, J., Mouly, V., et al. (2016). Spell Checking Nature: Versatility of CRISPR/Cas9 for Developing Treatments for Inherited Disorders. Am. J. Hum. Genet. *98*, 90–101.

Xu, L., Park, K.H., Zhao, L., Xu, J., El Refaey, M., Gao, Y., Zhu, H., Ma, J., and Han, R. (2015). CRISPR-mediated genome editing restores dystrophin expression and function in mdx mice. Mol. Ther., in press. Published online October 9, 2015. http://dx.doi.org/10.1038/mt.2015.192.

Even more ways to access Cell Press journals

Sponsored By

R&D SYSTEMS
a bio-techne brand

App Features:

- Available to subscribers at no additional charge
- Free 30-day trial to all 30 Cell Press journals
- Interactive reading experience
- Browse abstracts and full-text articles
- Download now, read later
- Make article notes
- Share with your friends and colleagues
- Now on the iPhone, iPad, Android tablet and Android phone

Download the app at cell.com/mobile

CellPress

Cell

Resource

Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency

Davide Cacchiarelli,[1,2,3,7] Cole Trapnell,[6,7] Michael J. Ziller,[1,2,3,7] Magali Soumillon,[1,2,3] Marcella Cesana,[3,4] Rahul Karnik,[1,2,3] Julie Donaghey,[1,2,3] Zachary D. Smith,[1,2,3] Sutheera Ratanasirintrawoot,[3,4] Xiaolan Zhang,[2] Shannan J. Ho Sui,[3,5] Zhaoting Wu,[3,4] Veronika Akopian,[1,2,3] Casey A. Gifford,[1,2,3] John Doench,[2] John L. Rinn,[1,2,3] George Q. Daley,[3,4] Alexander Meissner,[1,2,3] Eric S. Lander,[2] and Tarjei S. Mikkelsen[1,2,3,*]

[1]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA

[2]Broad Institute, Cambridge, MA 02142, USA

[3]Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA

[4]Boston Children's Hospital and Dana-Farber Cancer Institute, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Manton Center for Orphan Disease Research, Howard Hughes Medical Institute, Boston, MA 02115, USA

[5]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

[6]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

[7]Co-first author

*Correspondence: tarjei@broadinstitute.org

http://dx.doi.org/10.1016/j.cell.2015.06.016

SUMMARY

Induced pluripotency is a promising avenue for disease modeling and therapy, but the molecular principles underlying this process, particularly in human cells, remain poorly understood due to donor-to-donor variability and intercellular heterogeneity. Here, we constructed and characterized a clonal, inducible human reprogramming system that provides a reliable source of cells at any stage of the process. This system enabled integrative transcriptional and epigenomic analysis across the human reprogramming timeline at high resolution. We observed distinct waves of gene network activation, including the ordered re-activation of broad developmental regulators followed by early embryonic patterning genes and culminating in the emergence of a signature reminiscent of pre-implantation stages. Moreover, complementary functional analyses allowed us to identify and validate novel regulators of the reprogramming process. Altogether, this study sheds light on the molecular underpinnings of induced pluripotency in human cells and provides a robust cell platform for further studies.

INTRODUCTION

Engineered reprogramming systems have provided useful tools for the study of induced pluripotency. In "secondary" reprogramming systems, somatic cells are first transduced with lentiviral constructs carrying drug-inducible transcription factors. Clonal induced pluripotent stem cells (IPSCs) are then derived and next differentiated back to a somatic state that can be reprogrammed a second time, often with greater efficiency (Hockemeyer et al., 2008; Stadtfeld et al., 2010; Wernig et al., 2008). Because the resulting somatic cells are clonal, this strategy eliminates biases and heterogeneity caused by variable lentiviral delivery and transgene stoichiometry present in "primary" reprogramming experiments (Stadtfeld et al., 2010).

Secondary reprogramming systems engineered from transgenic mouse embryonic fibroblasts (MEFs) have enabled large-scale genomic and epigenomic profiling studies of cells as they reacquire pluripotency (Hussein et al., 2014; Mikkelsen et al., 2008; Polo et al., 2012). These analyses have revealed that somatic identity is rapidly lost upon induction of the reprogramming factors and pluripotency is promoted by an early mesenchymal to epithelial transition (MET) (Li et al., 2010), a process accompanied by removal of several epigenetic roadblocks (Apostolou and Hochedlinger, 2013). It has, however, been difficult to directly compare the reprogramming MEFs to the same process in human cells, due to differences in culture conditions, differential expression of key markers, and other factors. Under standard conditions, murine IPSCs also appear to reprogram with faster kinetics and higher efficiency than human IPSCs and reach a more naive, pre-implantation-like cellular state (Hanna et al., 2010; Nichols and Smith, 2009). Moreover, analyses of intermediate states in previous systems have been complicated by heterogeneity in the initial cell populations and progressive loss of reprogramming capacity over serial passaging (Utikal et al., 2009). Thus, a well-controlled model system that generates intermediately and fully reprogrammed cells with consistent kinetics and efficiency even after extensive expansion in vitro would be a valuable asset for efforts to characterize reprogramming in human cells.

We hypothesized that senescence would be a major contributor to the variability and passage-dependent loss of reprogramming capacity that has been observed in previous attempts to generate human reprogramming systems (Park et al., 2008). We therefore sought to extend the lifespan of human secondary fibroblasts by overexpression of the telomerase gene (hTERT). Here, we employ this approach to generate a robust model system that enables continual propagation of clonal cells with a defined reprogramming capacity. We leverage this model to systematically characterize the transcriptional and epigenomic

CrossMark

changes during reprogramming. Through integrative analyses, we find that OCT4/KLF4/c-MYC/SOX2 (OKMS) induction leads to transient reactivation of genes in a pattern that is suggestive of a reversal of normal development. Unexpectedly, these changes culminate in the emergence of a subpopulation of cells with transcriptional and epigenomic signatures with pre-implantation-like characteristics. Finally, we demonstrate the utility of our secondary system for discovery and characterization of a variety of modulators of reprogramming in human cells.

RESULTS

hTERT Confers Robustness to Secondary Reprogramming Systems

We generated human IPSCs (hIPSCs) from primary BJ foreskin fibroblasts using a doxycycline (DOX)-inducible, polycistronic human OCT4/KLF4/c-MYC/SOX2 (OKMS) cassette. We then differentiated these hIPSCs in a serum-based media (Park et al., 2008) to obtain human inducible fibroblasts-like cells (hiF) that could be subsequently reprogrammed by DOX treatment (Figure 1A). Consistent with previous attempts, both primary BJ cells and secondary hiF generated IPSC colonies that were highly heterogeneous in size and appeared asynchronously over 3 weeks following OKMS induction (Figure 1B). Moreover, secondary hiFs rapidly lost their reprogramming potential with successive passages in culture, which correlated with the appearance of senescent cells (Figure 1C). Foreskin fibroblasts from different donors also displayed variability in proliferation and senescence (Figures S1A–S1E), which influenced reprogramming efficiencies in a passage-dependent manner (Figure S1F). We observed similar variability across different batches of secondary cells generated from the same pluripotent stem cell (PSC) clone (dH1f or hiF; Figures S1A–S1F). These observations highlight the variable reprogramming efficiency of previous primary and secondary reprogramming systems.

To increase the expansion potential of hiF cells, we delivered a lentiviral vector driving constitutive human telomerase (hTERT) expression (Stewart et al., 2003) and derived clonal cell lines (hiF-T; Figure 1A, lower scheme). Inclusion of hTERT in reprogramming cocktails is known to be compatible with IPSC generation (Park et al., 2008). hiF-Ts displayed a lower reprogramming efficiency than early passage hiFs. But unlike both hiF and primary cells, they showed no evidence of senescence and maintained a stable reprogramming efficiency even after 3 months of continuous culture (Figure 1D; Figure S1F). At the same time, hiF-Ts showed growth and apoptosis rates that were equivalent to healthy primary cells and were negative in assays of cellular transformation (Figures S1A–S1E). Thus, hTERT expression also appears to reduce a key source of experimental variability in reprogramming systems.

To better understand the differences between primary (BJ), secondary (hIF), and secondary immortalized (hiF-T) reprogramming systems, we performed expression profiling by RNA sequencing (RNA-seq). We found that hiF cells downregulated proliferative genes after only limited passaging, while hiF-Ts maintained expression of these genes in long-term cultures (Figure 1E). Moreover, hiF cultures expressed high levels of genes associated with "stemness," even after the emergence of senescent cells, indicating either the persistence of primed or privileged subpopulations or incomplete differentiation. Either case might explain the high reprogramming efficiency of early passage hiFs. In contrast, the stemness genes were silenced in BJ and hiF-T cells. The consistency of hiF-T reprogramming may therefore reflect hTERT's ability to block senescence, but may also ensure that the secondary cells can be cultured long enough to acquire a fully differentiated state (Figure 1F).

Importantly, hiF-T cells showed high levels of genomic and transcriptional stability (Figures S1G and S1H). Moreover, gene expression profiles of hIPSCs derived from reprogrammed hiF-T cells (hIPSC-T) are equivalent to reference PSCs, showing expression of expected protein markers and complete silencing of all transgenes, including hTERT (Figures S1I and S2). hIPSC-Ts also maintain the capacity to form all three embryonic germ layers in vitro through directed differentiation (Figures S1J and S1K). We therefore conclude that hIF-T secondary cells provide a faithful and reliable model system for large-scale studies of human reprogramming.

Genome-wide Profiling of Reprogramming Cells

We next leveraged the increased proliferative capacity and decreased heterogeneity of hiF-Ts to collect large numbers of cells for comprehensive immunophenotypic, genomic, and epigenomic analyses of reprogramming. hiF-Ts rapidly lost the somatic cell marker CD13 upon OKMS induction and then acquired the embryonic marker SSEA-3. A subset of the most SSEA-3 positive cells subsequently acquired the pluripotency-associated marker TRA-1-60 (Figure 2A). We collected cells from key stages throughout this process. In the most advanced stages, we fractionated cells based on SSEA-3 or TRA-1-60 expression to isolate those that were transitioning toward pluripotency (Figure 2B). We then profiled mRNA and small RNA by RNA-seq, major histone modifications (H3K4me1, me2, me3; H3K27ac, me3; H3K36me3) by chromatin immunoprecipitation sequencing (ChIP-seq) and DNA methylation by reduced representation bisulfite sequencing (RRBS).

To characterize the major transitions in cellular states during reprogramming, we performed multi-dimensional scaling (MDS) analyses on the resulting data (Figures 2C–2E).

MDS of the RNA-seq data revealed a continuous trajectory of transcriptional changes beginning with uninduced hiF-Ts and ending with established hIPSC-Ts (Figure 2C). Cells with higher SSEA3 or TRA-1-60 expression were closer to the hIPSC-T state than cells with lower expression at the same day (Figure S2A), confirming the specificity of these surface antigens for reprogramming cells. Notably, the expression patterns of transgene-expressing TRA-1-60^{+} cells isolated at days 20–24 were clearly distinct from those observed after DOX withdrawal and establishment of hIPSC-Ts, as well as from reference human embryonic stem cells (hESCs).

MDS of H3K4me2 (Figure 2D), a histone methylation mark associated with open chromatin and active regulatory elements (Zhou et al., 2011), suggested two major transitions in its distribution: one occurring during the first 5 days and a second coinciding with the acquisition of TRA-1-60. In contrast, genome-wide DNA methylation patterns (Figure 2E) appeared to remain largely constant throughout the first 14 days of the

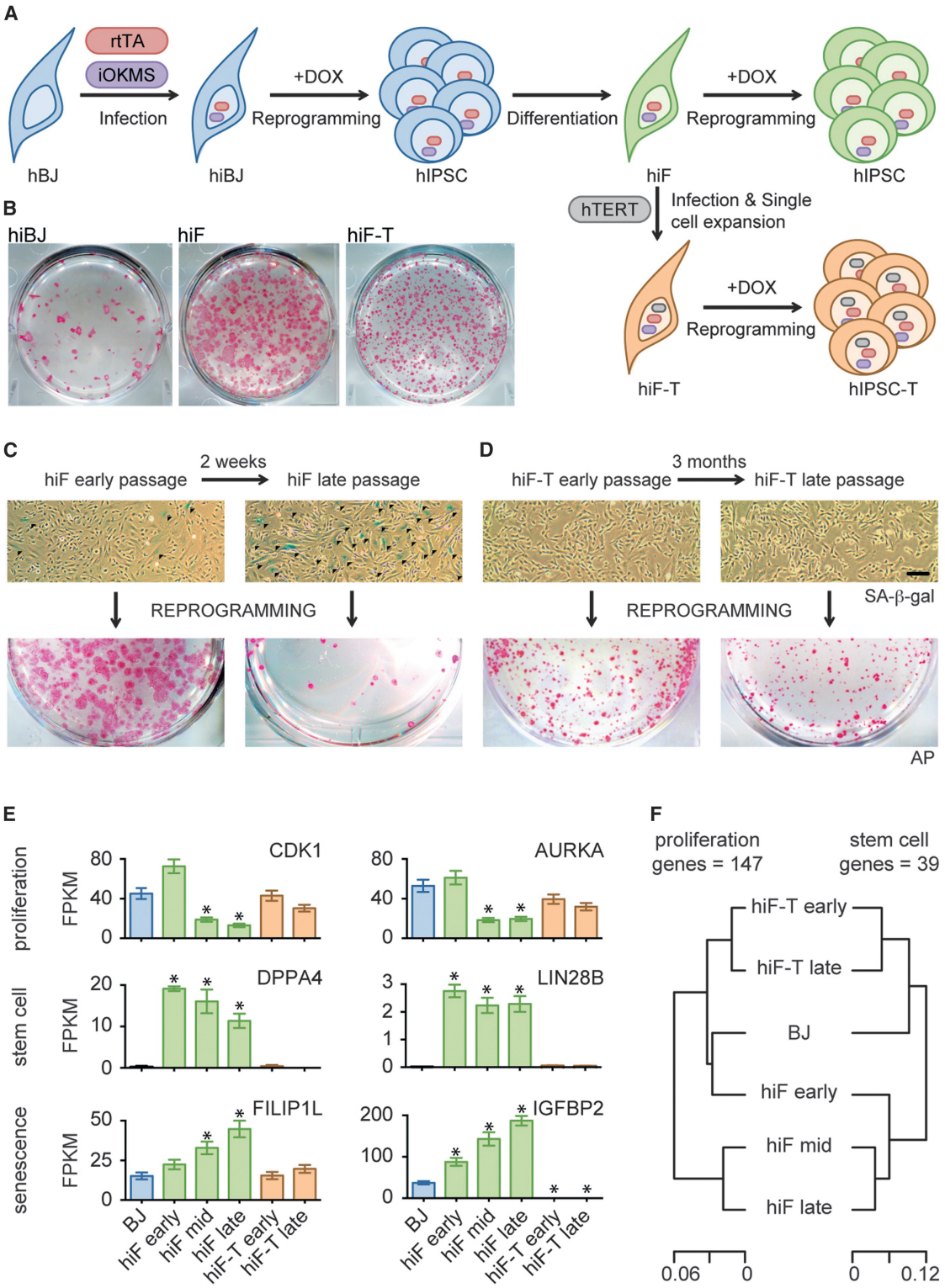


Figure 1. An Optimized Secondary Reprogramming System for Human Reprogramming

(A) Schematic representation of secondary reprogramming strategy and hiF-T engineering using inducible reprogramming factors (iOKMS) under the control of the reverse tetracycline transactivator (rtTA). Fibroblasts and IPSCs are shown as belonging to a primary hiBJ (blue), derived secondary hiF (green) and hiF-hTERT (brown) reprogramming system.

(B) Alkaline phosphatase (AP) staining of the indicated reprogrammed cells.

(legend continued on next page)

reprogramming course but then changed rapidly upon TRA-1-60 acquisition and again upon removal of DOX and derivation of hIPSC-T lines, eventually reaching a pattern equivalent to that of reference hESCs.

The different genome-wide profiles therefore all suggested that TRA-1-60$^+$ cells obtained at the end of the reprogramming course were qualitatively different from both derived hIPSC-T cell lines and reference hESCs when maintained under standard culture conditions. In the following sections, we characterize these differences and the preceding dynamics in more detail.

Ordered Re-activation of Developmental and Pluripotency Pathways

Clustering analysis of the RNA-seq data identified ten major dynamic expression patterns (Figure 3A). We applied two complementary approaches to interpret these clusters: (1) a comparison with gene expression signatures obtained from hESCs in their undifferentiated state or upon differentiation toward the three major germ layers; and (2) gene ontology enrichment analysis for both biological processes and developmental cell identity (Edgar et al., 2013) (Figure S2D; representative genes in Figures 3B and S2B).

Similar to murine reprogramming systems (Mikkelsen et al., 2008; Polo et al., 2012), OKMS rapidly downregulated mesenchymal signature genes, including genes encoding structural components, like collagen, and transcription factors, like SNAI2 and FOSL1. We also observed rapid, but transient, downregulation of genes that are known to be expressed in both stem cells and terminally differentiated cells, most notably the bHLH Inhibitor of DNA-binding (ID) proteins. Pluripotency-related genes were subsequently activated in two waves, distinguishing early pluripotency signature genes already detected by day 5, such as DPPA4, from late core regulators, like LIN28A that were fully activated only in TRA-1-60$^+$ cells at day 20. A final set of genes reached maximal levels only after derivation of hIPSC-T. This set included neuro-ectodermal and epiblast-related factors like SOX3 and OTX2, likely reflecting a priming of human PSCs under standard culture conditions (Tesar et al., 2007).

While rapid downregulation of somatic genes and subsequent activation of the pluripotency network have been described (Chan et al., 2009; Park et al., 2008), comprehensive characterization of the transitions between these states has been limited by the heterogeneity of previous reprogramming systems. We found that the more synchronized changes of the hiF-T system allowed us to characterize several transient waves of gene activation (last panel in Figure 3A).

The first transient wave peaked at day 5 and was enriched for genes promoting proliferation and metabolic changes. Genes characteristic of this wave included IGF2, AFP, GSN, and ALDH1A1, which are known to exert complementary proliferative and/or anti-apoptotic functions. The second wave, which peaked in the SSEA3$^+$ fraction at day 10, included genes expressed during body patterning in late embryogenesis, such as HOX genes (see also Figure 3B), as well as markers of developing mesoderm (e.g., H19) and endoderm (e.g., HNF transcription factors). The third wave brought activation of genes associated with early embryogenesis and primitive endoderm, including NANOG, UTF1, LEFTY2, NODAL, and GDF3. Similar to murine reprogramming, NANOG activation reached its peak prior to activation of the core pluripotency network (Samavarchi-Tehrani et al., 2010). Finally, and in parallel with activation of the core pluripotency network in TRA-1-60$^+$ cells around day 20, we detected substantial expression of pre-implantation- or trophoblast-associated markers (e.g., DPPA2/3/5, DNMT3L, ALPPL2, FGF4, and TFCP2L1) and lowly expressed primitive streak genes (e.g., T, CER1, MIXL1). Notably, these markers were lost upon withdrawal of DOX and derivation of clonal hIPSC-T lines (e.g., DPPA3/STELLA [Hayashi et al., 2008] in Figures 3B, 3C, and S3A). This is consistent with the final step of hIPSC-T derivation from TRA-1-60$^+$ colonies being accompanied by a shift from a gene expression program with pre-implantation-like characteristics to a program with more post-implantation-like characteristics (Figures S3B and S3C).

To examine whether the patterns we observed from bulk RNA profiling actually reflected changes that occurred within individual cells, or were in fact averages over multiple distinct subpopulations, we collected single cell RNA-seq profiles from 52 unfractionated hiF-T cells after 10 days of reprogramming (Figure 3D). We found low expression of genes from the previously defined somatic cluster and high expression of genes from the developmental clusters across the majority of the cells (Figure S3D). Notably, consistent with the bulk profiles, many of the single cells displayed simultaneous expression of genes associated with early and late embryogenesis, as well as LIN28A and other pluripotency markers (Figures 3D and S3D).

To test whether the expression dynamics in hiF-Ts were representative of other reprogramming systems, we also profiled non-immortalized hiFs at multiple time points (Figure S2C). These secondary cells were derived similarly to the frequently used dH1f (Park et al., 2008). We found that all patterns described above, including the late expression of pre-implantation-associated markers, also emerged from the hiFs, although with lower magnitude. Our data therefore suggest a general model of human OKMS-mediated reprogramming where cells first enter a highly proliferative state and lose their somatic identity, then display preferential re-activation of late, followed by early developmental genes, finally leading to the emergence of a distinct transient expression program with pre-implantation-like characteristics in the TRA-1-60$^+$ subpopulation.

(C and D) Representative bright fields of hiF and hiF-T cultures at different passages, after senescence-associated-beta-galactosidase (SA-β-GAL, upper panels) assay and corresponding alkaline phosphatase (AP) staining after 24 days of reprogramming (lower panels). Scale Bar, 100 μm. Senescent cells are stained in blue and indicated with dark arrows.

(E) Histograms showing absolute expression levels of selected proliferation, stemness, and senescence related genes, as measured by RNA-seq. FPKM, fragments per kilobase per million fragments mapped. Error bars represent a 95% confidence interval around the average values. *Significant difference with respect to control BJ at false discovery rate (FDR) <1%.

(F) Hierarchical clustering of BJ, hiF, and hiF-T cells according to expression levels of proliferation (left) or stemness-related (right) genes.

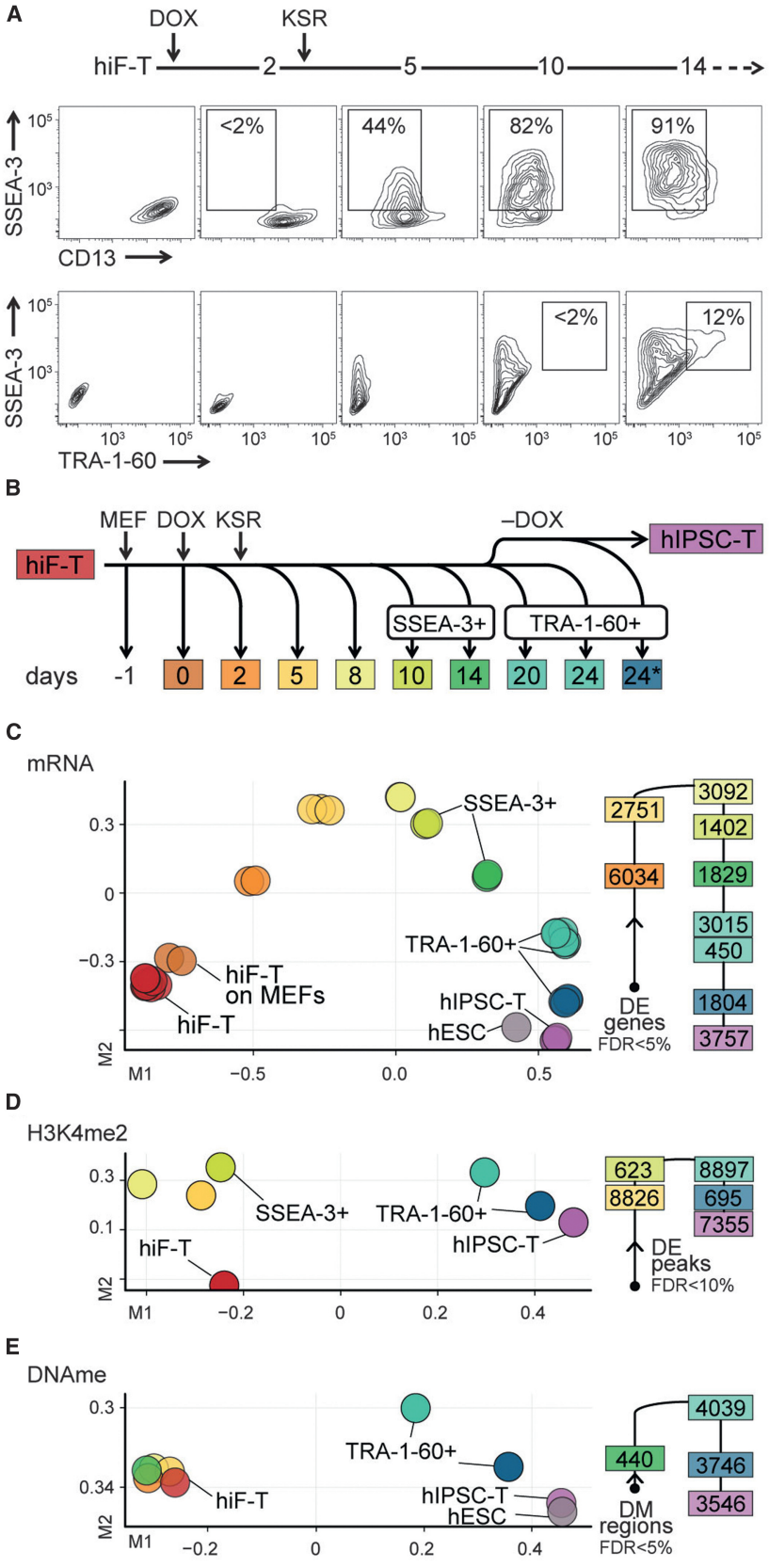


Figure 2. Integrative Analysis of Human Cellular Reprogramming

(A) Flow cytometry analysis of surface markers during reprogramming of hiF-T cells.

(B) Schematic representation of time course collection of reprogramming intermediates, including fractionation by the indicated surface markers. For reprogramming, hiF-T cells were seeded on MEF feeder layer (MEF). Reprogramming was initiated at day 0 (DOX) and a switch to KSR-hESC media was performed at day 2 (KSR). The collection point labeled 24* represents cells reprogrammed for 20 days in DOX followed by 4 days without DOX.

(C) MDS analysis of RNA-seq data (left) along with the number of differentially expressed (DE) genes associated with each transition (right). More comparisons are shown in Figure S2A.

(D and E) MDS analyses of epigenomic data from regions differentially enriched (DE) in H3K4me2 or differentially DNA methylated (DM).

In all the representations, samples are color coded to the reference time points in (B).

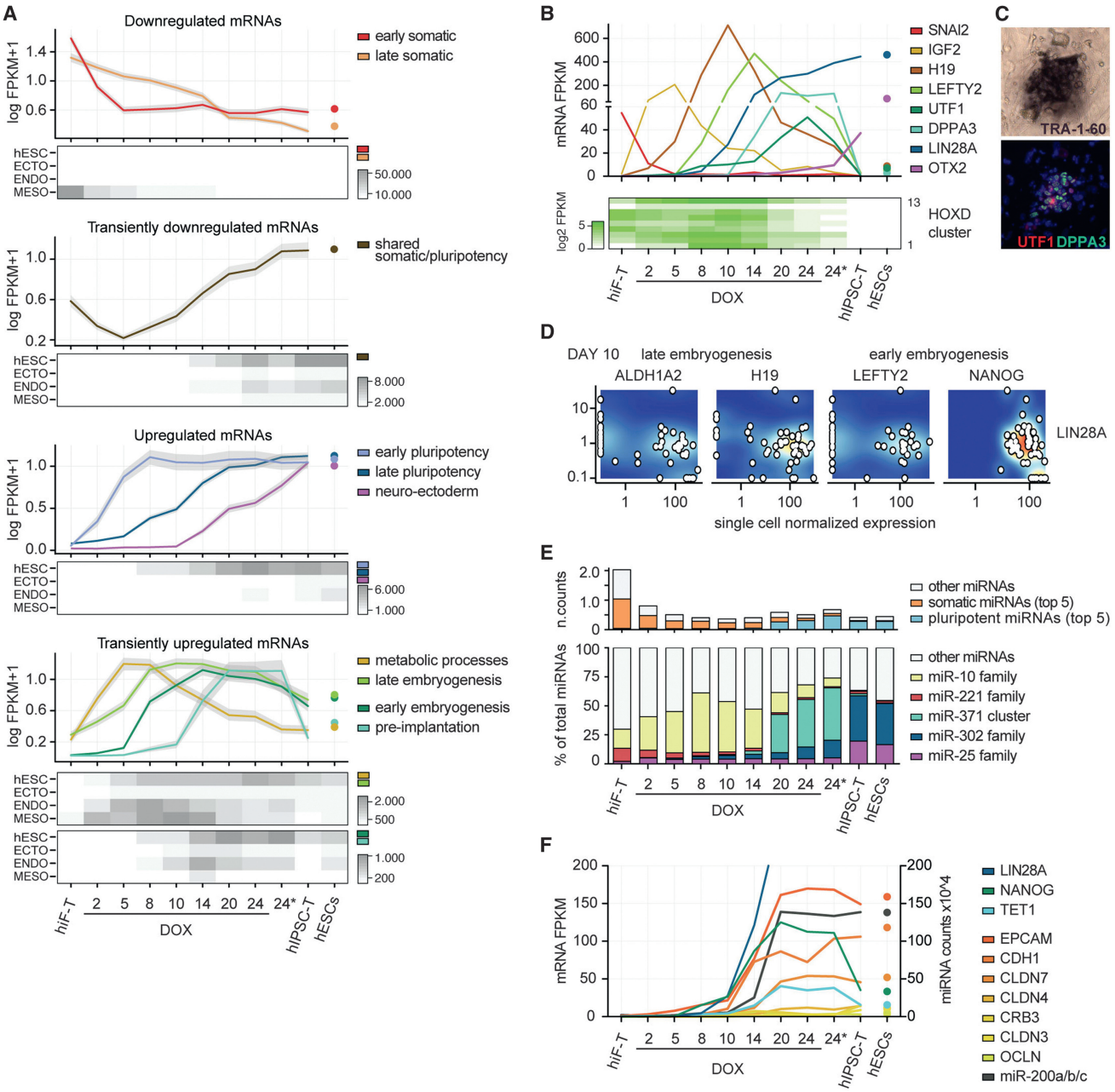


Figure 3. Transcriptional Dynamics during Human Cellular Reprogramming

(A) Line plots showing expression dynamics of differentially expressed genes during reprogramming, grouped by k-medoids clustering. Refer to Figure 2B for reprogramming time points. Gray shades represent a 95% bootstrap confidence interval around the median values. For each cluster, the median expression value in hESC using 18 reference hESC lines is also reported. Heatmaps below each set of clusters (upregulated, downregulated, and transients) show the total expression of genes in each cluster with respect to gene sets that define pluripotent (hESC) or embryonic germ layer-specific cells (ECTOderm, ENDOderm, MESOderm).

(B) Absolute expression levels (FPKM) of selected dynamic genes reported as line plot or heatmap.

(C) Representative colony of reprogrammed hIF-T cells identified by TRA-1-60 chromogenic staining in bright field and overlapping UTF1 and DPPA3 fluorescent staining. Complete field and staining controls are reported in Figure S3A.

(D) Co-expression relationships between representative markers of the identified developmental transitions with respect to LIN28A as reference pluripotent marker, measured by single-cell RNA-seq. Additional single-cell data are shown in Figure S3D.

(E) Absolute expression levels of categorized miRNAs (as normalized counts; see Experimental Procedures) (upper panel) and the relative expression levels of specific miRNA families with respect to total miRNA abundance (lower panel) at the indicated time points and in hESC line HUES64.

(F) Line plot showing absolute expression values of mRNA and miRNA involved in MET during reprogramming and in reference hESCs. Pluripotent genes with different onset during reprogramming (LIN28A, NANOG, TET1) are also included to illustrate the relationship between epithelialization and acquisition of pluripotency.

Analysis of miRNA captured in the small RNA-seq data (Figure 3E) reinforced this model. We similarly observed the rapid loss of somatic miRNAs (e.g., the miR-221 family), followed by upregulation of miRNAs under developmental control (e.g., miR-10) and eventually of pluripotency-associated families (the miR-371 cluster and the miR-302 family). Strikingly, while many different miRNA families were expressed at comparable levels in hiF-Ts, each phase of the reprogramming process was defined by only one specific miRNA family that alone accounted for almost 50% of the sequenced molecules from that phase. Notably, miR-10b alone represented 49% of total miRNAs sequenced during the reactivation of late mesoderm pathways, which is consistent with its location within the HOXD cluster that is activated during this phase. The miR-371 cluster accounted for the majority of the mature miRNAs in TRA-1-60$^+$ cells after day 20, while the miR-302 family, which is enriched in the post-implantation mouse embryo (Parchem et al., 2014), dominated in derived hIPSC-Ts. Consistent with neuro-ectodermal priming, derived hIPSC-Ts also expressed higher levels of miRNAs associated with neuronal progenitors (e.g., members of miR-25 family) (Nowakowski et al., 2013).

A possible driver of the differences in gene expression between TRA-1-60$^+$ cells in DOX-containing media and derived hIPSC-Ts is elevated expression of KLF4 in the former. While the combined expression levels of the lentiviral and endogenous OCT4, SOX2, and c-MYC genes were largely similar, KLF4 expression was ~100-fold higher in the TRA-1-60$^+$ cells (Figure S2B). This would be consistent with recent reports that higher expression of KLF proteins can push IPSC colonies to a more pre-implantation-like state (Gafni et al., 2013; Takashima et al., 2014).

Interestingly, while a mesenchymal to epithelial transition (MET) has been reported as a critical early event in reprogramming of mouse cells (Li et al., 2010; Samavarchi-Tehrani et al., 2010), we did not observe clear enrichment of epithelial associations in the early phases of hiF-T reprogramming in our gene ontology analysis. We therefore looked specifically for the activation of key epithelial marker genes (Figure 3F). In mouse cells, activation of these markers precedes activation of Nanog and the core pluripotency network (Samavarchi-Tehrani et al., 2010). In contrast, during hiF-T reprogramming, these markers appeared to be activated at the same time as NANOG, LIN28A, TET1, and other core regulators of pluripotency.

A Pre-implantation-like Chromatin State Is Transiently Established during Reprogramming

A characteristic feature of pluripotent cell lines is that the promoters of many developmental regulatory genes display a bivalent chromatin state, where histone marks associated both with activation (H3K4me2/me3) and repression (H3K27me3) co-occur (Zhou et al., 2011). To study the acquisition of bivalency during hiF-T reprogramming, we focused on 6,615 promoters that showed a significant change in either H3K4me3 or H3K27me3 signal between day 0 and the iPSC state (Figures 4A and 4B). As expected, bivalent promoters were rare at the somatic and early stages. The reprogrammed TRA-1-60$^+$ subpopulation at 24 days displayed a significant number of bivalent promoters, but only approximately half of that found in clonal hIPSC-Ts derived from them. Interestingly, other groups have recently reported that an increase in bivalency is a key feature of the transition between naive and primed states in ESCs (Marks et al., 2012; Theunissen et al., 2014), although the extent to which this reflects epigenetic changes occurring during human embryogenesis remains unknown.

The distinction between TRA-1-60$^+$ cells and established hIPSC-T was, however, further clarified by analysis of promoter DNA methylation. Consistent with the genome-wide MDS analysis (Figure 2E), we found that the vast majority of changes in promoter DNA methylation took place during the late stages of reprogramming (Figures 2E and 4C). A comparison of our data with methylation profiles from human blastocysts and reference hESCs (Smith et al., 2014) allowed us to identify 722 differentially methylated promoters that each followed one of three major patterns: (1) promoters with high methylation in fibroblasts that were hypomethylated in all reprogrammed or pluripotent samples (24 days TRA-1-60$^+$ cells, derivative hIPSC-T, in vivo blastocysts and in-vitro derived hESCs); (2) promoters with low to moderate methylation in fibroblasts that became hypermethylated upon hIPSC-T derivation were also hypermethylated in hESCs but showed lower methylation in both 24 days TRA-1-60$^+$ cells and blastocysts; and (3) hypermethylated promoters that were transiently de-methylated in 24 days TRA-1-60$^+$ cells showed low methylation in blastocysts but hypermethylation in both hIPSC-Ts and hESCs. This third pattern included promoters of key pre-implantation markers, such as DNMT3L, DPPA3, and miR-371. Although we did not observe global DNA hypomethylation, the differences in promoter DNA methylation between TRA-1-60$^+$ cells at 24 days of reprogramming and derived hIPSC-Ts are in agreement with those previously described between human pre-implantation blastocyst and derived hESCs (Smith et al., 2014).

Transient Chromatin Remodeling at Lineage-Specific Regulatory Elements

To corroborate our finding that OKMS transiently re-activates diverse developmental pathways prior to acquisition of pluripotency, we next examined changes in H3K4me2, a histone methylation mark associated with both active promoters and enhancers (Zhou et al., 2011).

We detected altered H3K4me2 levels in 26,122 distinct genomic regions throughout hiF-T reprogramming. These regions could be grouped into 14 dynamic clusters (Figures 5A, S4A, and S4B). 4,815 regions (clusters 1 and 2) displayed H3K4me2 in hiF-Ts but lost this mark within 5 days of OKMS induction. 8,794 regions (clusters 9–14) were unmarked until day 10 or later, with clusters 10 and 12 showing maximal signals only after derivation of hIPSC-Ts. The remaining 12,513 regions (clusters 3–8) showed variable patterns of transient H3K4me2, many with maximal signal at day 5.

Comparing these dynamic H3K4me2 patterns to reference chromatin maps from various human cell types and tissues (Kundaje et al., 2015) (Figure 5A), we found that regions that were marked by H3K4me2 in hiF-T cells but rapidly lost this mark upon OKMS induction were largely specific to in-vitro-derived somatic cell lines. Regions that gained H3K4me2 during the early stages of reprogramming often displayed active chromatin

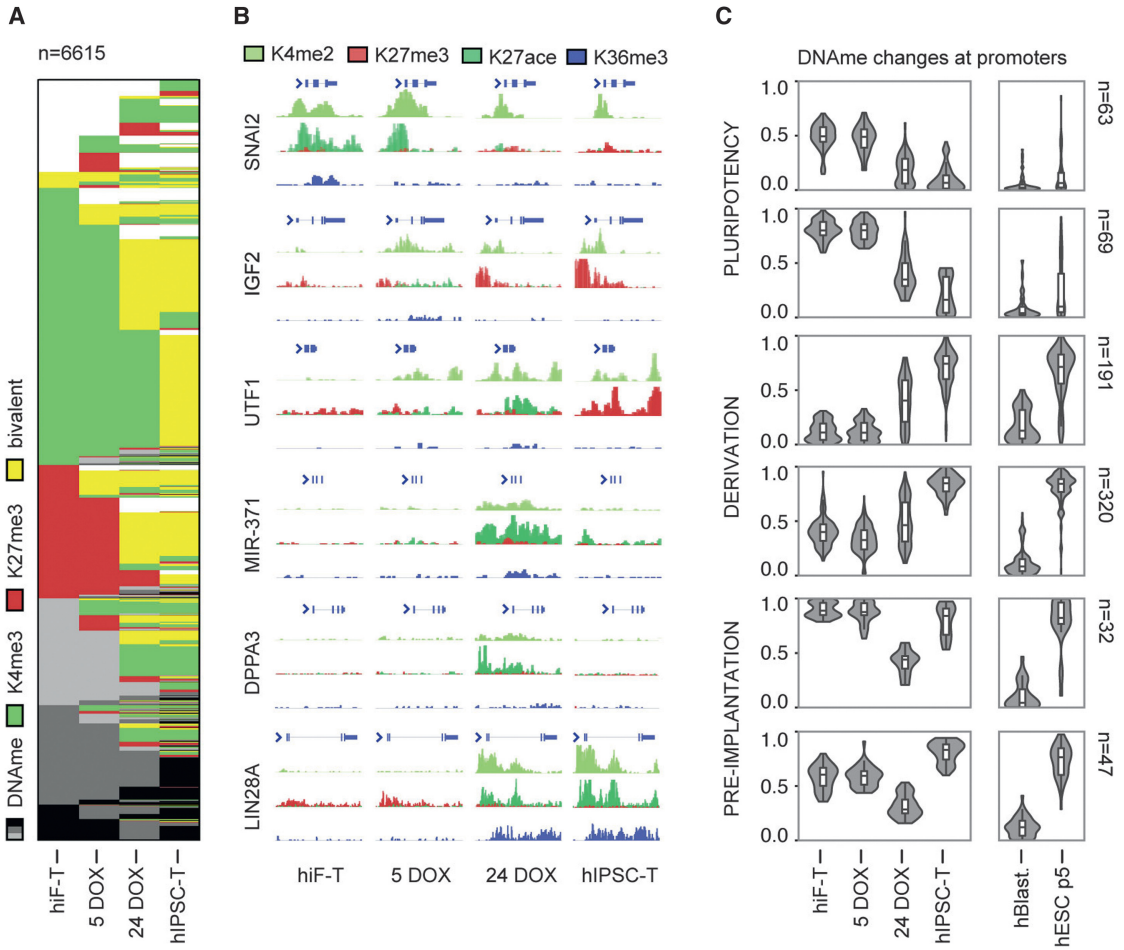


Figure 4. Changes in Bivalency and DNA Methylation during Reprogramming

(A) Chromatin state maps of 6,615 dynamic promoter regions, showing active (H3K4me3), repressed (H3K27me3), or bivalent regions (functionally poised by the co-enrichment of both H3K4me3 and H3K27me3). Promoters with none of these histone marks are marked by different degrees of DNA methylation (DNAme, three shades of gray for the ranges 25%–50%, 50%–75%, and 75%–100%).

(B) Histone methylation at representative 5- to 50-kb loci from the major transcriptional clusters in Figure 3B. A similar map of the broad pre-implantation region around the miR-371 cluster is also reported in Figures S4C and S4D.

(C) Violin plots showing promoter DNA methylation dynamics across the indicated reprogramming time points and sample types, grouped by k-means clustering. n, size of each cluster. The box plots show the first and the third quartiles, along with the medians.

marks in tissues of both mesodermal and ectodermal origin, while regions that gained H3K4me2 in the late stages often also display active chromatin marks in tissues of endodermal origin. We therefore conclude that the transient activation of developmental pathways during reprogramming is accompanied by chromatin remodeling at regulatory elements associated with lineages of different developmental stages.

OKMS Induction Supports Direct Lineage Conversion

The transient re-activation of developmental pathways and regulatory elements suggested to us that OKMS expression might induce epigenetic changes that could enable direct reprogramming to alternative lineages in addition to pluripotency. In support of this, using the Transcription factor Epigenetic Remodeling Activity (TERA) framework (Ziller et al., 2015), we found that the DNA sequences of genomic regions showing changing H3K4me2 were associated with potential binding sites for numerous transcription factor families involved not only in pluripotency, but also in developmental patterning and differentiation (Figures 5B and S5).

The TERA analysis revealed that regions that gained H3K4me2 at early time points frequently coincided with OCT4 and SOX2 binding sites, which is consistent with a more dominant role for these transgenes in chromatin remodeling during the early stages of reprogramming relative to the late stages of IPSC establishment (Soufi et al., 2012). In contrast, regions that gained H3K4me2 at later time points were enriched for potential binding sites for other transcription factors that were upregulated toward the end of the time course, such as ZIC3 and REST. In agreement with our previous analyses, the transient waves of developmental gene expression were accompanied by exposure of binding sites for transcription factors acting in fetal development

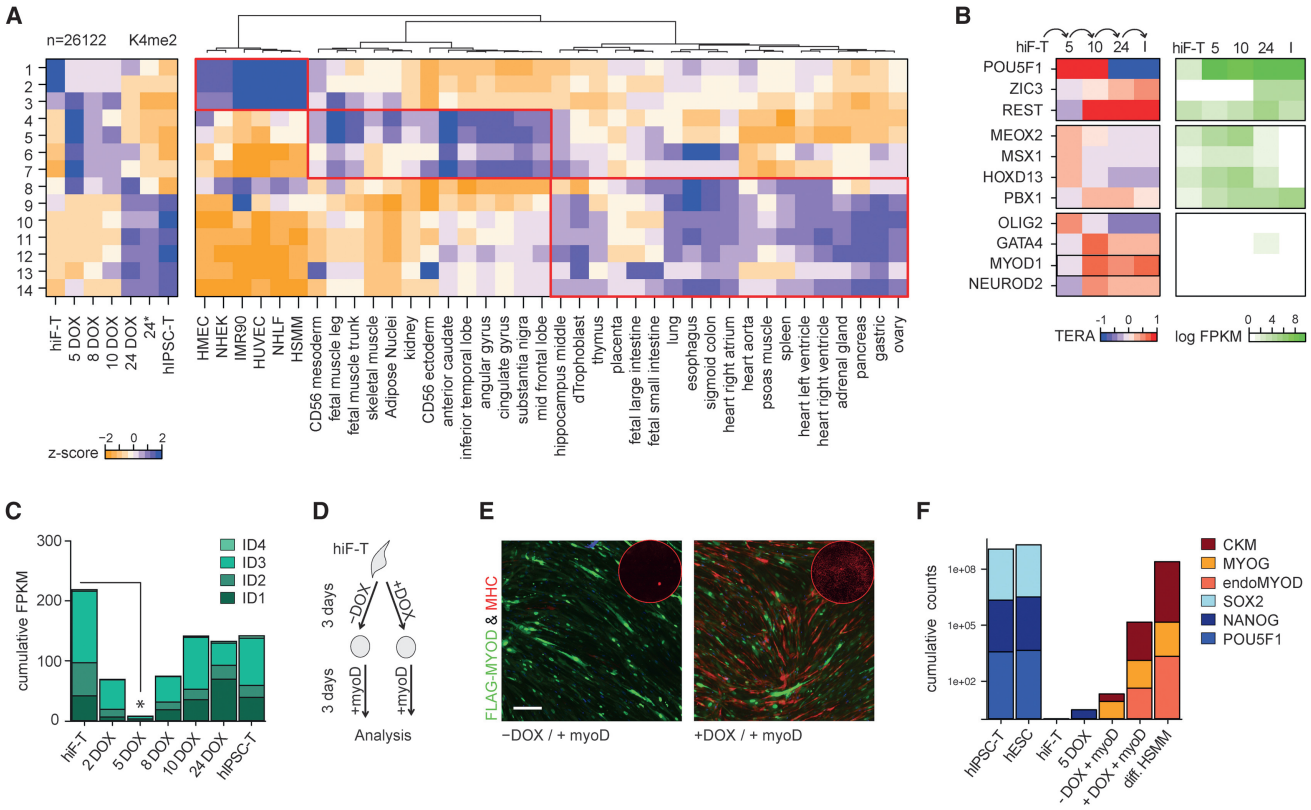


Figure 5. Transient Chromatin Remodeling and Epigenetic Priming during Reprogramming

(A) Left: heatmap showing the *Z* scores of the mean H3K4me2 enrichment in 26,122 dynamic genomic regions, grouped into 14 clusters. For full representation of the clusters and corresponding H3K27me3 dynamics refer to Figures S4A and S4B. Right: heatmap showing the corresponding z-scores of the mean H3K27ac enrichment across tissues of different identity.

(B) Heatmap showing the TERA score of selected transcription factors predicted to be activated during the indicated reprogramming transitions, based on H3K4me2 footprints. Corresponding absolute gene expression values (FPKM) during reprogramming are reported on the right. A full list of the top transcription factors groups and their predicted co-binding relationships are shown in Figure S5.

(C) Bar plot showing the cumulative absolute expression values (FPKM) of the ID gene family during reprogramming.

(D) Schematic representation of the OKMS-enhanced MYOD reprogramming of hiF-Ts.

(E) Representative field of MYOD-mediated myogenic conversion without (–DOX) or with (+DOX) prior OKMS activation for 3 days. Cells positive for ectopic FLAG-MYOD are green while cells positive for the late muscle marker MHC are red. The corresponding MHC whole-well staining is shown in the corner of each condition. Scale Bar, 200 μm.

(F) Bar plot showing the normalized cumulative expression counts of pluripotent (SOX2, NANOG, POU5F1) and muscle -specific genes (CKM, muscle creatine kinase; MYOG, myogenin; endoMYOD, endogenous MYOD). Controls are reference PSCs, hiF-T reprogramming time points, and differentiated human skeletal muscle myoblasts (HSMM).

like HOXD13 and PBX1. Interestingly, some factors which binding sites were also enriched in transiently H3K4me2-marked regions, including bHLH transcription factors MYOD, NEUROD2, and OLIG2, were not detected by RNA-seq at any time point. This enrichment might therefore reflect lineage priming by other factors. In addition, we noticed that OKMS induction led to a transient downregulation of the ID transcriptional repressor family (Figure 5C), which are known to restrict lineage commitment by inhibition of bHLH activity (Perk et al., 2005).

To test our hypothesis of epigenetic priming toward alternative fates, we performed ectopic expression of MYOD, a master regulator of skeletal muscle cell fate. MYOD was first discovered as a factor that reprogrammed somatic cells toward a myotube fate (Fong and Tapscott, 2013), although the efficiency of this phenomenon differs across cell types (including fibroblasts subtypes; Salvatori et al., 1995), due to variation in epigenetic states and expression of inhibitors such as ID1 (Perk et al., 2005). We found that introduction of a MYOD lentivirus led to a very low rate of myogenic conversion of otherwise unperturbed hiF-T cells, as judged by low expression of skeletal muscle markers at both the protein and mRNA levels (Figures 5D–5F). In contrast, introduction of MYOD after a 3-day pulse of OKMS expression drove efficient conversion to myosin heavy chain (MHC) positive cells and induced key muscle genes to levels that approached those of mature skeletal muscle cells within another 3 days. Notably, introduction of MYOD after OKMS induction uniquely activated endogenous MYOD, which may support stabilization of the converted state through its auto-regulation (Hanna et al., 2010). The rapid rate of conversion and the complete absence of pluripotency markers throughout the time course (Figure 5F)

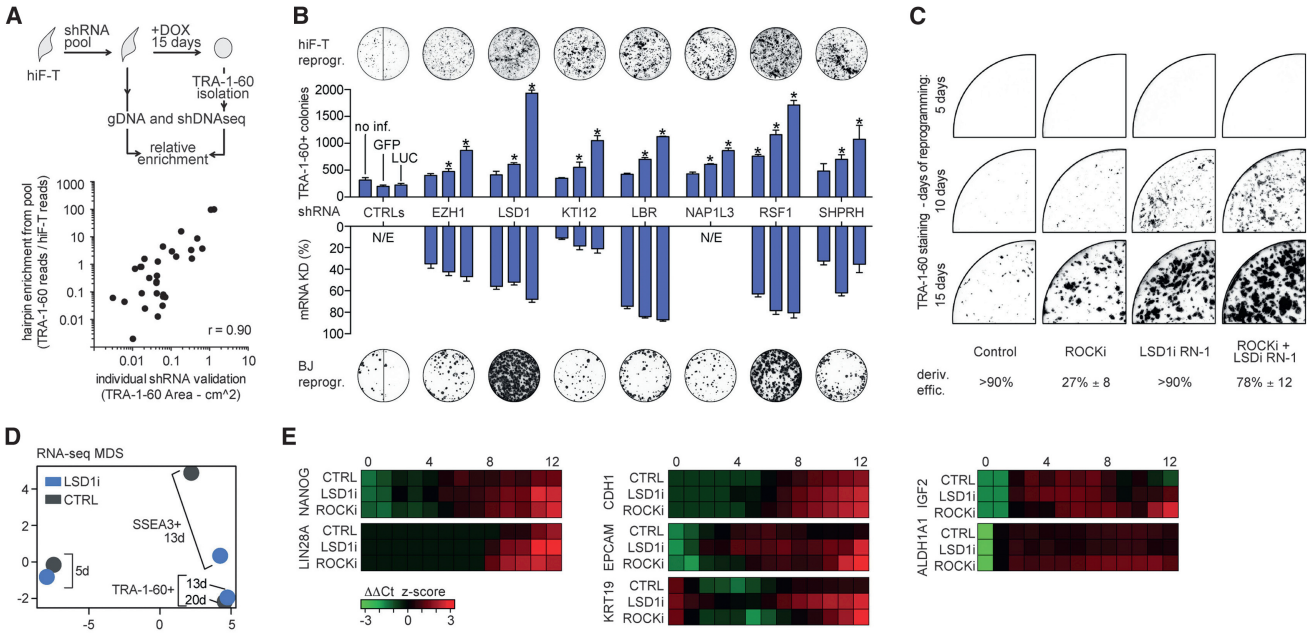


Figure 6. Characterization of Negative Regulators of Reprogramming

(A) Schematic representation of the pooled shRNA screening strategy.

(B) Scatter plot showing comparison of selected reprogramming efficiencies in shRNA-perturbed hIF-T cells at day 15 in a pooled screening format (y axis, enrichment of shRNA sequence reads from TRA-1-60$^+$ cells versus cells prior to induction of reprogramming) versus an arrayed format (x axis, area of TRA-1-60$^+$ colonies). The reported values are the mean of biological duplicates.

(C) Bar plots showing reprogramming efficiency (number of TRA-1-60$^+$ colonies) upon shRNA-mediated perturbation of candidate regulators (upper histogram) and the corresponding change in mRNA expression levels in hIF-T cells relative to the effect of a control shRNA targeting luciferase (LUC) mRNA (lower histogram). Additional controls are shRNAs targeting GFP mRNA or uninfected cells. Three distinct hairpins were tested for each gene and representative TRA-1-60 stainings for each shRNA group are displayed above each set (control lane shows wells of both shGFP and shLUC treatments). Similar effects were observed in primary BJ reprogramming experiments, as shown by TRA-1-60 stainings below each set. Error bars indicate SD from the average. All reported values in histograms are significant with respect to controls at FDR <5%. *Significant difference with respect to control at FDR <1%. Representative TRA-1-60 staining at indicated time points of hiF-T reprogrammed in the continuous presence of the indicated inhibitors.

(D) Gene expression profiles, from RNA-seq, of hiF-T reprogramming with or without LSD1 inhibition (LSD1i and CTRL respectively), represented as points in two-dimensional MDS component space. A gene set enrichment analysis is described in Figure S6C.

(E) Heatmap showing expression (normalized z-score) of indicated genes during reprogramming in different conditions (untreated control, LSD1 inhibition, ROCK inhibition). Key time points for reprogramming transitions are indicated (0, 4, 8, 12). Corresponding MDS plot of the RNA-seq time course utilized to identify these genes is reported in Figure S6D.

strongly suggest that it did not involve transition through an intermediate pluripotent state. We therefore conclude that OKMS rapidly drives cells into an epigenetic state that is conducive not only to derivation of pluripotent cells, but also to direct conversion into alternate lineages.

Identification of Regulatory Genes that Inhibit Reprogramming

We next attempted to identify regulatory factors that might inhibit or delay reprogramming to pluripotency. We again leveraged the expansion capacity and uniformity of hiF-Ts to perform a quantitative RNAi screen (Luo et al., 2008) using a pooled lentiviral library encoding ~2,900 small hairpin RNAs (shRNAs) targeting 370 distinct regulatory genes with known or putative roles in chromatin remodeling and other epigenetic processes. The library complexity was chosen to maximize the representation of neutral shRNAs in the final TRA-1-60$^+$ population despite the bottleneck imposed by limited reprogramming efficiency (see Experimental Procedures).

We noted that hiF-T cultures infected with the lentiviral pool generated TRA-1-60$^+$ colonies much more efficiently than controls. To identify the genes for which knockdown led to enhanced reprogramming, we compared shRNA abundances before (hiF-T) and after (TRA-1-60$^+$) reprogramming using deep sequencing (Figure 6A). This primary screen identified 23 candidate genes with at least one strongly enriched hairpin. In a secondary screen, we confirmed seven of these candidates (EZH1, KTI12, LBR, NAP1L3, RSF1, SHPRH, and LSD1) based on consistent phenotypes from three distinct shRNAs (Figure 6B). The majority of the validated regulators have not been well characterized, but they span a variety of different chromatin modification and remodeling processes. Six of the seven have to our knowledge not been reported as barriers to human reprogramming (Onder et al., 2012). The positive effect of inhibiting the Polycomb complex member EZH1 is surprising given that its homolog EZH2 is required for reprogramming (Onder et al., 2012), but we note that the two appear to regulate different targets (Shen et al., 2008). Inhibition of the histone lysine demethylase LSD1

(KDM1A) has been reported to enhance reprogramming (Li et al., 2009), but its mode of action remains unknown.

Comparison of the Effects of LSD1 and ROCK1 Inhibition on Reprogramming

As LSD1 was the only validated hit in our RNAi screen with available chemical inhibitors, we sought study the effects of its perturbation in more detail. We found that inhibition of LSD1 with either a standard inhibitor (parnate) or a potent analog (Histone Lysine Demethylase Inhibitor RN-1; Neelamegam et al., 2012) dramatically enhanced hiF-T reprogramming. A 10-nM dose of RN-1 over the first 5 days was sufficient to both accelerate and increase the efficiency of reprogramming (Figure S6A), generating TRA-1-60$^+$ cells in less than 10 days (Figure 6C), as opposed to at least 15 days in untreated cultures. This effect was higher than the previously reported effect of ROCK1 inhibitor Y-27632, which also requires much higher concentrations ($\geq$1 μM). In addition, chemical inhibition of LSD1 enhanced reprogramming even in the presence of saturating doses of the ROCK1 inhibitor (Figure S6B), which suggests synergistic modulation of distinct pathways.

Notably, TRA-1-60$^+$ cells from LSD1i-treated cultures at day 13 were indistinguishable from those from untreated cultures from day 20, as judged by RNA-seq profiling (Figure 6D), and could be used to generate stable hIPSC-T clones. The efficiency of hIPSC-T derivation from LSD1i-treated cells was significantly higher than that of ROCK1i-treated cells, which indicates more consistent and complete reprogramming (Figure 6C, lower panel). Moreover, the accelerated reprogramming caused by LSD1 inhibition did not appear to simply be a result of increased proliferation or decreased apoptosis (Figure S6C), as has been suggested for ROCK1 inhibition (Watanabe et al., 2007). On the contrary, the replication rate of LSD1i-treated hiF-T is not affected in short-term cultures.

To gain deeper insights into the effects of LSD1 and ROCK1 inhibition, we collected additional data from each of the first 12 days of a new reprogramming experiment with LSD1i- and ROCK1i-treated and -untreated hiF-T cells using RNA-seq. We were particularly interested in whether the gene expression changes in the treated cells followed the same trajectory as untreated cells. MDS analysis suggested that this was indeed the case (Figure S6D). Both LSD1i and ROCK1i treatments led to the similar patterns of rapid downregulation of somatic genes, followed by transient upregulation of various developmental and embryonic genes. The expression of early and late markers of pluripotency, such as NANOG and LIN28A, were not accelerated with respect to the control cells in early time points, but they became significantly upregulated with respect to untreated cells by day 8 (Figure 6E).

Interestingly, the only significant change in gene expression unique to LSD1i-treated cells in the early stages of the reprogramming was an accelerated upregulation of a small set of genes enriched in epithelial markers, most notably e-cadherin (CDH1), EPCAM, KRT19, and CLDN10. CDH1 is known to be both a major driver of the MET process and essential for maintenance of the pluripotent state (Samavarchi-Tehrani et al., 2010). In fact, CDH1, EPCAM, and KRT19 are among the most highly expressed epithelial genes in human PSCs (data not shown). This suggests that LSD1 inhibition may enhance reprogramming through epithelialization, a phenomenon that we found to be delayed under standard conditions (see Figure 3F).

In contrast, the ROCK1i treatment was characterized by elevated and persistent expression of growth promoting genes like IGF2 and ALDH1A1 during the later stages (Figure 6E). This is consistent with increased proliferation and survival but could also explain the lower efficiency of hIPSC-T derivation from ROCK1i-treated cells (Figure 6C, lower panel). IGF2 and ALDH1A has both been found to be highly expressed in cancer stem cells (Baccelli and Trumpp, 2012; Pollak, 2008) and persistent IGF2 expression has recently been reported as a maker of transformation in vivo during cellular reprogramming (Ohnishi et al., 2014). This indicates that prolonged ROCKi treatment may favor the emergence of a highly proliferative and potentially aberrant reprogramming environment.

DISCUSSION

Ever since the first successful attempts to reprogram mouse and human cells were reported (Takahashi and Yamanaka, 2006), it has been clear that there are significant species-specific differences in these processes. The hiF-T secondary reprogramming system now provides a convenient and representative model system for dissection of reprogramming in human cells.

Recent work by Yamanaka and colleagues has shown that reprogramming human cells enter into an early mesendodermal state just prior to acquisition of pluripotency (Takahashi et al., 2014). Extending this observation, we found evidence of multiple transient waves of gene expression changes that begin with a rapid loss of somatic identity, followed by re-activation of early developmental pathways and embryonic patterning genes in the reverse order of normal development, eventually reaching a pre-implantation-like state that is only lost upon derivation of IPSC lines under standard conditions.

In fully committed somatic cells, the reprogramming factors appear to facilitate re-activation of related developmental lineages that reflect their ontology. Thus, fibroblast-like hiF-T cells largely de-differentiate to first express a broad mesodermal signature and then features of early development. We hypothesize that somatic cells of different origins might activate different transient gene sets. Eventually, only a small fraction of those cells will activate the core pluripotency network. A key remaining question is whether passing through either one of these transient cell states is strictly required for successful derivation of human IPSC lines.

Several recent studies have argued that human PSCs can be brought into a more naive state of pluripotency with cocktails of chemical inhibitors and/or transcription factors (Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014). Our data show that OKMS expression alone is sufficient to induce key mRNA and miRNA genes that are uniquely expressed in the pre-implantation embryo, and to reduce DNA methylation of promoters that are known to be hypomethylated in the inner cell mass but hypermethylated in standard hESC cultures. We also show that prior to reaching the stabilized pluripotent state, OKMS expression induces a less restricted epigenetic state that is particularly amenable to direct lineage conversion. This

is consistent with recent reports of derivation of mesodermal and endodermal cell types from the early stages of reprogramming cell populations (Efe et al., 2011; Zhu et al., 2014).

We expect that the consistency and virtually unlimited expansion potential of the hIF-T system will enable new approaches to dissection of human cellular reprogramming.

EXPERIMENTAL PROCEDURES

Cell Culture and Reprogramming

Human fibroblasts were cultured in an optimized DMEM/F12 culture media supplemented with 10% FBS. Pluripotent stem cells were cultured in 20% knockout serum replacement (KSR)-based DMEM/F12 culture media with irradiated MEF feeders or mTeSR1 or Essential 8 media without feeders. Reprogramming was performed on a confluent irradiated MEF layer using the KSR media formulation and doxycycline as indicated. BJ fibroblasts were first reprogrammed with a dox-inducible, polycistronic OKMS lentiviral vector (Addgene 51543). This gave rise to the first hIPSC line, which was then differentiated in vitro (Park et al., 2008) to obtain the hiF line. Infection of the hiF line with a CMV-hTERT lentivirus (Applied Biological Materials) and clonal isolation generated the final hiF-T. Directed differentiation of hIPSC-T was performed as previously reported (Gifford et al., 2013). Additional details of cell culture and media formulations, reprogramming, and sampling are reported in the Supplemental Experimental Procedures.

Cellular Assays

Senescence (beta-galactosidase), alkaline phosphatase, flow cytometry, and immunostaining analyses were performed with commercial kits and antibodies, as detailed in the Supplemental Experimental Procedures. Cell proliferation, senescence, and apoptosis during fibroblasts expansion and reprogramming efficiency were assessed using manual cell or colony counting as indicated. For quantitative analysis of reprogramming efficiency in some comparative approaches (RNAi screening, LSD1, and ROCK1 inhibition), digital acquisition of chromogenic TRA-1-60 staining was performed and followed by colony identification and counting by ImageJ (http://imagej.nih.gov/ij/). Further details are reported in Supplemental Experimental Procedures.

Genomic and Epigenomic Profiling

Cells were prepared for profiling using MEF depletion and, in some instances, SSEA-3 or TRA-1-60 enrichment/depletion using magnetic beads separation (Miltenyi Biotec) as indicated in the text. Bulk mRNA-seq and small RNA-seq were performed using TruSeq kits (Illumina). Single cell RNA-seq was performed using the Smart-Seq2 protocol with minor modifications. RRBS and ChIP-seq were performed as previously described (Mikkelsen et al., 2010; Boyle et al., 2012). Assessment of MYOD-mediated direct differentiation was performed using a NanoString nCounter with a custom codeset. High-throughput 3′ digital gene expression (DGE) was performed using a modified single-cell RNA barcoding sequencing (SCRB-seq) protocol with barcoded poly-dT RT primers and a hybrid Nextera/TruSeq sequencing strategy. Details of all the library construction and sequencing procedures as well as downstream bioinformatics analyses are reported in Supplemental Experimental Procedures.

RNAi Screening

RNAi screening was performed by infecting at least 6×10^7 hiF-T cells with a pool of The RNAi Consortium (TRC) lentiviral shRNA constructs targeting 370 distinct epigenetic regulators and then reprogramming the infected cells for 15 days. Integrated shRNA templates were recovered from TRA-1-60$^+$ cell gDNA by PCR and counted using Illumina sequencing. Further details are reported in Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for all sequencing-derived data reported in this paper is GEO: GSE62777.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.06.016.

AUTHOR CONTRIBUTIONS

T.S.M. and D.C. conceived the experimental plan and wrote the manuscript. D.C. derived and reprogrammed the hiF and hiF-T cell lines. S.R., Z.W., V.A., C.A.G., and J.D. assisted with cell culture and screening. D.C., M.S., M.C., J.D., and X.Z. performed library preps and sequencing. C.T., M.J.Z., D.C., R.K., Z.D.S., and S.J.H.S. processed and analyzed data. J.L.R., G.Q.D., A.M., E.S.L., and T.S.M. provided mentoring and assisted with data interpretation.

ACKNOWLEDGMENTS

We are grateful to Martin Sauvageau and Alexander Tsankov for technical assistance and for helpful discussion. D.C. is a Human Frontier Science Program Fellow. C.T. is a Damon Runyon Fellow. M.C. is an EMBO and Leukemia and Lymphoma Society Fellow. M.S. was supported by a Swiss National Science Foundation postdoctoral fellowship, and D.C., M.S., S.J.H.S., and T.S.M. were supported by the Harvard Stem Cell Institute. A.M. is a New York Stem Cell Foundation, Robertson Investigator. This work was supported by grants from NIGMS to T.S.M., A.M., and J.L.R. (P01GM099117) and G.Q.D. (R01GM107536).

Received: November 11, 2014
Revised: March 18, 2015
Accepted: June 1, 2015
Published: July 16, 2015

REFERENCES

Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. Nature *502*, 462–471.

Baccelli, I., and Trumpp, A. (2012). The evolving concept of cancer and metastasis stem cells. J. Cell Biol. *198*, 281–293.

Boyle, P., Clement, K., Gu, H., Smith, Z.D., Ziller, M., Fostel, J.L., Holmes, L., Meldrim, J., Kelley, F., Gnirke, A., and Meissner, A. (2012). Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. Genome Biol. *13*, R92.

Chan, E.M., Ratanasirintrawoot, S., Park, I.-H., Manos, P.D., Loh, Y.-H., Huo, H., Miller, J.D., Hartung, O., Rho, J., Ince, T.A., et al. (2009). Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat. Biotechnol. *27*, 1033–1037.

Edgar, R., Mazor, Y., Rinon, A., Blumenthal, J., Golan, Y., Buzhor, E., Livnat, I., Ben-Ari, S., Lieder, I., Shitrit, A., et al. (2013). LifeMap Discovery™: the embryonic development, stem cells, and regenerative medicine research portal. PLoS ONE *8*, e66629.

Efe, J.A., Hilcove, S., Kim, J., Zhou, H., Ouyang, K., Wang, G., Chen, J., and Ding, S. (2011). Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. Nat. Cell Biol. *13*, 215–222.

Fong, A.P., and Tapscott, S.J. (2013). Skeletal muscle programming and reprogramming. Curr. Opin. Genet. Dev. *23*, 568–573.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282–286.

Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R., et al. (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. Cell *153*, 1149–1163.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell *143*, 508–525.

Hayashi, K., Lopes, S.M.C. de S., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*, 391–401.

Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M., and Jaenisch, R. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell Stem Cell *3*, 346–353.

Hussein, S.M.I., Puri, M.C., Tonge, P.D., Benevento, M., Corso, A.J., Clancy, J.L., Mosbergen, R., Li, M., Lee, D.-S., Cloonan, N., et al. (2014). Genome-wide characterization of the routes to pluripotency. Nature *516*, 198–206.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al.; Roadmap Epigenomics Consortium (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Li, W., Zhou, H., Abujarour, R., Zhu, S., Young Joo, J., Lin, T., Hao, E., Schöler, H.R., Hayek, A., and Ding, S. (2009). Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. Stem Cells *27*, 2992–3000.

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.

Luo, B., Cheung, H.W., Subramanian, A., Sharifnia, T., Okamoto, M., Yang, X., Hinkle, G., Boehm, J.S., Beroukhim, R., Weir, B.A., et al. (2008). Highly parallel identification of essential genes in cancer cells. Proc. Natl. Acad. Sci. USA *105*, 20380–20385.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A., and Stunnenberg, H.G. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590–604.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Mikkelsen, T.S., Xu, Z., Zhang, X., Wang, L., Gimble, J.M., Lander, E.S., and Rosen, E.D. (2010). Comparative epigenomic analysis of murine and human adipogenesis. Cell *143*, 156–169.

Neelamegam, R., Ricq, E.L., Malvaez, M., Patnaik, D., Norton, S., Carlin, S.M., Hill, I.T., Wood, M.A., Haggarty, S.J., and Hooker, J.M. (2012). Brain-penetrant LSD1 inhibitors can block memory consolidation. ACS Chem. Neurosci. *3*, 120–128.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.

Nowakowski, T.J., Fotaki, V., Pollock, A., Sun, T., Pratt, T., and Price, D.J. (2013). MicroRNA-92b regulates the development of intermediate cortical progenitors in embryonic mouse brain. Proc. Natl. Acad. Sci. USA *110*, 7056–7061.

Ohnishi, K., Semi, K., Yamamoto, T., Shimizu, M., Tanaka, A., Mitsunaga, K., Okita, K., Osafune, K., Arioka, Y., Maeda, T., et al. (2014). Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. Cell *156*, 663–677.

Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., Cahan, P., Marcarci, B.O., Unternaehrer, J., Gupta, P.B., et al. (2012). Chromatin-modifying enzymes as modulators of reprogramming. Nature *483*, 598–602.

Parchem, R.J., Ye, J., Judson, R.L., LaRussa, M.F., Krishnakumar, R., Blelloch, A., Oldham, M.C., and Blelloch, R. (2014). Two miRNA clusters reveal alternative paths in late-stage reprogramming. Cell Stem Cell *14*, 617–631.

Park, I.-H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. Nature *451*, 141–146.

Perk, J., Iavarone, A., and Benezra, R. (2005). Id family of helix-loop-helix proteins in cancer. Nat. Rev. Cancer *5*, 603–614.

Pollak, M. (2008). Insulin and insulin-like growth factor signalling in neoplasia. Nat. Rev. Cancer *8*, 915–928.

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.

Salvatori, G., Lattanzi, L., Coletta, M., Aguanno, S., Vivarelli, E., Kelly, R., Ferrari, G., Harris, A.J., Mavilio, F., Molinaro, M., et al. (1995). Myogenic conversion of mammalian fibroblasts induced by differentiating muscle cells. J. Cell Sci. *108*, 2733–2739.

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.

Shen, X., Liu, Y., Hsu, Y.-J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.-C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol. Cell *32*, 491–502.

Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. Nature *511*, 611–615.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994–1004.

Stadtfeld, M., Maherali, N., Borkent, M., and Hochedlinger, K. (2010). A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nat. Methods *7*, 53–55.

Stewart, S.A., Ben-Porath, I., Carey, V.J., O'Connor, B.F., Hahn, W.C., and Weinberg, R.A. (2003). Erosion of the telomeric single-strand overhang at replicative senescence. Nat. Genet. *33*, 492–496.

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., Sasaki, A., Yamamoto, M., Nakamura, M., Sutou, K., Osafune, K., and Yamanaka, S. (2014). Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. Nat. Commun. *5*, 3678.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell *158*, 1254–1269.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell *15*, 471–487.

Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature *460*, 1145–1148.

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S., Muguruma, K., and Sasai, Y. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat. Biotechnol. *25*, 681–686.

Wernig, M., Lengner, C.J., Hanna, J., Lodato, M.A., Steine, E., Foreman, R., Staerk, J., Markoulaki, S., and Jaenisch, R. (2008). A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nat. Biotechnol. *26*, 916–924.

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. *12*, 7–18.

Zhu, S., Rezvani, M., Harbell, J., Mattis, A.N., Wolfe, A.R., Benet, L.Z., Willenbring, H., and Ding, S. (2014). Mouse liver repopulation with hepatocytes generated from human fibroblasts. Nature *508*, 93–97.

Ziller, M.J., Edri, R., Yaffe, Y., Donaghey, J., Pop, R., Mallard, W., Issner, R., Gifford, C.A., Goren, A., Xing, J., et al. (2015). Dissecting neural differentiation regulatory networks through epigenetic footprinting. Nature *518*, 355–359.

Cell Stem Cell
Resource

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

Mohammad A. Mandegar,[1,*] Nathaniel Huebsch,[1,2] Ekaterina B. Frolov,[1] Edward Shin,[1] Annie Truong,[1] Michael P. Olvera,[1] Amanda H. Chan,[1] Yuichiro Miyaoka,[1,12] Kristin Holmes,[1] C. Ian Spencer,[1] Luke M. Judge,[1,2] David E. Gordon,[3,4,5] Tilde V. Eskildsen,[6,7] Jacqueline E. Villalta,[3,4,8,9] Max A. Horlbeck,[3,4,8,9] Luke A. Gilbert,[3,4,8,9] Nevan J. Krogan,[3,4,5] Søren P. Sheikh,[6,7] Jonathan S. Weissman,[3,4,8,9] Lei S. Qi,[10] Po-Lin So,[1] and Bruce R. Conklin[1,3,4,11,*]
[1]Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158, USA
[2]Department of Pediatrics, University of California, San Francisco, San Francisco, CA 94158, USA
[3]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA
[4]California Institute for Quantitative Biosciences, QB3, University of California, San Francisco, San Francisco, CA 94158, USA
[5]Gladstone Institute of Virology and Immunology, San Francisco, CA 94158, USA
[6]Department of Cardiovascular and Renal Research, University of Southern Denmark, 5000 Odense C, Denmark
[7]Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, 5000 Odense C, Denmark
[8]Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA
[9]Center for RNA Systems Biology, University of California, San Francisco, San Francisco, CA 94158, USA
[10]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA
[11]Department of Medicine and Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA
[12]Present address: Regenerative Medicine Project, Tokyo Metropolitan Institute of Medical Science, Tokyo, 156-8506, Japan
*Correspondence: mo.mandegar@gladstone.ucsf.edu (M.A.M.), bconklin@gladstone.ucsf.edu (B.R.C.)
http://dx.doi.org/10.1016/j.stem.2016.01.022

SUMMARY

Developing technologies for efficient and scalable disruption of gene expression will provide powerful tools for studying gene function, developmental pathways, and disease mechanisms. Here, we develop clustered regularly interspaced short palindromic repeat interference (CRISPRi) to repress gene expression in human induced pluripotent stem cells (iPSCs). CRISPRi, in which a doxycycline-inducible deactivated Cas9 is fused to a KRAB repression domain, can specifically and reversibly inhibit gene expression in iPSCs and iPSC-derived cardiac progenitors, cardiomyocytes, and T lymphocytes. This gene repression system is tunable and has the potential to silence single alleles. Compared with CRISPR nuclease (CRISPRn), CRISPRi gene repression is more efficient and homogenous across cell populations. The CRISPRi system in iPSCs provides a powerful platform to perform genome-scale screens in a wide range of iPSC-derived cell types, dissect developmental pathways, and model disease.

INTRODUCTION

To understand the biological roles of genes in development and disease, we must decipher the relationships between genotype and phenotype. Until recently, RNAi has been the most commonly used loss-of-function tool to study human biology (Boettcher and McManus, 2015). However, RNAi suffers from off-target effects and incomplete silencing of the desired gene (Jackson et al., 2003; Kim et al., 2013b; Krueger et al., 2007). Alternatively, programmable nucleases, such as zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), allow more precise gene editing in model organisms, particularly in mammalian and human systems (Gaj et al., 2013; Kim and Kim, 2014). While ZFNs and TALENs are efficient tools for targeting single alleles, they cannot be easily used for library-scale loss-of-function studies.

In 2012, clustered regularly interspaced short palindromic repeat (CRISPR) technology emerged as a new tool for gene editing. This technology is a microbial adaptive-immune system that uses RNA-guided nucleases to recognize and cleave foreign genetic elements (Doudna and Charpentier, 2014; Wiedenheft et al., 2012). The recently engineered CRISPR/Cas9 system consists of two components: a single-chimeric guide RNA (gRNA) that provides target specificity and a CRISPR-associated protein (Cas9) that acts as a helicase and a nuclease to unwind and cut the target DNA (Cong et al., 2013; Mali et al., 2013). In this system, the only restriction for targeting a specific locus is the protospacer adjacent motif (PAM) sequence ("NGG" in the case of *Sp*Cas9) (Doudna and Charpentier, 2014).

CRISPR nuclease (CRISPRn) has been used for genome-scale screens to identify essential genes for cell viability in cancer and embryonic stem cells (Shalem et al., 2014) and human leukemic cell lines (Wang et al., 2014, 2015). However, CRISPRn may not be the most robust system for loss-of-function studies, because it is limited by the number of cells within a population that do not produce knockout phenotypes (González et al., 2014). In addition, partial loss- or gain-of-function phenotypes can be generated by Cas9-induced in-frame insertion/deletions (INDELs) and hypomorphic alleles (Shi et al., 2015), which can obscure the readout.

The nuclease deactivated version of Cas9 (dCas9) blocks transcription in prokaryotic and eukaryotic cells (known as CRISPR interference; CRISPRi) (Qi et al., 2013). More recently, dCas9 was fused to the Krüppel-associated box (KRAB) repression domain to generate dCas9-KRAB, producing a

more efficient transcriptional interference (Gilbert et al., 2013, 2014; Kearns et al., 2014). To further this effort, we aimed to use CRISPRi technology to efficiently repress genes to study early differentiation and model disease with human induced pluripotent stem cells (iPSCs) (Takahashi et al., 2007).

iPSCs are well suited to study early embryonic development and disease since they can produce different functional cell types in vitro (Sterneckert et al., 2014). Early embryonic development consists of a series of accurately timed events that affect gene activation and repression (Bolouri and Davidson, 2003). Therefore, precisely regulating the timing and dosage of transcription factors critically affects embryonic development (McFadden et al., 2005; Takeuchi et al., 2011), and dysregulation in the timing and dosage of transcripts can lead to disease development (Theodoris et al., 2015). In this study, we compared inducible CRISPR systems for gene knockout (using Cas9) or knockdown (using dCas9-KRAB) to enable temporal control of loss-of-function phenotypes in iPSCs and differentiated cell types.

RESULTS

Generation of CRISPRi and CRISPRn iPSC Lines

For loss-of-function studies, we independently derived multiple stable CRISPRi and CRISPRn human iPSC clones in two genetic backgrounds: wild-type B (WTB) and wild-type C (WTC) (Miyaoka et al., 2014). In separate targeting events, the CRISPRi and CRISPRn constructs (see Supplemental Experimental Procedures) were integrated into the AAVS1 locus of WTB and WTC iPSCs using a TALEN-assisted gene-trap approach (Figures 1A, 1B, and S1). Transgenes integrated at the AAVS1 locus remain transcriptionally active in both iPSCs and differentiated cell types (Hockemeyer et al., 2011; Lombardo et al., 2011). We generated several different versions of the CRISPRi system that are either inducible or constitutive; the inducible CRISPRi (Gen1 and Gen2) clones express dCas9-KRAB (KRAB domain fused at the N terminus) from the inducible TetO promoter, while the constitutive CRISPRi clones (Gen3) express dCas9-KRAB under the constitutively active CAG promoter. The CRISPRn (Gen1) clones express Cas9 under the inducible TetO promoter (Figure S1).

The average efficiency of forming stable clones was ~350 colonies per million iPSCs transfected with AAVS1 TALENs and donor plasmid (data not shown). From each condition, multiple independent colonies were isolated and expanded. A subset of the stable colonies from each targeting vector was screened using junction PCR. Two putative colonies from each targeting event were further characterized by stably introducing an *OCT4*-specific gRNA and performing knockdown or knockout assays with immunofluorescence and western blot analysis. All putative CRISPRi clones containing an *OCT4*-specific gRNA showed efficient knockdown (>95%) of OCT4 in bulk populations, while a significant fraction of the CRISPRn cells remained OCT4 positive (~30%–40%) in bulk populations containing *OCT4*-specific gRNA (Figure S1). One clone each from CRISPRi and CRISPRn (Gen1 lines in the WTC genetic background) were subsequently used as lead clones for further studies.

To enable non-invasive and high-throughput phenotypic analysis in iPSC-derived cardiomyocytes (iPS-CMs), we performed a second targeting event that introduced the green fluorescent calcium-modulated protein 6 fast type (GCaMP) calcium sensor (Chen et al., 2013) into the other AAVS1 locus of the CRISPRi cell line. The GCaMP transgene is driven off the strong, constitutive CAG promoter (Figure S1). We found that CRISPRi iPSCs could differentiate into iPS-CMs, so that we could measure calcium transients based on the GCaMP-fluorescent intensity (Movie S1) (Huebsch et al., 2015). Lead CRISPRi and CRISPRn iPSCs were karyotypically normal (Figures S2A and S2B) and expressed pluripotency markers, as expected (Figures S2C and S2D).

RNA-sequencing (RNA-seq) analysis indicated that expression of dCas9-KRAB or Cas9 was undetectable in the absence of doxycycline, and addition of doxycycline without any gRNA resulted in robust selective induction of dCas9-KRAB or Cas9, while the rest of the transcriptome remained virtually unchanged (Figures S2E and S2F). Furthermore, the RNA-seq data suggest that the addition of the KRAB domain has no detectable off-target effects when compared to expression of Cas9 alone. Remarkably the one gene that appeared to be upregulated upon doxycycline induction (without gRNA) was the same gene (Vimentin; *VIM*) for both CRISPRi and CRISPRn cells (Figures S2E and S2F). Since the same gene is upregulated for CRISPRi and CRISPRn cells, we suspect it may represent an off-target activity of the doxycycline-induced transactivator. Importantly, our experiments suggest that the expression of dCas9-KRAB alone has no additional effects on gene expression.

We also expressed dCas9-KRAB and Cas9 by continuously culturing CRISPRi and CRISPRn lines with doxycycline for 3 weeks (four passages). With this long-term treatment, we observed no cytotoxicity, decrease in proliferation, or change in morphology in these cells (Figures S2G and S2H). Using a droplet digital PCR (ddPCR)-based copy-number assay, we measured the number of integration events (Figure S2I). We further validated on-target integration sites on the lead CRISPRi and CRISPRn clones with junction PCR (Figure S2J) and verified their sequences (data not shown).

To further ensure there was no leaky expression of the single doxycycline-inducible vector, we measured the protein levels of dCas9-KRAB and Cas9 in iPSCs. With immunostaining, flow cytometry and western blots did not detect dCas9-KRAB or Cas9 protein without doxycycline in either CRISPRi or CRISPRn iPSCs, indicating that the TetO promoter has high fidelity in the AAVS1 locus. After doxycycline treatment, all cells in the CRISPRi and CRISPRn lines expressed dCas9-KRAB or Cas9 within 48 hr, respectively (Figures 1C–1H). dCas9-KRAB and Cas9 were expressed at similar levels after induction, and both proteins rapidly degraded after removing doxycycline (Figures 1F, 1H, and S2K). These data showed that dCas9-KRAB and Cas9 expression could be tightly regulated with the TetO promoter, which would support studies that rely on precisely timing gene knockdown or knockout.

Comparison of Loss of Function between CRISPRi and CRISPRn

To compare CRISPRi and CRISPRn for loss-of-function studies, we designed a gRNA that targets the first exon of *NANOG*, a transcription factor necessary for maintaining the pluripotency network. We selected *NANOG* as our first target gene because its deficiency is sufficient to give an immediate readout, as

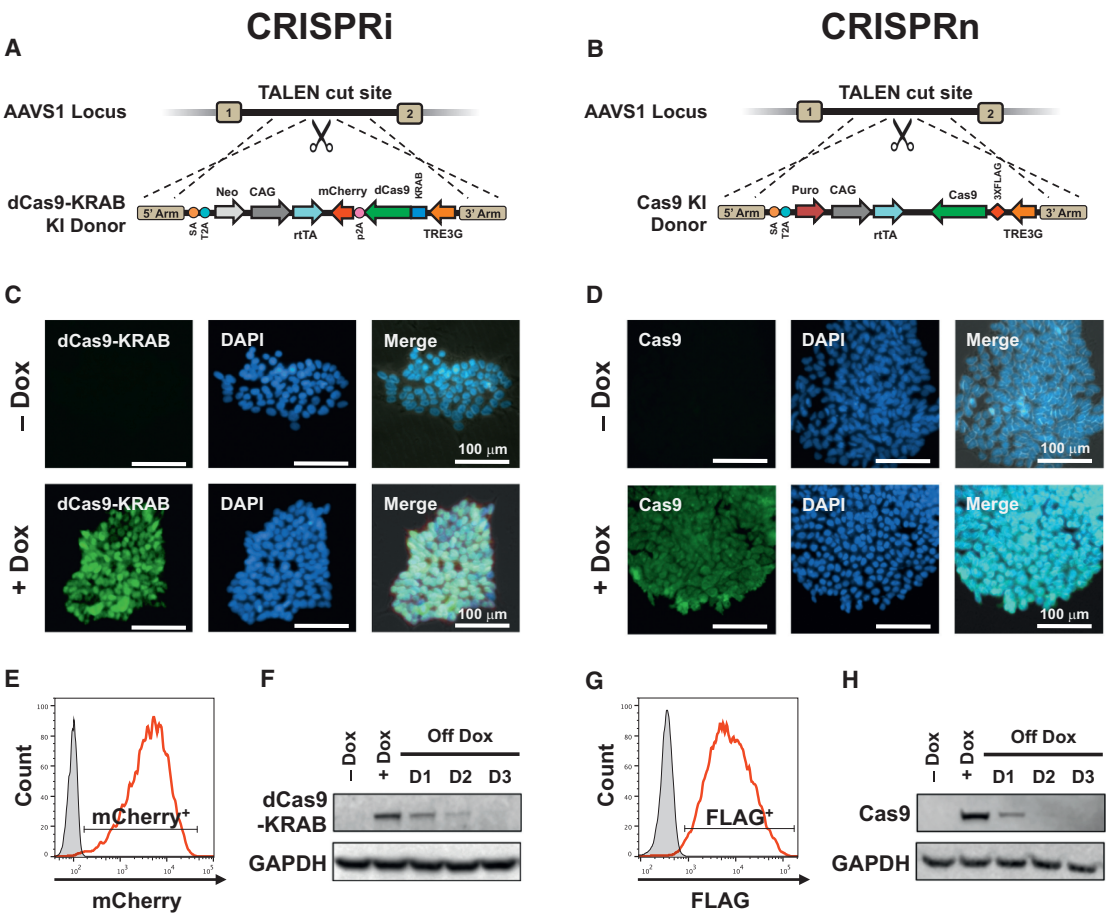


Figure 1. Generation and Characterization of Inducible CRISPRi and CRISPRn iPSCs

(A and B) Schematic overview of the strategy for TALEN-mediated targeting to the AAVS1 locus to generate the CRISPRi and CRISPRn iPSC lines. The doxycycline-controlled reverse transcriptional activator (rtTA) is driven by a strong constitutive promoter (CAG). The third-generation doxycycline-response element (TRE3G) drives transcription of either Cas9 (CRISPRn) or dCas9-KRAB-P2A-mCherry (CRISPRi) and is oriented in the opposite direction of the transactivator to ensure no leaky expression without doxycycline treatment.

(C and D) Immunostaining of CRISPRi and CRISPRn colonies before and after 48 hr of doxycycline treatment with an antibody against Cas9 (green). Nuclei are stained with DAPI (blue). All nuclei showed expression of dCas9-KRAB or Cas9 after adding doxycycline.

(E and G) Flow cytometry analysis of CRISPRi and CRISPRn iPSC lines before and after 48 hr of doxycycline treatment. Doxycycline treatment of CRISPRi and CRISPRn produced expression of mCherry and FLAG in all cells, respectively. The doxycycline-untreated sample is plotted in gray.

(F and H) CRISPRi and CRISPRn iPSC lines were treated with doxycycline (2 μM) for 24 hr, which was then removed to measure the protein half-life of dCas9-KRAB and Cas9. Total protein was extracted from samples and analyzed by western blot with antibodies against Cas9 and GAPDH as a loading control. Both the CRISPRi and CRISPRn clones express dCas9-KRAB and Cas9 at similar levels after doxycycline treatment, and the half-life of both proteins was ~12 hr in iPSCs. Scale bars, 100 μm.

indicated by a clear loss of pluripotent cell morphology (Hayashi et al., 2015). In general, Cas9 can disrupt gene function at any given exon (Doench et al., 2014), while dCas9-KRAB knocks down gene expression only when gRNAs are targeted to the transcription start site (TSS) (Gilbert et al., 2014). Hence, for this comparative study, we used the same gRNA sequence for both CRISPRi and CRISPRn. Here, we introduced a gRNA targeting 358 bp downstream of the *NANOG* TSS (142 bp into exon 1 of *NANOG*) into the CRISPRi and CRISPRn clones and selected subclones (as described in Experimental Procedures). We then treated multiple independent subclones of CRISPRi and CRISPRn iPSCs containing the *NANOG* gRNA-expression vector (as indicated by mKate2 expression) with doxycycline (Figure 2).

With CRISPRi, we found that NANOG expression was completely lost (>99%) in multiple independent iPSC subclones after doxycycline treatment (Figures 2A, 2C, 2E, S3A, and S3C). However, with CRISPRn, only 60%–70% of the cells lost NANOG expression in multiple independent subclones post-doxycycline induction (Figures 2B, 2D, 2G, S3B, and S3D). Next, we extracted genomic DNA from *NANOG* gRNA-containing CRISPRi and CRISPRn iPSCs and performed sequence analysis. As expected, we found that CRISPRi iPSCs did not harbor any mutations in the *NANOG* locus pre- or post-doxycycline treatment (Figure 2F). However, with CRISPRn, after 12–17 days of continuous doxycycline treatment, among the mutated alleles, 30%–50% of the sequences contained in-frame INDELs at the cut site (a total of 77 sequenced clones) (Figure 2H).

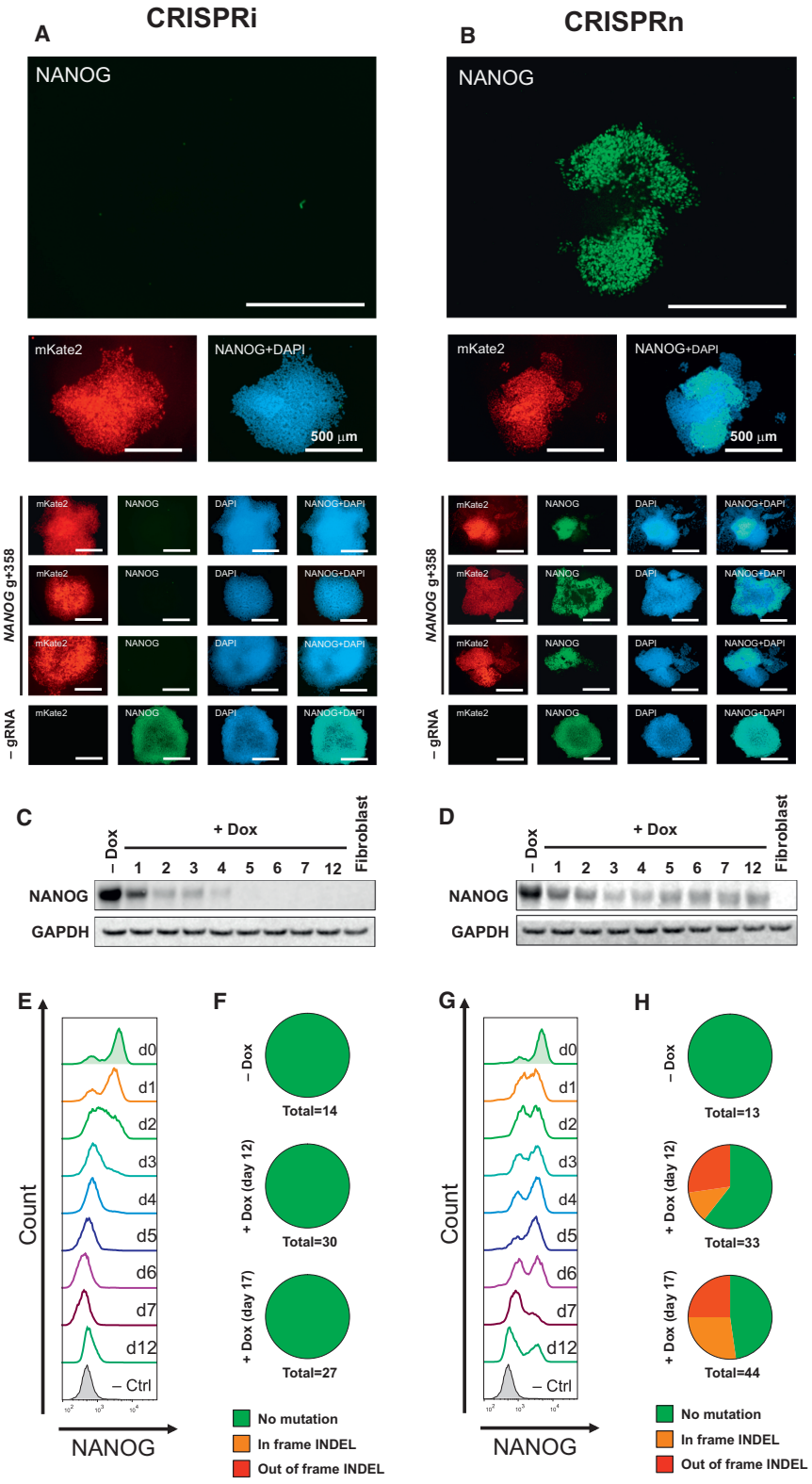


Figure 2. Comparison of the Efficiency of CRISPRi Knockdown and CRISPRn Knockout

(A and B) Immunostaining of representative (A) CRISPRi and (B) CRISPRn stable clones, each containing the same gRNA targeting the first exon of *NANOG* (*NANOG* g+358). After 7 days of doxycycline treatment, *NANOG* expression (green) was completely lost in all CRISPRi clones but showed a variegated pattern of knockout in multiple independent CRISPRn clones. The mKate2 signal indicates the presence of the gRNA-expression vector in all cells within the clone. Nuclei are counterstained with DAPI.

(C, D, E, and G) Western blot and flow cytometry analyses of (C and E) CRISPRi and (D and G) CRISPRn stable clones that contain the same gRNA against the first exon of *NANOG*. With CRISPRi, *NANOG* expression was uniformly decreased during doxycycline treatment and did not increase thereafter; however, with CRISPRn, the percentage of NANOG-positive cells fluctuated during doxycycline treatment. Even after 12 days of continuous doxycycline treatment, ~30% of the population stained positive for NANOG.

(F and H) Genomic DNA was extracted from (F) CRISPRi and (H) CRISPRn stable lines containing a gRNA against *NANOG* before and after continuous doxycycline treatment for up to 17 days and subjected to sequencing. Red, out-of-frame INDELs; orange, in-frame INDELs; green, non-mutated alleles. Even after 12–17 days of continuous doxycycline treatment, 50%–70% of sequenced alleles from CRISPRn contained no mutation, and 30%–50% of mutated alleles were in-frame INDELs. No mutations were observed in either CRISPRi or CRISPRn without doxycycline, and the CRISPRi clones did not contain any mutations after doxycycline treatment. The total number of sequenced colonies is listed below each pie graph.

Scale bars, 500 μm.

To further compare CRISPRi with CRISPRn, we targeted another pluripotency transcription factor, *OCT4*, with two independent gRNAs. Similar to our findings with *NANOG*, *OCT4* was completely knocked down in independent CRISPRi clones expressing the gRNA vector after doxycycline treatment (Figure S3E). In contrast, the attempted knockout of OCT4 with CRISPRn again yielded incomplete effects (Figure S3F). These findings were also replicated in a completely different iPSC line (WTB genetic background; CRISPRi Gen1B and CRISPRn Gen1B) (Figures S1D and S1F). We analyzed the genomic DNA of CRISPRn cells after 14 days of continuous doxycycline treatment and found 30%–40% of the mutated alleles had in-frame INDELs (a total of 91 sequenced clones) (Figure S3G). These results suggested that, in the context of targeting pluripotency factors, CRISPRi more rapidly generates loss-of-function phenotypes in bulk populations than CRISPRn. CRISPRi caused a complete

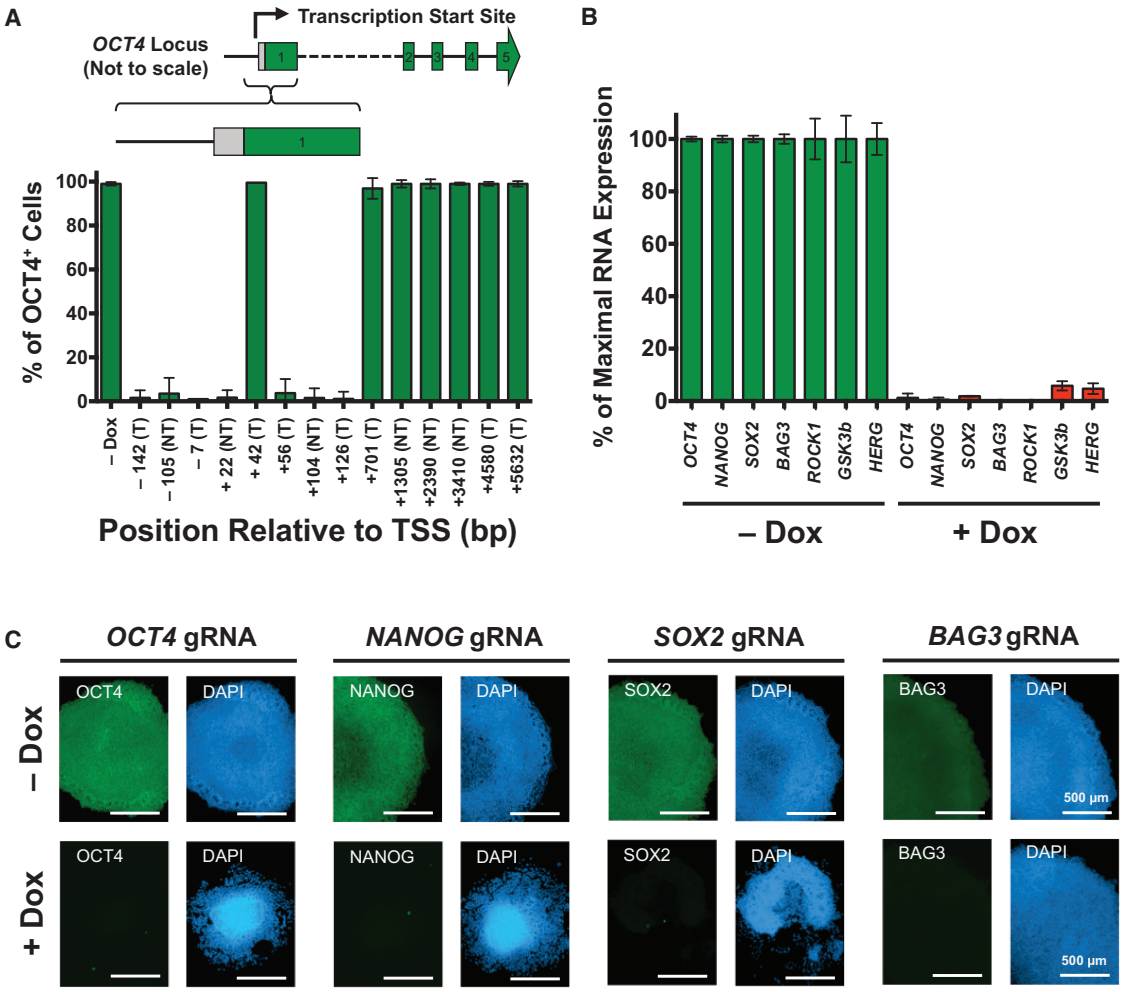


Figure 3. CRISPRi Knockdown Is Efficient in iPSCs

(A) Efficiency of gRNA knockdown based on proximity to the transcription start site (TSS). The binding location of each gRNA is indicated relative to the TSS of the *OCT4* locus and whether it targets the template (T) or non-template (NT) strand. Only gRNAs targeting near the TSS (approximately ±150 bp) effectively knocked down *OCT4*.

(B) TaqMan qPCR analysis of stable iPSCs containing gRNA against the gene of interest showed greater than 90% knockdown efficiency after 7 days of doxycycline induction in different endogenous genetic loci.

(C) Immunostaining of stable clones containing a single gRNA against the gene of interest (*OCT4, SOX2*, *NANOG,* and *BAG3*). After 7 days of doxycycline treatment, there was a complete knockdown of the protein of interest (green). As expected, DAPI staining revealed that knocking down *OCT4*, *NANOG* and *SOX2* resulted in loss of pluripotency and clear morphological changes. Also, knocking down *BAG3* did not cause a loss of pluripotent morphology, as indicated by the distinct and round colony edges.

Error bar represents SD.

loss of transcript expression and rapid cell differentiation when targeting *NANOG* and *OCT4* within 5–7 days of knockdown initiation. With CRISPRn, even after ~2 weeks of doxycycline treatment, a significant fraction (30%–40%) of the cells remained NANOG and OCT4 positive and maintained their pluripotency. Therefore, we focused on using CRISPRi as a loss-of-function tool in subsequent experiments.

CRISPRi Is Most Effective near the TSS

To further test the efficacy of gRNAs in CRISPRi, we designed multiple gRNAs that target near the TSS of OCT4. With flow cytometry assays for OCT4 staining (Figure 3A), we found that most gRNAs targeting near the TSS (approximately −150 bp to +150 bp around the TSS in this study) were highly effective at gene knockdown, but gRNAs targeting significantly (>700 bp) downstream of the TSS were not. This result agrees with previous data (Gilbert et al., 2014) and suggests that CRISPRi primarily blocks transcription at initiation, which reduces the likelihood of off-target effects from transcript interference elsewhere in the genome. Following these design criteria, for subsequent gene targets, we designed gRNAs to target near the TSS.

CRISPRi Efficiently Knocks Down a Broad Range of Genetic Loci

To test the efficiency of CRISPRi across a broad range of genetic loci in both iPSCs and differentiating/differentiated cell types, we

designed gRNAs against a total of nine genomic loci. The loci included core pluripotency transcription factors (*OCT4*, *NANOG*, and *SOX2*), kinases (*ROCK1* and *GSK3-β*), a cardiac mesoderm-transcription factor (*MESP1*), and cardiac disease-associated genes (*BAG3*, *MYBPC3*, and *HERG*). Except for *MESP1* (expressed only transiently in cardiac mesoderm cells) and *MYBPC3* (expressed only in cardiomyocytes), all other genes are expressed in iPSCs at different levels. We generated populations of CRISPRi iPSCs containing stably integrated gRNA-expression constructs. We then cultured these stable polyclones or clonal populations either with or without doxycycline for at least 7 days.

Three to five gRNAs were designed to target near the TSS of each gene and initially were tested individually in polyclonal populations. Approximately half of the tested gRNAs were active in polyclonal populations with a silencing activity of over 70% (Figure S4A). We did not observe a difference in the knockdown efficiency between gRNAs targeting either the template or non-template strands (Figures 3A, S4A, and S4B). The most active gRNA-containing polyclonal line was further passaged and subcloned for more detailed knockdown analysis. Using the most active gRNA, we achieved 90%–99% knockdown of the gene of interest in a selected population of iPSCs after doxycycline treatment (Figure 3B). As expected, when we subcloned polyclonal populations via single-cell cloning, we observed a higher percentage of transcriptional knockdown. With immunofluorescence analysis we found that iPSC clones expressing gRNAs against *OCT4*, *NANOG*, *SOX2*, and *BAG3* showed complete loss of target protein expression 7 days after doxycycline induction. In cells expressing gRNAs against the core pluripotency transcription factors *OCT4*, *NANOG*, and *SOX2*, we observed clear morphological changes and a loss of pluripotency after doxycycline induction; however, loss of a non-pluripotency gene (*BAG3*) did not affect pluripotent morphology (Figure 3C).

Using the Gen1 CRISPRi knockin vector, we targeted non-iPSCs with a different genetic background to determine how broadly this technology can be applied to other cell types. A T-lymphocyte (CEM) CRISPRi line was generated, as described in Experimental Procedures. Similar to the iPSC experiments, gRNAs were introduced to the stable CEM CRISPRi cell line, and cells cultured in either the presence or absence of doxycycline for 10 days. Three gRNAs were tested to knock down *CD4* in CEM-CRISPRi cells, and all showed greater than 70% knockdown efficiency in polyclonal populations (Figure S4B). The most active gRNA-containing polyclone was subcloned, and three independent clonal lines were isolated and assayed for knockdown, where greater than 95% knockdown efficiency was observed (Figure S4C). These results clearly demonstrate the doxycycline-inducible CRISPRi vector system is highly versatile and transportable to other cell lines and shows high efficiency of knockdown across a range of cell types and genetic loci.

CRISPRi Knockdown Is Reversible and Tunable and Can Be Allele Specific

GCaMP is a calcium-sensitive modified GFP and, thus, can be used as a fluorescent reporter under steady-state levels of cytoplasmic Ca^{2+} (Apáti et al., 2013). Using GCaMP (driven off the strong constitutive promoter, CAG), we monitored the green-fluorescence signal in iPSCs to determine if we could knock down GCaMP and then reverse its expression by removing doxycycline from the culture. We found that adding doxycycline for 7 days knocked down GCaMP expression by 98%, which was completely restored after removing doxycycline for 14 days (Figure 4A). Similarly, we targeted the *BAG3* endogenous locus and achieved efficient transcript knockdown post-doxycycline treatment. *BAG3* expression was fully restored after doxycycline withdrawal (Figure 4B). These findings indicate that CRISPRi knockdown is fully reversible in iPSCs.

To determine if we could achieve variable levels of knockdown with different gRNA sequences, we tested two additional gRNAs targeting GCaMP (g+24 and g+91) (Figure 4C). These gRNAs knocked down GCaMP expression by only ~30% and ~50%, as measured by flow cytometry (Figures 4D and 4E). Therefore, by changing the location of the gRNA-binding site, we can tune the level of knockdown when trying to mimic haploinsufficiency or reduced protein levels (rather than complete loss of function). In addition, we tested whether the knockdown level is tunable by titrating the doxycycline concentration. Careful titration of the doxycycline concentration enabled homogenous modulation of GCaMP expression (Figure S5).

We next sought to further test the tunability of knockdown with CRISPRi. We determined if we could use single-nucleotide polymorphisms (SNPs) to specifically target one allele for knockdown to achieve a heterozygous-like state. In our CRISPRi iPSCs, there is a SNP near the TSS of *OCT4*. Thus, we designed a gRNA in which the heterozygous SNP is located in the PAM sequence (AGG versus AGA). Because an "NGG" sequence is required for dCas9 to target DNA, we could selectively target only one *OCT4* allele (Figure 4F). After doxycycline induction, we found that the iPSC population carrying the SNP-specific *OCT4* gRNA (*OCT4* g−4) remained OCT4 positive (~97%) by flow cytometry analysis. However, the median intensity of OCT4 staining was reduced by ~40% after 7 days of doxycycline treatment, implying that OCT4 expression was homogeneously reduced in all cells and not just a fraction of them (Figures 4G and 4H). We confirmed this finding with immunocytochemistry and TaqMan qPCR (data not shown).

CRISPRi Knockdown Is Highly Specific

To assess the specificity of CRISPRi targeting, we designed a gRNA that targets the GCaMP transgene, since its silencing should have few downstream transcriptional and cellular consequences. Indeed, expression of the GCaMP transcript was over 30-fold lower in the doxycycline-treated condition, while few other endogenous transcripts changed expression level with the exception of *VIM* as previously discussed (Figure 5A).

CRISPRi to Promote iPSC Differentiation

To show that our system can release iPSCs from their pluripotent state to promote differentiation, we tested the efficiency of CRISPRi in knocking down core pluripotency transcription factors (*OCT4*, *SOX2*, and *NANOG*) without adding small molecules or cytokines to the mTeSR media. We targeted gRNA against these genes and performed a time-course analysis of a selected number of transcripts by TaqMan qPCR (Figure 5B). We found that knocking down these target transcripts caused cell differentiation, as indicated by morphological changes and transient expression of the lineage-specific transcript *T* (mesoderm marker), and expression of *PAX6* (neuronal progenitor marker). After 3 days

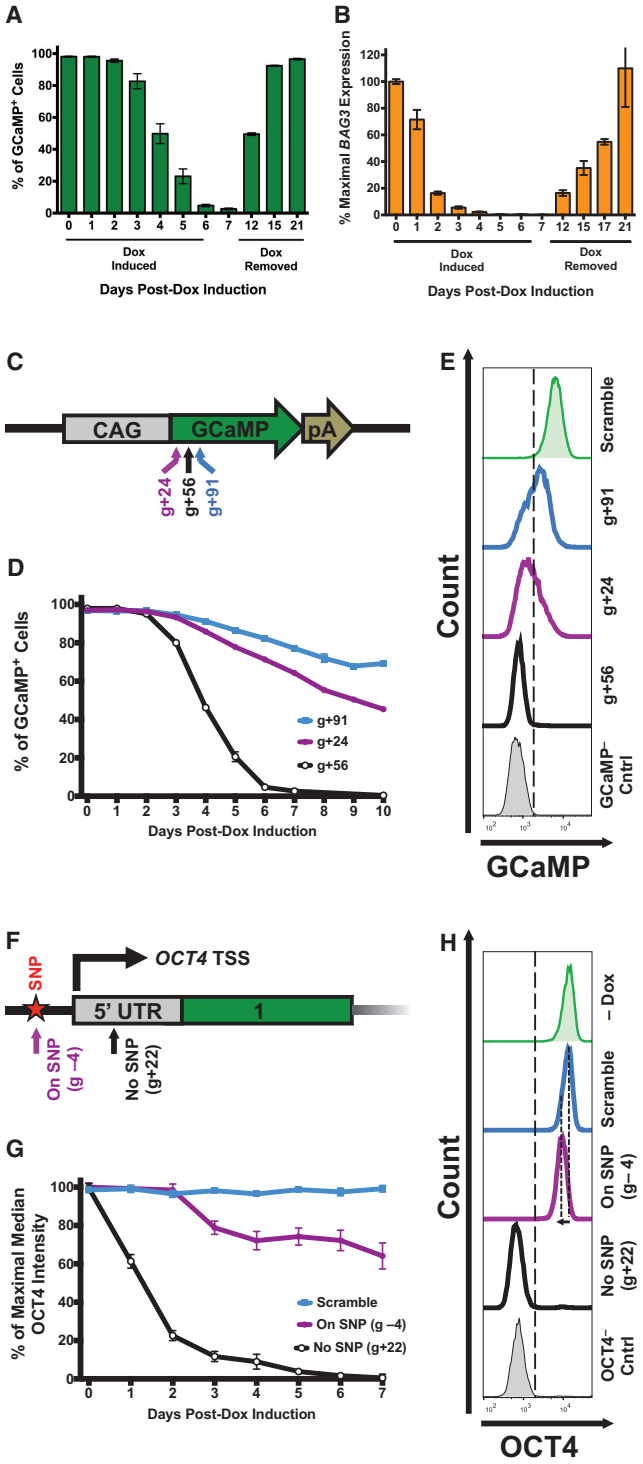


Figure 4. CRISPRi Knockdown Is Reversible and Tunable

A CRISPRi clone containing gRNA against the GCaMP transgene (GCaMP g+56) and endogenous *BAG3* locus were used to test the knockdown efficiency and reversibility of the CRISPRi system in iPSCs.

(A) Flow cytometry analysis of GCaMP expression showed that after 7 days of doxycycline induction, GCaMP was knocked down by ~99% and was completely restored after doxycycline withdrawal for 14 days.

(B) Using TaqMan qPCR, *BAG3* transcript levels were knocked down to nearly undetectable levels, and expression was restored after doxycycline withdrawal.

(C) Schematic diagram of the GCaMP-expression vector in which the GCaMP open reading frame (ORF) is driven off the CAG promoter. The locations of three gRNAs (g+24, g+56, and g+91) are schematically highlighted on the GCaMP ORF. The coordinates of GCaMP gRNA are based on the translation start site. pA, poly A signal.

(D) Three stable CRISPRi colonies, each containing a different gRNA against GCaMP, were selected using blasticidin and cultured with doxycycline for 10 days. The percentage of GCaMP-positive cells for each gRNA-containing clone was plotted as a function of time based on flow cytometry analysis. Variable levels of GCaMP knockdown (~30%, ~50%, and ~99%) were achieved with different gRNA sequences. n = 1–3 technical replicates for each time point.

(E) Flow cytometry plots of GCaMP fluorescence of stable CRISPRi clones on day 10 of doxycycline treatment. Using different gRNAs that target near the same region, variable levels of knockdown can be achieved. A scramble gRNA-containing CRISPRi and a GCaMP-negative iPSC population are displayed as controls.

(F) Partial schematic diagram of the *OCT4* locus marked with the location of the TSS and two gRNA-binding locations. Asterisk, an SNP; green box, exon 1; gray box, 5′ UTR.

(G) Three stable CRISPRi colonies, two with different gRNAs against *OCT4* and one with a scrambled control, were selected with blasticidin. Stable clones that contain either a scramble gRNA, a gRNA that targets a PAM sequence containing a SNP (*OCT4* g–4), or a gRNA that does not target a SNP (*OCT4* g+22) were treated with doxycycline. The percentage of the maximal median intensity of OCT4 staining for each gRNA-containing clone is plotted as a function of time by flow cytometry analysis. Complete loss of *OCT4* expression (>98% knockdown) was observed after 7 days of doxycycline induction only when both alleles were targeted using *OCT4* g+22. While using *OCT4* g–4, which targets only one *OCT4* allele (due to SNP in the PAM sequence), a gradual loss of OCT4 staining intensity is observed over time (down by ~40% by day 7). Error bars represent SD; n = 1–3 technical replicates for each time point.

(H) Flow cytometry plots of OCT4 staining on day 7 of doxycycline treatment. Dashed lines highlight the loss of OCT4-staining intensity (~40%) when using *OCT4* g–4 compared to the scramble control. By targeting only one allele of *OCT4*, the OCT4-staining intensity homogeneously shifts (while remaining OCT4-positive), indicating that each cell experiences approximately the same level of knockdown. Note that the x axis is a log-scale of OCT4 intensity. Differentiated iPSC-derived fibroblasts (OCT4$^-$ Cntrl) and a non-doxycycline-treated (–Dox) sample are displayed as controls.

Error bars represent SD.

of doxycycline treatment, over 80% of the target transcript was depleted, indicating that CRISPRi can precisely and temporally control efficient knockdown of the transcript of interest.

CRISPRi Knockdown in Cardiac Mesoderm and iPS-CMs

To determine if loss-of-function approaches using CRISPRi can be applied in differentiated cell types, we targeted the cardiac mesoderm-specific transcription factor (*MESP1*) and two known cardiac-related disease-causing genes (*MYBPC3* and *HERG*). We established stable polyclonal lines of iPSCs containing gRNA against these three genes and differentiated them into cardiac mesoderm or iPS-CMs as described in Experimental Procedures (Figures S6A and S6B). Using a gRNA against these genes, *MESP1* was knocked down by ~90% in cardiac progenitor cells, and *MYBPC3* and *HERG* by ~90% and 60%, respectively, in lactate-purified iPS-CMs (Figure 6A). With western blots and immunocytochemistry, we observed ~90% MYBPC3 protein knockdown on day-35 lactate-purified iPS-CMs (Figures 6B and 6C).

Using flow cytometry, we analyzed the doxycycline response of CRISPRi cells based on mCherry expression (as a surrogate

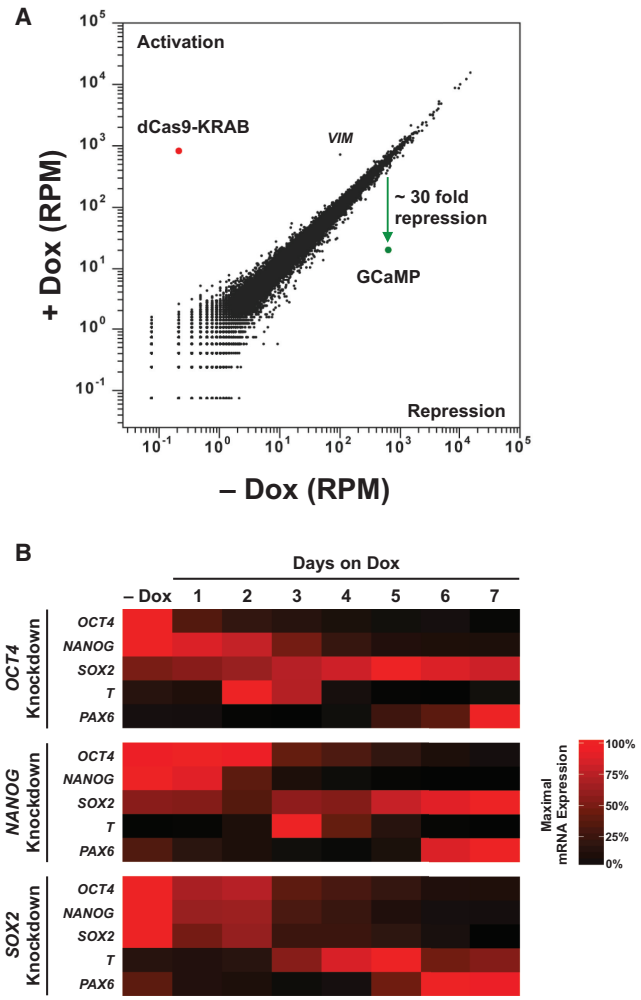


Figure 5. RNA-Seq and TaqMan qPCR Analysis

(A) RNA-sequencing RPMs (reads per million) are plotted for CRISPRi cells stably expressing a gRNA targeting the GCaMP transgene (GCaMP g+56) cultured in the absence or presence of doxycycline. CRISPRi knockdown is specific to the GCaMP transcript, and few off-target transcriptional changes were observed. Data represent two independent biological replicates.

(B) Heatmap of TaqMan qPCR of stable clones containing a single gRNA against the gene of interest (*OCT4*, *NANOG*, and *SOX2*) as a function of days after doxycycline treatment. Analysis shows that by day 3, over 80% of the target transcript is depleted. Three housekeeping genes (*18S*, *GAPDH*, and *UBC*) were used to measure relative transcript levels. Each data point is an average of two to four technical replicates. TaqMan probes are listed in Supplemental Experimental Procedures.

for dCas9-KRAB expression; Figure S5A). There was no silencing of the TetO promoter in low-passage and high-passage iPSCs, suggesting that long-term culturing (>3 months) does not cause silencing. However, cardiac progenitors (day 5) and iPS-CMs (day 15) lose ~20% and 50%–80% of the doxycycline response, respectively. Prolonging the duration of doxycycline treatment (from 2 to 7 days) and splitting the cells improved doxycycline response (as measured by mCherry expression) in iPS-CMs (Figure S6C). For this reason, we initiated all of our knockdowns on day 5 post-differentiation to obtain the maximum amount of target gene silencing. It is worth noting that with CRISPRi, only minute amounts of the dCas9-KRAB protein are necessary to induce a knockdown. Hence, knockdown might occur even in cells that do not show detectable mCherry expression (Figure S5).

The knockdown of the *HERG* potassium channel in iPSCs was highly efficient (>95%), while in iPS-CMs it was only 60% effective. We hypothesize that the reduction in the efficiency of *HERG* knockdown is partially due to activation of other *HERG* isoforms in iPS-CMs. We further investigated whether knocking down the *HERG* potassium channel in iPS-CMs would recapitulate a physiologically relevant cellular phenotype. We found that knocking down *HERG* in iPS-CMs lead to a prolonged beat duration and the appearance of a shoulder during the downstroke, as measured using the GCaMP signal (which can be used as a surrogate for the action potential) (Huebsch et al., 2015) (Figures 6D and 6E). We confirmed the prolongation of action potential duration by patch-clamp electrophysiology in the *HERG* knockdown samples (Figures 6F). We expected this result, because the *HERG* potassium channel pumps potassium ions out of cells to lower the inner membrane potential during diastole. This cellular phenotype recapitulates aspects of the phenotype observed in LQT patients and their iPS-CMs (Schwartz et al., 2012; Spencer et al., 2014).

DISCUSSION

In this study, we combined the power of human iPSC technology, which generates functional human cells, with inducible CRISPR-based genome editing and modulation technologies. Using the TetO inducible system, we deploy the newly developed CRISPRi system in the AAVS1 safe–harbor locus of human iPSCs to enable precise control of transcript silencing upon addition of doxycycline. With this approach, we rapidly and efficiently generated loss-of-function phenotypes in iPSCs and their cell-type derivatives to study mechanisms in development and disease. We introduced a single doxycycline-inducible vector system into the AAVS1 safe-harbor locus to gain tight transcriptional control of dCas9-KRAB (for CRISPRi) and Cas9 (for CRISPRn) for gene knockdown and knockout studies, respectively. This inducible vector system helped us precisely control the timing of knocking down the expression of target genes in a clonal iPSC line carrying the gRNA of interest. We were also able to efficiently target the CRISPRi vector into non-iPSC human cells (T-lymphocytes) and show efficient levels of transgene knockdown, which demonstrates the versatility of using the CRISPRi system in a wide range of cell types. This system can be readily targeted to other human cellular models in vitro and also to mouse models (Soriano, 1999) by exchanging the AAVS1-homology arms with the ROSA26-specific knockin arms.

We found that in iPSC populations, CRISPRi produced a homogeneous and rapid loss-of-function phenotype compared to CRISPRn. CRISPRi avoids potential complications associated with incomplete loss-of-function and gain-of-function phenotypes in cell populations produced by Cas9-induced hypomorphic alleles. Therefore, CRISPRi represents a powerful technology for repressing gene expression in bulk populations and especially when performing genome-scale phenotypic screens. Every CRISPRi iPSC that contained a target-specific gRNA

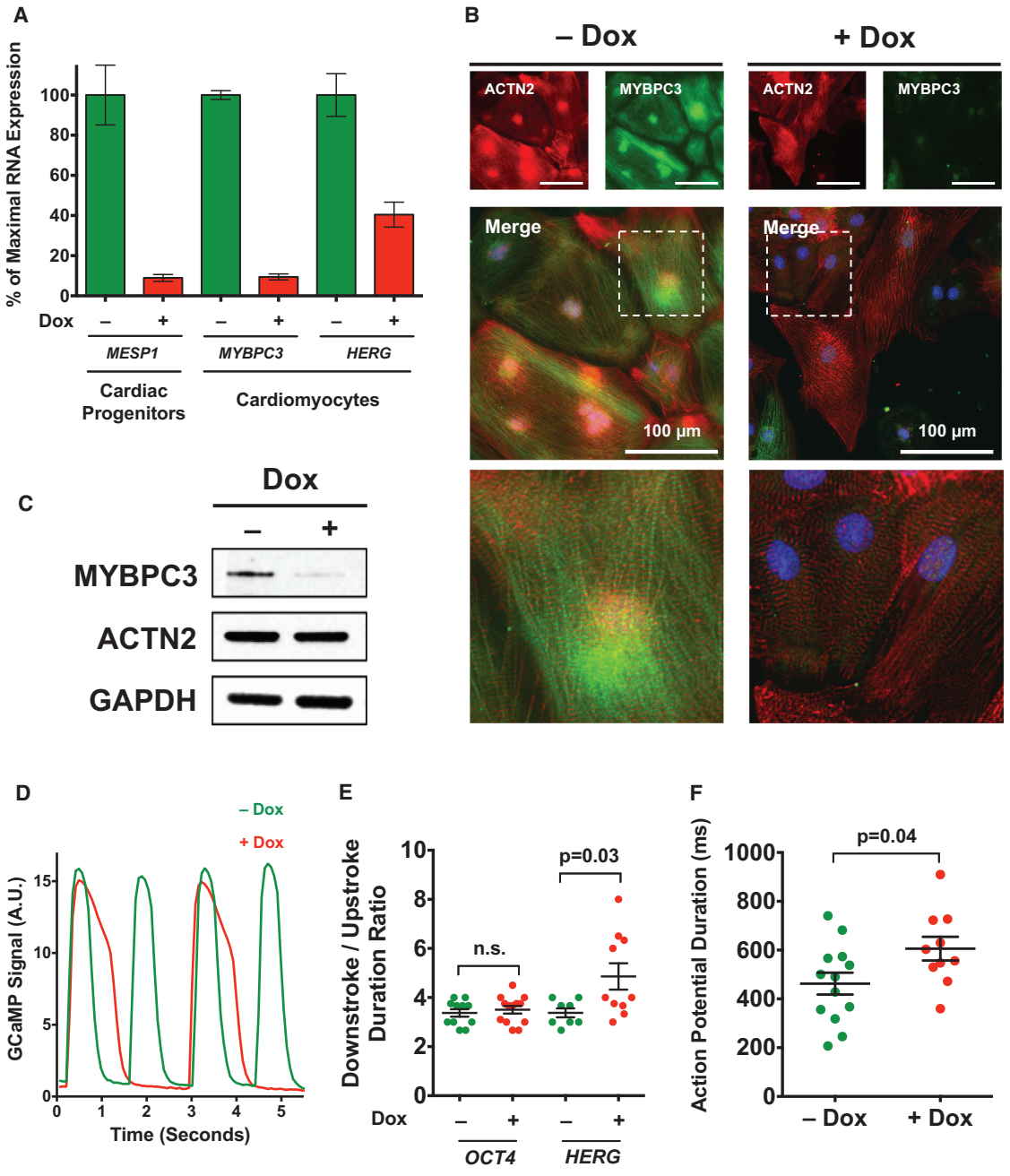


Figure 6. CRISPRi Knockdown in Differentiated Cell Types and Cardiac Disease Modeling

(A) Using CRISPRi, *MESP1* was knocked down by ~90% in polyclonal cardiac progenitors, and *MYBPC3* and *HERG* were knocked down by ~90% and 60% in polyclonal iPS-CMs, respectively.

(B) Immunostaining of day-35 lactate-purified iPS-CMs stained with antibodies against MYBPC3 (green) and ACTN2 (red). Using CRISPRi knockdown, loss of MYBPC3 was observed in over 85% of analyzed cells in a polyclonal population. Nuclei were counterstained with DAPI. Scale bar, 100 μm.

(C) Western blot of day-35 lactate-purified iPS-CMs with antibodies against MYBPC3, ACTN2, and GAPDH. Using CRISPRi, MYBPC3 protein was knocked down by ~90%.

(D) GCaMP fluorescence in iPS-CMs containing gRNA against *HERG* and cultured in doxycycline (red). Recordings show a prolonged beat duration compared to untreated controls (green).

(E) Quantified ratio of the downstroke-to-upstroke duration of doxycycline-treated iPS-CMs shows a significant difference in untreated iPS-CMs containing a gRNA against *HERG*, but not in iPS-CMs containing gRNA against *OCT4* (negative control).

(F) Patch-clamp recordings from single iPS-CMs show prolonged action potential durations in doxycycline-treated samples containing *HERG* gRNA.

Error bars represent SD.

displayed a rapid, uniform, and efficient transcriptional knockdown. This result was also validated across multiple endogenous loci in iPSCs, cardiac progenitors, and iPS-CMs. By contrast, using CRISPRn, we found that while all cells harbored the gRNA-expression vector and had continuous expression of Cas9, they did not all display complete loss-of-function phenotypes. Indeed, up to one-third of the cells maintained expression of the target gene. When we sequenced the target alleles, we found that of the mutated alleles, over one-third had in-frame INDELs, potentially resulting in a hypomorphic protein encoded by a gene that is now resistant to further Cas9 cutting using the target gRNA. Statistically, we expect that one-third of the INDELs generated by double-strand breaks induced by Cas9 through the non-homologous end-joining pathway would produce in-frame mutations. This effect could cause partial loss-of-function or gain-of-function phenotypes. Additionally, the location and size of the in-frame INDEL might not change the function of the mutated protein compared with the wild-type protein (Boettcher and McManus, 2015; Shi et al., 2015; Sung et al., 2013).

CRISPRi gRNAs were only effective at promoter regions close to the TSS, which may reduce the likelihood of off-target effects by transcriptional interference elsewhere in the genome. Indeed, RNA-seq analysis showed that the knockdown of GCaMP was highly specific. Furthermore, expression of dCas9-KRAB did not cause significant off-target transcriptional changes as compared to Cas9 expression alone. Although CRISPRi is highly effective, there are cases when other genetic tools such as CRISPRn, TALENs, and RNAi may have advantages. For instance, we and others (Gilbert et al., 2014) have shown that CRISPRi gRNAs are only effective near the TSS, which restricts the efficiency of transcript for genes that have poorly defined or multiple TSSs. CRISPRn and TALENs can be effective at any exon as long as the genomic region is accessible (Doench et al., 2014; Kim et al., 2013b). Additionally, RNAi can target any constitutive portion of the mRNA and has already been approved for human therapy (Davidson and McCray, 2011; Haussecker, 2012); however, RNAi has been shown to have many off-target effects (Jackson et al., 2003; Kim et al., 2013b; Krueger et al., 2007).

We also demonstrated the feasibility of allele-specific interference and the tunable nature of CRISPRi-based knockdown, which can be used to study the dose-dependent effects of a gene involved in development and disease. The dosage of transcription factors plays a significant role during development and organogenesis (McFadden et al., 2005; Takeuchi et al., 2011). In addition, many human diseases result from haploinsufficiency in which a mutation in a single copy of a gene produces the disease phenotype (Armanios et al., 2005; Marston et al., 2012; Minami et al., 2014; Theodoris et al., 2015). Therefore, to study the dose-dependent effects of transcription factors in development and disease, CRISPRi can be used to homogeneously tune the level of repression in cells by either choosing the relevant gRNA sequences or empirically titrating the levels of doxycycline to achieve the desired knockdown level. Alternatively, introducing a single point mutation at different positions in the gRNA sequence (which leads to mismatches between the RNA-DNA homology sequence) can be used to tune CRISPRi knockdown activity (Gilbert et al., 2014). Finally, CRISPRi knockdown was reversible in iPSCs upon doxycycline withdrawal, which would support studies involving transient knockdown of transcripts within a specific window during cell differentiation.

Our studies with CRISPRi in iPSCs showed that knocking down transcripts involved in maintaining pluripotency is highly efficient and rapidly causes a complete loss of pluripotent morphology, followed by cell differentiation in all cells expressing the appropriate gRNA. We also used this approach to knock down the *HERG* potassium channel to mimic an LQT2-type phenotype in iPS-CMs. We found that the inducible TetO promoter is partially silenced during the cardiac differentiation process, which has been reported to be due to methylation at CpG dinucleotides (Oyer et al., 2009). This silencing is independent of integration at the AAVS1 locus, as CAG-driven transgenes integrated at the AAVS1 locus remain active after differentiation. To avoid the effects of promoter silencing, we initiated transcript knockdown in the iPSC state or progenitor cells (day 5 of differentiation), where the vast majority of the cells respond to doxycycline. This strategy has proved highly effective at transgene knockdown in cardiac progenitors and iPS-CMs. To circumvent issues with silencing in future studies, we generated a non-inducible CRISPRi iPSC line (Gen3; in which dCas9-KRAB is driven off the CAG promoter), and the knockdown can be initiated upon introduction of gRNA. With this cell line, we expect to achieve highly efficient knockdown in differentiated cell types, such as iPS-CMs.

Several groups have used the CRISPR/Cas9 system for loss-of-function genetic screens in human cells (Shalem et al., 2014; Wang et al., 2014). Furthermore, some groups have used genome-scale screens with CRISPRi and CRISPR activation (CRISPRa) to identify known and novel genes that control cell growth and sensitivity to cholera-diphtheria toxin (Gilbert et al., 2014). In this study, we present our CRISPRi iPSC lines as suitable model systems for performing screens to identify novel transcripts of pluripotency, drug resistance, and cell survival at the pluripotent stem cell stage. With genome-scale screens, we can identify factors that improve cell-specific differentiation into functional cell types that have been traditionally hard to obtain, and we can more rapidly generate mature functional cell types that better mimic in vivo cell counterparts. In addition, with CRISPRi, we can repress putative disease-associated genes in a medium- to high-throughput manner to unravel the molecular mechanisms underlying human disease in vitro. Finally, we can build on the current power of CRISPRi for developmental screens by using an orthogonal dCas9-effector system for gene activation via CRISPRa, which can synergistically modulate gene knockdown and activation and direct cell fate toward a particular lineage.

EXPERIMENTAL PROCEDURES

iPSC Culture

WTB and WTC iPSCs and derivative lines were maintained under feeder-free conditions on growth factor-reduced Matrigel (BD Biosciences) and fed daily with mTeSR medium (STEMCELL Technologies) (Ludwig et al., 2006). Accutase (STEMCELL Technologies) was used to enzymatically dissociate iPSCs into single cells. To promote cell survival during enzymatic passaging, cells were passaged with the p160-Rho-associated coiled-coil kinase (ROCK) inhibitor Y-27632 (10 μM; Selleckchem) (Watanabe et al., 2007). iPSCs were frozen in 90% fetal bovine serum (HyClone) and 10%

DMSO (Sigma). The committee on Human Research at the University of California, San Francisco approved the iPSC research protocol (#10-02521).

Generation of Stable CRISPRi and CRISPRn iPSC Lines

iPSCs were singularized with accutase, resuspended in PBS, and counted with a Countess automated cell counter (Life Technologies). For plasmid transfections, the human stem cell nucleofector kit 1 solution was used on the Amaxa nucleofector 2b device (program A-23; Lonza). To generate the CRISPRi and CRISPRn iPSC lines, two million WTC or WTB iPSCs were nucleofected with the appropriate knockin vector (5 μg) and each AAVS1 TALEN pair (2 μg). Cells were then seeded in six-well plates in serial dilutions in mTeSR supplemented with Y-27632 (10 μM). Selection was applied 3 days post-nucleofection with the appropriate antibiotic in mTeSR plus Y-27632 (10 μM). To knock in the CRISPRi construct (carrying the Neomycin resistance gene cassette), Geneticin (Life Technologies) was applied at 100 μg/ml. To knock in the CRISPRn and GCaMP constructs (carrying the Puromycin resistance gene cassette), 0.5 μg/ml Puromycin (Life Technologies) was added. Selection was maintained for ~10 days until stable colonies appeared. Colonies with a diameter of greater than ~500 μm were manually picked using a P200 pipette tip under an EVOS FL picking microscope (Life Technologies) and transferred to individual wells of a 24-well plate containing mTeSR medium supplemented with Y-27632 (10 μM). Clones were then expanded into larger vessel formats.

Generation of CEM CRISPRi Cell Line

CEM CRISPRi cells were generated by electroporation of 0.5 μg of each AAVS1 TALEN pair and 1 μg of the Gen1 CRISPRi vector with an Amaxa nucleofector 2b device and Amaxa cell line nucleofector kit C (Lonza). Cells were selected in 1 μg/μl G418, and clonal lines were generated by dilution in 96-well plates. Clonal populations were selected based on doxycycline induction of mCherry expression. Oligos encoding the CD4 protospacer were annealed and cloned into the pSLQ1371 vector using restriction sites BstXI and BlpI, and lentivirus was produced in HEK293T cells (Gilbert et al., 2014). To compare performance of CD4 gRNAs, each was transduced into CEM-CRISPRi cells. Transduced populations were incubated for 96 hr with doxycycline (2 μM). Knockdown efficiency was calculated by gating all mCherry-expressing cells, and comparing cell-surface CD4 expression in the presence or absence of gRNA-expressing cells (BFP^+). Three independent stable CEM CRISPRi clones were selected with 0.6 μg/ml Puromycin and incubated in the presence or absence of doxycycline (2 μM) for 14 days to assess maximal CD4 knockdown. Cells were stained using anti-CD4 APC-conjugated antibody and cell surface CD4 staining was quantified using a BD LSRII flow cytometer. CD4 knockdown was quantified as percent reduction relative to no doxycycline treatment condition.

gRNA Design and Cloning into the gRNA-Expression Vector

For CRISPRi, three to five gRNAs were designed to target near the TSS of the gene of interest (250 bp upstream and downstream, respectively). The location of the TSS was determined using NCBI (http://www.ncbi.nlm.nih.gov/). gRNA oligos were designed, phosphorylated, annealed, and cloned into the pgRNA-CKB vector using BsmBI ligation strategy. Additional details and a list of gRNA sequences are listed in supplemental experimental procedures.

gRNA Nucleofection and Selection of Stable CRISPRi and CRISPRn Clones

The gRNA-expression vector (pgRNA-CKB) was transfected into either the CRISPRi or CRISPRn cells with the human stem cell nucleofector kit 1 solution on the Amaxa nucleofector 2b device (program A-23; Lonza). Two million CRISPRi or CRISPRn iPSCs and 5 μg of the circular gRNA-expression plasmid were used per nucleofection. Nucleofected cells were then seeded in a single well of a six-well plate in mTeSR supplemented with Y-27632 (10 μM). Blasticidin selection (10 μg/ml) was applied 24 hr post-nucleofection in mTeSR supplemented with Y-27632 (10 μM) for 7–10 days, until stable colonies appeared. Stable colonies were then pooled and passaged at least three times in mTeSR plus Blasticidin and Y-27632 to enrich for cells with integration at transcriptionally active sites (Figure S3).

RNA Sequencing

For each sample, 1 μg of total RNA was prepared using TRIzol as previously described. Strand-specific mRNaseq libraries were prepared using TruSeq Stranded mRNA Library Prep Kit (Illumina). Upon completion, libraries were quantified and pooled using Qubit dsDNA HS assay and Agilent's Bioanalyzer high-sensitivity DNA assay. The indexed libraries were pooled and sequenced on Illumina HiSeq 4000 as 50-bp single-end reads. Reads were aligned to the hg19 genome assembly using the Ensembl 75 reference transcriptome customized to include the GCaMP6f constructs using TopHat2 (Kim et al., 2013a). Unaligned reads were subsequently aligned to the CRISPRi or CRISPRn knockin constructs where appropriate. Transcript alignments were then counted using SubRead v1.4.6 and analyzed with custom scripts written in Python (Liao et al., 2013). All data are displayed as reads per million (RPM) with a pseudocount of 0.075.

iPS-CM Differentiation and Lactate Purification

iPSCs were differentiated into iPS-CMs using the WNT modulation-differentiation method (Lian et al., 2012) (Figure S5A). iPS-CMs were purified via a modified version of the lactate metabolic-selection method (Tohyama et al., 2013). Additional details are outlined in Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the RNA-seq data reported in this paper is GEO: PRJNA307261.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2016.01.022.

AUTHOR CONTRIBUTIONS

M.A.M. and B.R.C. were primarily responsible for conception, design, and interpretation of the experiments. M.A.M. conducted most experiments with help from N.H., E.F., E.S., A.T., M.P.O., T.V.E., K.H., and L.M.J. Y.M. and A.H.C. generated the CRISPRn Gen1C iPSC line. C.I.S. performed electrophysiology experiments. D.E.G. generated the CEM CRISPRi cell line and provided knockdown analysis. L.A.G., J.S.W., and L.S.Q. provided technical expertise, the CRISPRi fusion cassette, and gRNA expression constructs. J.E.V. and M.A.H. conducted and analyzed the RNA-seq experiments. M.A.M., P.L.S., and B.R.C. wrote the manuscript with support from all authors.

ACKNOWLEDGMENTS

We thank members of the Conklin laboratory, Gladstone Institute of Cardiovascular Disease, Roddenberry Stem Cell Core, BioFulcrum, a Gladstone Institutes Enterprise, and Innovative Genomics Institute for technical assistance and helpful comments on the manuscript. We thank Tim Rand and Knut Woltjen for valuable discussions and helpful comments on the manuscript. We thank S. John Liu for RNA-seq analysis advice. We thank Jen Berman and Samantha Cooper at Bio-Rad for assistance with designing ddPCR probe-primer sets. Summer students Matthew Keller and Monique Morrison assisted with preliminary experiments. CEM $CD4^+$ cells were obtained from Dr. J.P. Jacobs through the AIDS Reagent Program, Division of AIDS, NIAID, NIH. M.A.M. is supported by the Canadian Institutes of Health Research postdoctoral fellowship 129844. N.H. was supported by the CIRM training program TG2-01160 and T32 HL007544. E.F. was supported by a Bridges to Stem Cell Training grant TB1-01188 from CIRM. L.M.J. is supported by the CIRM Training Grant TG2-01160 and NICHD Career Development Award 1K12HD072222. Y.M. received fellowships from the Uehara Memorial Foundation Research and Gladstone-CIRM. D.G. is supported by UCSF-Gladstone Center for AIDS Research (CFAR), an NIH-funded program (P30 AI027763). T.V.E. was supported by Carlsberg Travel Grant (2013-01-0423), The Lundbeck Foundation (R140-2013-13348), and OUH Internationalisation Foundation. J.E.V., M.A.H., L.A.G., and J.S.W. were supported by the Howard Hughes

Medical Institute and the National Institutes of Health (R01 DA036858). B.R.C. received support from the US National Heart, Lung, and Blood Institute, the National Institutes of Health (U01-HL100406, U01-GM09614, R01-HL108677, U01-HL098179, U01-HL099997, P01-HL089707, R01HL130533, and R01-HL060664), and an Agilent University Relations Grant.

Received: August 20, 2015
Revised: December 21, 2015
Accepted: January 24, 2016
Published: March 10, 2016

REFERENCES

Apáti, Á., Pászty, K., Hegedűs, L., Kolacsek, O., Orbán, T.I., Erdei, Z., Szebényi, K., Péntek, A., Enyedi, Á., and Sarkadi, B. (2013). Characterization of calcium signals in human embryonic stem cells and in their differentiated offspring by a stably integrated calcium indicator protein. Cell. Signal. *25*, 752–759.

Armanios, M., Chen, J.-L., Chang, Y.-P.C., Brodsky, R.A., Hawkins, A., Griffin, C.A., Eshleman, J.R., Cohen, A.R., Chakravarti, A., Hamosh, A., and Greider, C.W. (2005). Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. Proc. Natl. Acad. Sci. USA *102*, 15960–15964.

Boettcher, M., and McManus, M.T. (2015). Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR. Mol. Cell *58*, 575–585.

Bolouri, H., and Davidson, E.H. (2003). Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. Proc. Natl. Acad. Sci. USA *100*, 9371–9376.

Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature *499*, 295–300.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Davidson, B.L., and McCray, P.B., Jr. (2011). Current prospects for RNA interference-based therapies. Nat. Rev. Genet. *12*, 329–340.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat. Biotechnol. *32*, 1262–1267.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Gaj, T., Gersbach, C.A., and Barbas, C.F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. *31*, 397–405.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442–451.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell *159*, 647–661.

González, F., Zhu, Z., Shi, Z.-D., Lelli, K., Verma, N., Li, Q.V., and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell *15*, 215–226.

Haussecker, D. (2012). The business of RNAi therapeutics in 2012. Mol. Ther. Nucleic Acids *1*, e8.

Hayashi, Y., Caboni, L., Das, D., Yumoto, F., Clayton, T., Deller, M.C., Nguyen, P., Farr, C.L., Chiu, H.-J., Miller, M.D., et al. (2015). Structure-based discovery of NANOG variant with enhanced properties to promote self-renewal and reprogramming of pluripotent stem cells. Proc. Natl. Acad. Sci. USA *112*, 4666–4671.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat. Biotechnol. *29*, 731–734.

Huebsch, N., Loskill, P., Mandegar, M.A., Marks, N.C., Sheehan, A.S., Ma, Z., Mathur, A., Nguyen, T.N., Yoo, J.C., Judge, L.M., et al. (2015). Automated video-based analysis of contractility and calcium flux in human-induced pluripotent stem cell-derived cardiomyocytes cultured over different spatial scales. Tissue Eng. Part C Methods *21*, 467–479.

Jackson, A.L., Bartz, S.R., Schelter, J., Kobayashi, S.V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P.S. (2003). Expression profiling reveals off-target gene regulation by RNAi. Nat. Biotechnol. *21*, 635–637.

Kearns, N.A., Genga, R.M.J., Enuameh, M.S., Garber, M., Wolfe, S.A., and Maehr, R. (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. Development *141*, 219–223.

Kim, H., and Kim, J.-S. (2014). A guide to genome engineering with programmable nucleases. Nat. Rev. Genet. *15*, 321–334.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013a). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kim, Y., Kweon, J., Kim, A., Chon, J.K., Yoo, J.Y., Kim, H.J., Kim, S., Lee, C., Jeong, E., Chung, E., et al. (2013b). A library of TAL effector nucleases spanning the human genome. Nat. Biotechnol. *31*, 251–258.

Krueger, U., Bergauer, T., Kaufmann, B., Wolter, I., Pilk, S., Heider-Fabian, M., Kirch, S., Artz-Oppitz, C., Isselhorst, M., and Konrad, J. (2007). Insights into effective RNAi gained from large-scale siRNA validation screening. Oligonucleotides *17*, 237–250.

Lian, X., Hsiao, C., Wilson, G., Zhu, K., Hazeltine, L.B., Azarin, S.M., Raval, K.K., Zhang, J., Kamp, T.J., and Palecek, S.P. (2012). Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. Proc. Natl. Acad. Sci. USA *109*, E1848–E1857.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res. *41*, e108.

Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provasi, E., Colombo, D.F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat. Methods *8*, 861–869.

Ludwig, T.E., Bergendahl, V., Levenstein, M.E., Yu, J., Probasco, M.D., and Thomson, J.A. (2006). Feeder-independent culture of human embryonic stem cells. Nat. Methods *3*, 637–646.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

Marston, S., Copeland, O., Gehmlich, K., Schlossarek, S., and Carrier, L. (2012). How do MYBPC3 mutations cause hypertrophic cardiomyopathy? J. Muscle Res. Cell Motil. *33*, 75–80.

McFadden, D.G., Barbosa, A.C., Richardson, J.A., Schneider, M.D., Srivastava, D., and Olson, E.N. (2005). The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. Development *132*, 189–201.

Minami, S.S., Min, S.-W., Krabbe, G., Wang, C., Zhou, Y., Asgarov, R., Li, Y., Martens, L.H., Elia, L.P., Ward, M.E., et al. (2014). Progranulin protects against amyloid β deposition and toxicity in Alzheimer's disease mouse models. Nat. Med. *20*, 1157–1164.

Miyaoka, Y., Chan, A.H., Judge, L.M., Yoo, J., Huang, M., Nguyen, T.D., Lizarraga, P.P., So, P.-L., and Conklin, B.R. (2014). Isolation of single-base genome-edited human iPS cells without antibiotic selection. Nat. Methods *11*, 291–293.

Oyer, J.A., Chu, A., Brar, S., and Turker, M.S. (2009). Aberrant epigenetic silencing is triggered by a transient reduction in gene expression. PLoS ONE *4*, e4832.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173–1183.

Schwartz, P.J., Crotti, L., and Insolia, R. (2012). Long-QT syndrome: from genetics to management. Circ Arrhythm Electrophysiol *5*, 868–877.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84–87.

Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., and Vakoc, C.R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat. Biotechnol. *33*, 661–667.

Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat. Genet. *21*, 70–71.

Spencer, C.I., Baba, S., Nakamura, K., Hua, E.A., Sears, M.A.F., Fu, C.C., Zhang, J., Balijepalli, S., Tomoda, K., Hayashi, Y., et al. (2014). Calcium transients closely reflect prolonged action potentials in iPSC models of inherited cardiac arrhythmia. Stem Cell Reports *3*, 269–281.

Sterneckert, J.L., Reinhardt, P., and Schöler, H.R. (2014). Investigating human disease using stem cell models. Nat. Rev. Genet. *15*, 625–639.

Sung, Y.H., Baek, I.-J., Kim, D.H., Jeon, J., Lee, J., Lee, K., Jeong, D., Kim, J.-S., and Lee, H.-W. (2013). Knockout mice created by TALEN-mediated gene targeting. Nat. Biotechnol. *31*, 23–24.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Takeuchi, J.K., Lou, X., Alexander, J.M., Sugizaki, H., Delgado-Olguín, P., Holloway, A.K., Mori, A.D., Wylie, J.N., Munson, C., Zhu, Y., et al. (2011). Chromatin remodelling complex dosage modulates transcription factor function in heart development. Nat. Commun. *2*, 187.

Theodoris, C.V., Li, M., White, M.P., Liu, L., He, D., Pollard, K.S., Bruneau, B.G., and Srivastava, D. (2015). Human disease modeling reveals integrated transcriptional and epigenetic mechanisms of NOTCH1 haploinsufficiency. Cell *160*, 1072–1086.

Tohyama, S., Hattori, F., Sano, M., Hishiki, T., Nagahata, Y., Matsuura, T., Hashimoto, H., Suzuki, T., Yamashita, H., Satoh, Y., et al. (2013). Distinct metabolic flow enables large-scale purification of mouse and human pluripotent stem cell-derived cardiomyocytes. Cell Stem Cell *12*, 127–137.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR/Cas9 system. Science *343*, 80–84.

Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. Science *350*, 1096–1101.

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S., Muguruma, K., and Sasai, Y. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat. Biotechnol. *25*, 681–686.

Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331–338.

Cell Stem Cell 18, March 3, 2016 ©2016 Elsevier Inc. DOI: http://dx.doi.org/10.1016/j.stem.2016.02.013

SnapShot: Key Advances in hiPSC Disease Modeling

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

Cell Stem Cell

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

2008

- Amyotrophic Lateral Sclerosis (ALS) iPSCs (Dimos, 18669821)

2009

- Parkinson's disease iPSCs (Soldner, 19269371)
- Spinal muscular atrophy iPSCs (Ebert, 19098894)
- Functional cardiomyocytes (Zhang, 19213953)
- Vascular and hematopoietic cells (Lengerke, 19796250; Taura, 19423866; Choi, 19259936)

2010

- Cardiovascular progenitors (Moretti, 19850773; Blin, 20335662)
- Intestinal tissue (Spence, 21151107)
- Functional hepatic endoderm (Sullivan, 19877180)
- Modeling liver metabolic disorder (Rashid, 20739751)

2011

- Schizophrenia (Brennand, 21490598)
- Treatment of rat model of Parkinson's disease (Rhee, 21576821)
- Transplantable astroglia (Krencik, 21602806)
- Optimizing cardiomyocyte differentiation (Kattman, 21295278)
- Cardiomyocyte surface markers (Elliot, 22020065; Dubois, 22020386; Uosaki, 21876760)
- Self-organizing 3D optic cup (Eiraku, 21475194)
- Anterior foregut endoderm (Green, 21358635)
- Treatment of a rat model of arterial disease (Rufaihah, 21836062)

2012

- Drug screening for ALS (Egawa, 22855461)
- Probing sporadic and familial Alzheimer's disease (Israel, 22278060)
- Modeling cardiac sodium channel disease (Davis, 22647976)
- Blood-brain barrier endothelium (Lippmann, 22729031)
- Vascular smooth muscle disease susceptibility (Cheung, 22252507)

2013

- Screen for candidate ALS therapeutics (Yang, 23602540)
- Astroglia regulate neuronal development (Tang, 23759711)
- Modeling of late-onset disease via induced aging (Miller, 24315443)
- Spinal cord astrocytes (Roybon, 23994478)
- Cardiomyocyte maturation (Nunes, 23793239)
- Drug screening for disease-specific cardiomyocyte toxicity (Liang, 23519760)
- Cerebral organoids model brain development and microcephaly (Lancaster, 23995685)
- Pre-rosette neural stem cell culture system (Ebert, 23474892)
- Drug screening and gene correction for liver disease (Choi, 23325555)
- Small molecule expansion and differentiation of hepatocytes (Shan, 23728495)
- Vascular endothelium (Adams, 24052946)
- Endothelium with limited variation (White, 23079999)
- Vascularized, functional liver from organ bud (Takebe, 23823721)
- Modeling and rescue of Williams-Beuren syndrome (Kinnear, 23283491)

2014

- cGMP expansion method for neural stem cells (Shelley, 24962813)
- Hyperexcitability of ALS motor neurons (Wainger, 24703839)
- Automated analysis of engineered heart tissue (Stoehr, 24585781)
- Defined chemical conditions for cardiomyocyte generation (Burridge, 24930130)
- Disease modeling and drug screening for diabetic cardiomyopathy (Drawnel, 25437537)
- Heart-on-chip disease modeling of Barth syndrome cardiomyopathy (Wang, 24813252)
- Epicardium (Witty, 25240927)
- Retinal pigment epithelium cell sheets for clinical application (Kamao, 24527394)
- Self-organizing kidney (Takasato, 24335651)
- 3D retina with photoreceptors (Zhong, 24915161)
- Gastric organoids (McCracken, 25363776)
- 3D kidney via redefined nephron progenitors (Taguchi, 24332837)
- Lung and airway epithelium (Huang, 24291815)
- Functional β cells (Pagliuca, 25303535)
- Prediction of interindividual hepatocyte drug sensitivity (Takayama, 25385620)
- Cord-blood endothelium (Prasain, 25306246)
- Hutchinson-Gilford Progeria syndrome drug responses (Blondel, 24598781)

2015

- Lithium responses in bipolar disorder neurons (Mertens, 26524527)
- Atrial cardiomyocyte pharmacology assesment (Devalla, 25700171)
- Contractile defect from MYBPC3 mutation (Birket, 26489474)
- Cardiomyocyte contractility depends on shape and substrate stiffness (Ribeiro, 26417073)
- Expansion and patterning of cardiovascular progenitors (Birket, 26192318)
- Hepatocyte and cardio-toxicity high content screening (Grimm, 26539751)
- Cardiac drug screening system (Mathur, 25748532)
- Single contracting cardiomyocytes (Feaster, 26429802)
- Bioprinted 3D mini-livers (Faulkner-Jones, 26486521)
- Cholangiocytes for disease modeling and drug validation (Sampaziotis, 26167629)
- Cholangiocytes (Ogawa, 26167630)
- 3D cortical neurons and astrocytes (Paşca, 26005811)
- Predicting neural toxicity (Scwartz, 26392547)
- Lung organoids (Dye, 25803487)
- Multi-lineage kidney organoids for modeling nephrogenesis (Takasato, 26444236)
- CRISPR-mutant kidney organoids for disease modeling (Freedman, 26493500)
- Nephron organoids model kidney development and injury (Moriazane, 26458176)
- Cell-cell interactions enhance hepatocyte maturation (Berger, 25421237)
- Influenza in IRF7 deficiency (Ciancanelli, 25814066)
- Endothelial instruction of pancreas fate (Kao, 25601205)
- Engineered pulmonary vascular in decellularized lungs (Ren, 26368048)
- Hemogenic and arterial endothelium have distinct origins (Ditadi, 25915127)
- NOTCH1 haploinsufficiency modeling (Theodoris, 25768904)
- Vascularized organ bud via mesenchymal cell condensation (Takebe, 25891906)
- Modeling Duchenne muscular dystrophy in muscle fibers (Chal, 26237517)

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

A Decade of iPS Cells

R&D SYSTEMS
a bio-techne brand

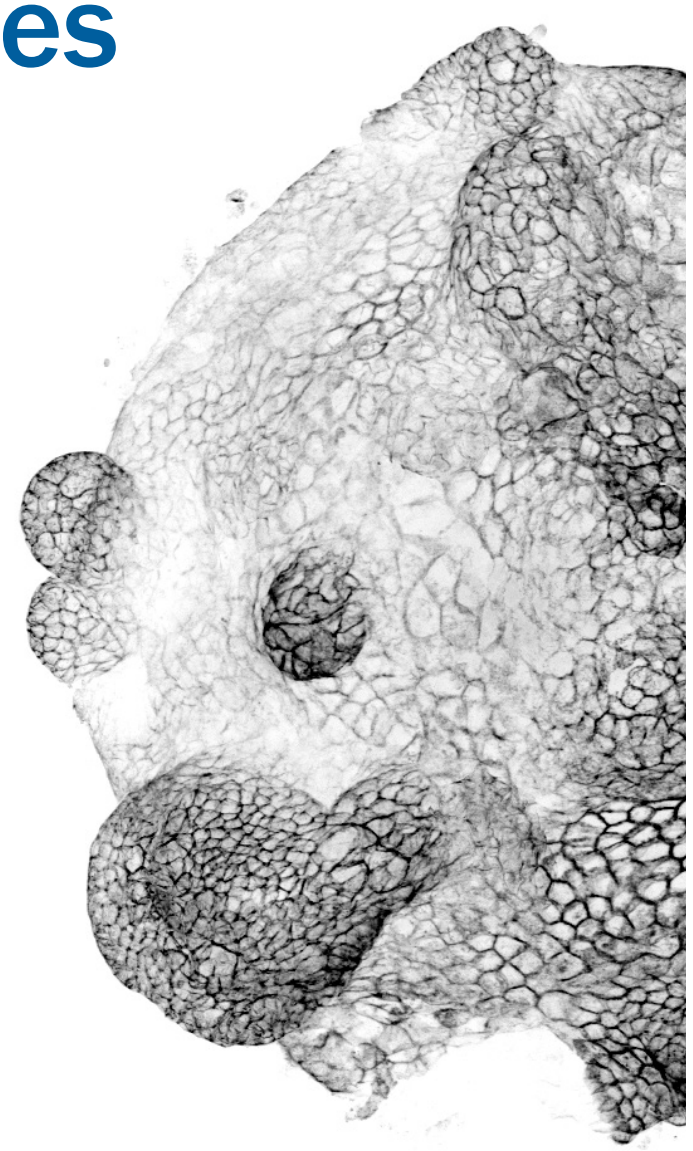
Reagents for Organoid and 3-D Cell Cultures

Reliably Generate Organoids for:

- Liver
- Intestine
- Brain
- Lung
- Pancreas
- Prostate
- Inner Ear

Using our High-Quality Reagents:

- Media supplements
- Matrix compounds
- Growth factors
- Small molecules
- Antibodies for cell characterization

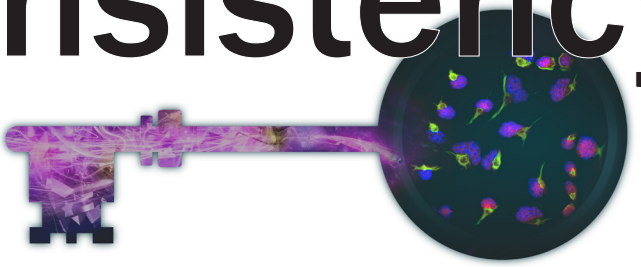


View Reagents | rndsystems.com/organoids

R&D SYSTEMS
a bio-techne brand

Get Your **Free Trial Vial** Today!

Consistency Is



N-2 MAX and N21-MAX Media Supplements for Neural and Stem Cell Cultures

Get your free trial vial | rndsystems.com/mediafreetrial

R&D SYSTEMS NOVUS BIOLOGICALS TOCRIS proteinsimple

bio-techne®
Global info@bio-techne.com bio-techne.com/find-us/distributors TEL +1 612 379 2956
North America TEL 800 343 7475 **Europe | Middle East | Africa** TEL +44 (0)1235 529449
China info.cn@bio-techne.com TEL +86 (21) 52380373
bio-techne.com