Title
Signaling and Transcriptional Mechanisms Regulating the Maintenance of and Reprogramming to the Pluripotent State

Author
Ho, Ritchie

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Signaling and Transcriptional Mechanisms Regulating the Maintenance of and Reprogramming to the Pluripotent State

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular Biology

by

Ritchie Ho

2013

ABSTRACT OF THE DISSERTATION

Signaling and Transcriptional Mechanisms Regulating the Maintenance of and Reprogramming to the Pluripotent State

by

Ritchie Ho

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2013

Professor Kathrin Plath, Chair

Induced pluripotent stem cells (iPSCs) are generated by overexpression of a combination of transcription factors. While this method to reprogram cells to the pluripotent state starts off using a defined number of factors, the complex cascade of events that it initiates during the inefficient process is far from being well defined. Therefore, many efforts are underway to characterize the molecular events involved in reprogramming to iPSCs. The work presented here aims to understand the mechanistic basis of reprogramming by examining biological processes that affect the maintenance of pluripotent cell populations. We find that the Wnt signaling pathway, which promotes the self-renewal of embryonic stem cells (ESCs), affects reprogramming efficiency in a stage-specific manner. We also find that while ESCs require the Polycomb repressive complexes (PRC1) and PRC2 to maintain the pluripotent state, components of the PRC1 and PRC2 are dispensable for reprogramming, revealing potential differences between how Polycomb group proteins may function in ESCs and during reprogramming. Finally, using a novel, single-cell secretion profiling platform, we discover that human ESC colonies are comprised of heterogeneous subpopulations that may serve to

coordinate colony growth and self-renewal. Together, these findings reveal complexities of regulating the establishment and maintenance of pluripotency that should be considered in future reprogramming studies.

The dissertation of Ritchie Ho is approved

Edward M. De Robertis

William E. Lowry

Bennett G. Novitch

Albert J. Courey

Kathrin Plath, Committee Chair

University of California, Los Angeles

2013

For Mom, Dad, Connie, and Seymour.

TABLE OF CONTENTS

Figures and Tables vii

Acknowledgements ix

Vita xii

Chapter 1

Introduction 1

References 7

Chapter 2

Mechanistic Insights into Reprogramming to Induced Pluripotency 9

References 40

Chapter 3

Stage-Specific Regulation of Reprogramming to Induced Pluripotent Stem Cells by Wnt Signaling and T Cell Factor Proteins 51

References 108

Chapter 4

Regulation of Reprogramming to Induced Pluripotency by Polycomb Repressive Complexes 111

References 148

Chapter 5

Multiplexed, Quantitative Analysis of Cytokine Secretion from Single Pluripotent Cells Reveals Heterogeneous, Secretory Subpopulations Throughout Colony Growth 151

References 183

Chapter 6

Conclusions and Outlook 185

References 191

FIGURES AND TABLES

Chapter 2

Table 2.1 Summary of Reprogramming Methods and Efficiencies 33

Table 2.2 Reprogramming Factor Replacements 35

Figure 2.1 Landmark Events on the Path to Induced Pluripotency 36

Figure 2.2 Considering Stochastic Events in Reprogramming 38

Chapter 3

Figure 3.1 Biphasic Role of Wnt Signaling in Reprogramming to iPSCs 78

Figure 3.2 Tcf1 and Lef1 Inhibit whereas Tcf3 and Tcf4 Promote the Early Phase of Reprogramming 80

Figure 3.3 *Tcf3/Tcf4* Depletion Enhances the Late Phase of Reprogramming in a *Tcf1/Lef1*-Dependent Manner 82

Figure 3.4 Ectopic Expression of Tcf3 Promotes Early Reprogramming Events but Inhibits the Induction of Pluripotency 84

Figure 3.5 *Tcf3* Ablation Allows Reprogramming in the Absence of Sox2 86

Figure 3.6 Regulation of OCK Reprogramming by Tcf3 98

Figure 3.7 Stepwise Modulation of *Tcf3* Levels Enables Efficient Reprogramming in the Absence of Sox2 90

Figure S3.1 Characterization of the Effect of Wnt Inhibition and Stimulation on Reprogramming 92

Figure S3.2 Characterization of the Requirement of Tcfs Early in Reprogramming 95

Figure S3.3 Characterization of the Requirement of Tcfs Late in Reprogramming 97

Figure S3.4 Characterization of the Tcf3 Overexpression Effect on Reprogramming 99

Figure S3.5 Characterization of Reprogramming in the Complete Absence of *Tcf3* 101

Figure S3.6 Characterization of *Tcf3*-Dependent Expression Changes in OCK Reprogramming 104
Figure S3.7 Characterization of OCK iPSCs Obtained by Stepwise Modulation of *Tcf3* Levels 106

Chapter 4

Figure 4.1 Ring1b Ablation Has No Effect on Reprogramming Efficiency 132
Figure 4.2 PRC2 is Required Throughout Reprogramming, But Ezh2 Activity is Dispensable in Late Stages 134
Figure 4.3 iPSCs Lacking Catalytic Ezh2 Can Be Established with Global Decrease in H3K27me3 136
Figure 4.4 Ectopic Expression of PRC2 Proteins Enhances Reprogramming 138
Figure S4.1 Characterization of PRC1 Expression During Reprogramming 140
Figure S4.2 Characterization of Ezh2 Expression and Inactivation Effects During Reprogramming 142
Figure S4.3 Characterization of iPSCs Lacking Catalytic Ezh2 144
Figure S4.4 Characterization of PRC2 Expression During Reprogramming 146

Chapter 5

Figure 5.1 Heterogeneous hESCs Bear Distinct Secretion Profiles 173
Figure 5.2 Secretion Intensity and Localization As a Function of Colony Size 175
Figure 5.3 Direct Intercellular Interaction Influences Secretion Profiles and Intensity 177
Figure 5.4 DKK3 Affects BMP2 and FGF4 Secretion 179
Figure S5.1 Selection of Reagents and Characterization of Culture Conditions 181

ACKNOWLEDGEMENTS

My enduring gratitude to those who have supported my endeavor to become an independently thinking scientist. Foremost, I thank my advisor, Professor Kathrin Plath, who to me, embodies dedication, commitment to high quality work, discipline, open mindedness, and a nurturing spirit. From her mentorship, I will continually strive for these ideals in my future professions. I thank the members of my doctoral committee, Professors William Lowry, Edward De Robertis, Bennett Novitch, and Albert Courey, for their critical evaluation, helpful guidance, and encouraging comments.

I am grateful to my colleagues, past and present, for making this journey bearable with their comradery in and out of the lab. Notably, I thank Rupa Sridharan, for whipping me into shape and infecting me with her laughter, Jason Tchieu for leading by example, and Bernadett Papp for allowing me to mature as a graduate student. Thanks to Victoria Ho, Kostas Chronis, Toh Hean Ch'ng, and Robin McKee for their companionship and countless diversions. I thank the rest of the Plath lab, my ACCESS 14 and Molecular Biology Institute classmates, and Biological Chemistry floormates for never providing a dull moment.

I will always be indebted to Professor Pamela Eversole-Cire, who unlocked the door to my career, and to Professor Benjamin Deneen, who fervently supported my goal of a doctoral education. I also thank Agnes Lukaszewicz, Wulf Haubensak, Sophia Vrontou, Gabriele Mosconi, Monica Martinez for these formative years.

Finally, to my special ones, I give my deepest appreciation for their patience and understanding throughout this process. I thank my dad for his proud and passionate support of my career path, my mom for filling me with love and my fridge with real food, my sister for

reminding me of what really matters, my niece and nephew for their joy, and my boyfriend for giving me everything I will ever need and more. We did it together!

Chapter 2 is a version of a publication authored by Ritchie Ho, Constantinos Chronis, and Kathrin Plath (Ho R, Chronis, C, Plath K. Mechanistic insights into reprogramming to induced pluripotency. *J Cell Physiol* 2010). We are grateful to Bill Lowry and Bernadett Papp for critical reading of the manuscript. R.H. is supported by a Training Grant of the National Institutes of Health 5T32AI060567-07, C.C. by a California Institute for Regenerative Medicine (CIRM) Training Award (TG2-01169), K.P. by the NIH Director's Young Innovator Award (DP2OD001686) and a CIRM Young Investigator Award (RN1-00564).

Chapter 3 is a version of a publication authored by Ritchie Ho, Bernadett Papp, Jackson A Hoffman, Bradley J. Merrill, and Kathrin Plath (Ho R, Papp B, Hoffman JA, Merrill BJ, Plath K. Stage-specific regulation of reprogramming to induced pluripotent stem cells by wnt signaling and T cell factor proteins. *Cell Rep* 2013) We thank Mark Chin, Sanjeet Patel, Rupa Sridharan, Robin McKee, Serena Lee, Amander Clark, and Gustavo Mostoslavsky for constructs, assistance with experiments, and data analysis and members of the Plath laboratory for helpful discussions. R.H. is supported by an NIH Training Grant (5T32AI060567-07) and the UCLA Graduate Division Dissertation Year Fellowship; K.P. is supported by the NIH (DP2OD001686 and P01 GM099134), CIRM (RN1-00564), and the UCLA Broad Center of Regenerative Medicine and Stem Cell Research; and B.J.M. is supported by the NIH (R01-CA128571).

Chapter 4 is a version of a manuscript in preparation for publication authored by Mark Chin, Constantinos Chronis, Ritchie Ho, Rupa Sridharan, Shahrad Hakimian, and Kathrin Plath.

Chapter 5 is a version of a manuscript in preparation for publication authored by Jun Wang, Ritchie Ho, Sanjeet Patel, Matthew Denholtz, James R. Heath, and Kathrin Plath.

VITA

EDUCATION

1/2001 – 5/2004	BA, University of California, Berkeley *College of Letters and Science* Molecular and Cell Biology, Neurobiology Minor: Asian American Studies GPA: 3.34

PUBLICATIONS

Ritchie Ho*, Bernadett Papp*, Jackson A. Hoffman, Bradley J. Merrill, and Kathrin Plath. (2013) Stage-specific regulation of reprogramming to induced pluripotent stem cells by Wnt and T cell factor proteins. Cell Reports 2013 6: 2113-2126 *equal authorship

Ritchie Ho*, Konstantinos Chronis*, and Kathrin Plath. (2010) The mechanistic basis of reprogramming. Journal of Cellular Physiology 2010 226: 868-878. *equal authorship

Benjamin Deneen, **Ritchie Ho**, Agnes Lukaszewicz, Christian J. Hochstim, Richard M. Gronostajski and David J. Anderson. (2006) The transcription factor NFIA controls the onset of gliogenesis in the developing spinal cord. Neuron 2006 52: 953-968.

AWARDS AND HONORS

2012	UCLA Graduate Division Dissertation Year Fellowship
2011	UCLA Molecular Biology Institute - Amgen Dissertation Award
2010	Paul D. Boyer Outstanding Teaching Award
2010	Molecular Biology Interdepartmental Program Poster Award
2008-2010	Viral Gene Therapy NIH Training Grant 5T32AI060567-07
2005	Special Recognition in Biocatalytics Award
2000/05/07	Kong Chow Benevolent Association Scholarship

MEETINGS AND PRESENTATIONS

2/2012	Gordon Research Conferences, Reprogramming Cell Fate, Galveston, TX (Poster)
10/2011	UCLA MBI Annual Retreat, Lake Arrowhead, CA (Speaker)
10/2011	World Stem Cell Summit, Pasadena, CA (Poster)
3/2011	Bridges to Stem Cells Graduate School Workshop, Pasadena City College, Pasadena, CA (Speaker/Panelist)
7/2010 and 4/2011	UCLA Virology and Gene Therapy Monthly Group Meeting, Los Angeles, CA (Speaker)
4/2010	Broad Institute Tri-Institutional Stem Cell Retreat, Asilomar, CA (Speaker)
5/2010	UCLA Biological Chemistry Annual Retreat, Los Angeles, CA (Speaker)
10/2009 and 10/2010	UCLA MBI Annual Retreat, Lake Arrowhead, CA (Poster)
10/2009 and 10/2012	UCLA Embryology Club, Los Angeles, CA (Speaker)
5/2009	UCLA Biological Chemistry Annual Retreat, Lake Arrowhead, CA (Poster)

CHAPTER 1

Introduction

Cellular reprogramming is the process of converting one cellular state, stabilized by an epigenetic layer of controlled gene expression, to another cellular state with an epigenetic profile that differs from that of the former. Somatic cell nuclear transfer experiments in frogs from over half a century ago demonstrated that this process is possible by converting the differentiated, somatic cellular state to an undifferentiated, totipotent state (Gurdon, 1962; Gurdon et al., 1958). In the intervening period, diverse methods of reprogramming have been developed, particularly in directing mammalian somatic cells to change their cellular state and functional capabilities into those of pluripotent, embryonic stem cells (ESCs) (Yamanaka, 2012). These cells are defined by their ability to give rise to all cell types of the adult body. The most refined of these cellular reprogramming methods was demonstrated by Kazutoshi Takahashi and Shinya Yamanaka, who induced pluripotent stem cells (iPSCs) from somatic cells using a defined formula of only four transcription factors that are highly expressed in ESCs: Oct4, Sox2, c-Myc, and Klf4 (Takahashi and Yamanaka, 2006). Furthermore, they and other groups including our own, demonstrated a significant molecular and functional equivalence between iPSCs and ESCs (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007).

This ability to induce pluripotency, due to the simplicity yet robustness of its technique, has since not only inspired a revolution in the field of developmental biology, but also renewed prospects in regenerative medicine and personalized cellular therapies. In recent years, many researchers worldwide have adopted this method of reprogramming to address a range of molecular and cellular, biological problems including basic questions in gene regulation to human disease modeling (Bellin et al., 2012; Papp and Plath, 2013). However, the iPSC method currently has a low efficiency of successfully converting a somatic cell to the pluripotent state, with less than 1% of cells transduced with the reprogramming factors giving rise to iPSCs. Many groups, including ours, are actively seeking to identify and characterize mechanisms that limit the efficiency of the reprogramming process in order to better understand and deconstruct the

control of cellular identity. The data presented in this dissertation describe our efforts to uncover such mechanisms regulating the stability and or transitions between cell states in the process of inducing pluripotency in somatic cells.

Chapter 2 of this dissertation provides an overview of what is currently known about the process of reprogramming somatic cells to iPSCs, particularly addressing how several intrinsic factors of the cells and different reprogramming conditions can impact the efficiency of the process. Firstly, it reviews basic aspects to the field of iPSC reprogramming, including a description of the many animal cell types amenable to induced pluripotency, techniques and reagents used to reprogram, and assays to assess the quality of pluripotency. Secondly, it discusses how characteristics of somatic cells such as the particular cell type and its differentiation potential can impact the efficiency of reprogramming and how some key experiments have fitted reprogramming between a deterministic versus a probabilistic model. Third, it describes landmark molecular events that occur in a particular order after the induction of reprogramming factors. Finally, at the heart of this chapter, is a discussion of several barriers that impede the progression towards induced pluripotency. These barriers include the ability of the reprogramming factors to bind to their correct genomic sites, the activation of key pluripotency genes that function to complete the reprogramming process, a plethora of chromatin marks that need to be reset to the ESC-specific landscape, the downregulation of cell cycle inhibitors, and optimal culturing conditions. The chapter concludes with an outlook on how the study of reprogramming mechanisms, as well as the molecular and functional equivalency between iPSCs and ESCs, will impact the future usage of iPSCs in regenerative medicine.

One type of barrier limiting the efficiency of reprogramming is the changing signaling requirements of cells as they transition from the somatic to the pluripotent state. Chapter 3 investigates the contribution of Wnt signaling to the reprogramming process. The Wnt signaling

pathway has been extensively characterized in many processes of cell biology including the maintenance of pluripotency, early embryonic lineage commitment, morphological patterning, tissue homeostasis, and cancer progression (Clevers and Nusse, 2012; Wray and Hartmann, 2012). Previous iPSC studies have noted minimal contribution by Wnt activation to the improvement of reprogramming efficiency (Marson et al., 2008; Takahashi and Yamanaka, 2006). By partitioning the reprogramming process into distinct early and late time windows, we discovered that the Wnt pathway has significant effects on reprogramming efficiency. Interestingly, these effects are opposite, and change depending on the time at which the Wnt pathway is stimulated in reprogramming cultures. Pursuing this biphasic model, the chapter further investigates the role of the transcriptional effectors of the canonical Wnt pathway and outlines the functional diversification and redundancies among them in the context of reprogramming.

Another reprogramming barrier is the global resetting of the somatic cell chromatin structure to resemble that of pluripotent cells. Chapter 4 explores the role of Polycomb repressive complexes (PRCs) in the reprogramming process, as they are widely studied chromatin modifying proteins that play a major role in regulating gene expression during development (Surface et al., 2010). PRCs are groups of proteins that covalently modify histone tails to signify transcriptional silencing and or chromatin compaction (Cao et al., 2002; Eskeland et al., 2010). In vertebrates, they are divided into subclasses PRC1 and PRC2, with many subunits within each subclass having multiple homologs to diversify function across several cell types (Simon and Kingston, 2009). While PRC1 and PRC2 bind to and cooperatively regulate a set of developmental genes in ESCs (Stock et al., 2007), they also have non-overlapping functions in both embryonic and somatic cells (Sauvageau and Sauvageau, 2010; Surface et al., 2010). While the earliest reprogramming studies demonstrated that the epigenetic profile of histone marks are reset to resemble an ESC-like pattern upon reprogramming of somatic cells

to iPSCs, particularly for H3K27me3, which is catalyzed by PRC2 (Maherali et al., 2007; Mikkelsen et al., 2008), the requirement for PRC1 and PRC2 subunits in reprogramming has remained uncharacterized. The experiments in Chapter 4 provide an initial attempt to define which PRC1 and PRC2 components are required for reprogramming to iPSCs. Among these, we find a changing requirement for Ezh2, the catalytic subunit of PRC2, during the reprogramming process, where it is required early on in reprogramming, but dispensable later on. iPSCs generated in a genetic background of catalytically inactive Ezh2 provide a system in which we can dissect where in the genome Ezh2 acts during reprogramming to mediate H3K27me3. Moving forward with these studies, we hope to distinguish the specialized functions from the overlapping functions of the PRC subunits during the reprogramming process.

Upon successful reprogramming of human somatic cells, iPSCs resemble ESCs both molecularly and functionally (Lowry et al., 2008; Yu et al., 2007), however, much remains unknown about how these cells propagate their pluripotent state in culture. Chapter 5 begins to explore how they might coordinate intercellular signaling to achieve this self renewal. Utilizing state-of-the-art technologies in microfluidics, we are able to analyze a multiplex panel of secreted proteins from a large number of single, dissociated somatic and pluripotent cells in order to determine whether cells in these populations have distinct secretion profiles. Strikingly, we discovered that human pluripotent stem cell colonies exist as heterogeneous subpopulations of cells secreting different profiles of cytokines from diverse pathways including IGFBP2, DKK3, BMP2, and FGF4. The percentage, as well as localization of these secretory subpopulations vary with colony size. Furthermore, *in silico* and functional experiments suggest that the secretory identity of a cell can be modulated based on signaling interactions with another cell. A more detailed characterization of heterogeneity in pluripotent cell cultures may provide a better understanding, and thereby improve the efficiency, of the reprogramming process.

Finally, Chapter 6 provides a summary and discussion of the results presented in this dissertation as a whole, the thematic concepts across chapters, and an outlook for future reprogramming studies in light of the findings described here.

REFERENCES

Bellin, M., Marchetto, M.C., Gage, F.H., and Mummery, C.L. (2012). Induced pluripotent stem cells: the new patient? Nature Reviews Molecular Cell Biology *13*, 713-726.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Clevers, H., and Nusse, R. (2012). Wnt/beta-catenin signaling and disease. Cell *149*, 1192-1205.

Eskeland, R., Leeb, M., Grimes, G.R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultchi, A.I., Wutz, A.*, et al.* (2010). Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. Molecular Cell *38*, 452-464.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. Journal of Embryology and Experimental Morphology *10*, 622-640.

Gurdon, J.B., Elsdale, T.R., and Fischberg, M. (1958). Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. Nature *182*, 64-65.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proceedings of the National Academy of Sciences of the United States of America *105*, 2883-2888.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55-70.

Marson, A., Foreman, R., Chevalier, B., Bilodeau, S., Kahn, M., Young, R.A., and Jaenisch, R. (2008). Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell *3*, 132-135.

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.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Papp, B., and Plath, K. (2013). Epigenetics of reprogramming to induced pluripotency. Cell *152*, 1324-1343.

Sauvageau, M., and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell *7*, 299-313.

Simon, J.A., and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nature Reviews Molecular Cell Biology *10*, 697-708.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nature Cell Biology *9*, 1428-1435.

Surface, L.E., Thornton, S.R., and Boyer, L.A. (2010). Polycomb group proteins set the stage for early lineage commitment. Cell Stem Cell *7*, 288-298.

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.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

Wray, J., and Hartmann, C. (2012). WNTing embryonic stem cells. Trends Cell Biol. *22*, 159-68.

Yamanaka, S. (2012). Induced pluripotent stem cells: past, present, and future. Cell Stem Cell *10*, 678-684.

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.

CHAPTER 2

Mechanistic Insights into Reprogramming to Induced Pluripotency

ABSTRACT

Induced pluripotent stem cells (iPSCs) can be generated from various embryonic and adult cell types upon expression of a set of few transcription factors, most commonly consisting of Oct4, Sox2, cMyc, and Klf4, following a strategy originally published by Takahashi and Yamanaka (Takahashi and Yamanaka, 2006). Since iPSCs are molecularly and functionally similar to embryonic stem cells (ESCs), they provide a source of patient-specific pluripotent cells for regenerative medicine and disease modeling, and therefore have generated enormous scientific and public interest. The generation of iPSCs also presents a powerful tool for dissecting mechanisms that stabilize the differentiated state and are required for the establishment of pluripotency. In this review, we discuss our current view of the molecular mechanisms underlying transcription factor-mediated reprogramming to induced pluripotency.

INTRODUCTION

Somatic cells can be reprogrammed to induced pluripotent stem cells (iPSCs) by the delivery of a few pluripotency-related transcription factors. Since the original description of iPSCs in Shinya Yamanaka's landmark report (Takahashi and Yamanaka, 2006), studies of transcription factor-induced reprogramming to the iPSC state have branched into two rapidly moving fields of research. First, no longer hindered by the technical and ethical limitations associated with somatic cell nuclear transfer (SCNT) and cell fusion, reprogramming via the Yamanaka approach provides a new avenue to investigate basic questions of cellular plasticity and pluripotency. Secondly, the iPSC technology enables the derivation of patient- and disease-specific pluripotent stem cell lines, which has widened the door to disease modeling, drug discovery, and cell replacement strategies.

Both of these branches of iPSC research are affected by the inefficiency of the reprogramming process (Table 2.1). Despite the variety of recent publications reporting DNA-

free or integration-free reprogramming via protein delivery of the reprogramming factors or the use of RNA viruses, the most efficient generation of iPSCs is based on genomic integration of DNA encoding the reprogramming factors, most commonly through lenti- or retroviral transduction (Table 2.1). The use of most iPSCs is therefore thought to be affected by genomic alterations that could lead to phenotypic artifacts arising from insertional mutagenesis or expression of the oncogenic reprogramming factors (Hochedlinger et al., 2005; Okita et al., 2007; Nakagawa et al., 2008; Wernig et al., 2008b). The hope is that a better understanding of the reprogramming process will lead to improved, more efficient reprogramming technologies that do not require genomic integration, linking the two major avenues of reprogramming research. Similarly, a better general understanding of how a small set of transcription factors can reset the epigenetic landscape of cells, gained from the reprogramming process, could also further the development of rational differentiation strategies for pluripotent cells for use in therapeutic applications in regenerative medicine.

Despite the numerous reports demonstrating tactics to boost the efficiency of reprogramming, the molecular requirements as well as barriers of the reprogramming process are only beginning to be defined. Many studies are looking for small molecules, miRNAs, siRNAs, or growth factors in efforts to substitute individual reprogramming factors to lower the need for genomic integration while allowing efficient reprogramming (Table 2.2). Others aim at uncovering pathways that are essential for the induction of pluripotency and contribute to overcoming reprogramming barriers. Changes in transcription factor function, chromatin state, and extracellular signaling are among the many obstacles a reprogramming cell must face in order to destabilize the somatic program and eventually establish a pluripotent state. This chapter aims to summarize the most recent studies describing the molecular events taking place during the reprogramming process, and to discuss the mechanistic obstacles proposed to limit the rate and efficiency of faithful conversion to pluripotency.

REPROGRAMMING BASICS

iPSCs have been generated upon ectopic expression of Oct4, Sox2, cMyc, and Klf4 from a number of species including human (Takahashi et al., 2007; Yu et al., 2007; Lowry et al., 2008; Park et al., 2008), mouse (Takahashi and Yamanaka, 2006; Maherali et al., 2007b; Okita et al., 2007; Wernig et al., 2007), rat (Li et al., 2009b), pig (Wu et al., 2009), and rhesus monkey (Liu et al., 2008), and many different cell types such as fibroblasts, terminally differentiated lymphocytes and other blood cells, stomach and liver cells, neural progenitors, keratinocytes, melanocytes, and pancreatic β cells (Aasen et al., 2008; Aoi et al., 2008; Hanna et al., 2008; Stadtfeld et al., 2008a; Kim et al., 2008b; Utikal et al., 2009a). While cMyc, Klf4, and Sox2 can be replaced in the reprogramming process by close homologs and small molecules, Oct4 appears more pivotal and so far could only be efficiently replaced by its upstream regulator, the orphan nuclear receptor Nr5a2 (Table 2.2 and references therein). The diversity of cell types and species that have been reprogrammed and the general applicability of the four original reprogramming factors suggests a generic fashion in which the four factors act and indicates that there probably is no cell type-specific barrier that cannot be overcome by the action of the reprogramming factors leading to an evolutionary conserved pluripotency network. Nevertheless, the starting cell type can alter the dependence on the reprogramming factors, efficiency, and kinetics. For example, the high expression level of endogenous Sox2 and moderate levels of cMyc and Klf4 in neural precursor cells (NPCs) allow reprogramming with only Oct4, albeit very slowly (Kim et al., 2009b). Addition of ectopic Sox2 may even interfere with the reprogramming of NPCs (Eminli et al., 2008; Silva et al., 2008), suggesting that there are ideal levels of the reprogramming factors in relation to each other to induce pluripotency.

In a typical reprogramming experiment, ectopic expression of Oct4, Sox2, cMyc, and Klf4 in the starting cell type leads to downregulation of somatic gene expression and formation

of embryonic stem cell (ESC)-like colonies, culminating in the upregulation of an ESC-like gene expression program and pluripotent capabilities after 2–4 weeks (Takahashi and Yamanaka, 2006; Brambrink et al., 2008; Mikkelsen et al., 2008; Stadtfeld et al., 2008b; Chan et al., 2009; Sridharan et al., 2009; Li et al., 2010; Samavarchi-Tehrani et al., 2010; Figure 2.1). As indicated by the low reprogramming efficiency (Table 2.1), most cells that receive and express the reprogramming factors, and their daughter cells, do not progress to the faithfully reprogrammed state indicating major epigenetic barriers to reprogramming. It should be noted though, that the reported efficiencies depend hugely on the criteria with which iPSC colonies are scored, whether all starting cells are considered or only those cells carrying all four reprogramming factors, and whether proliferation is taken into consideration.

To identify and/or quantify faithful reprogramming events, the best strategy generally is to observe the induced expression of endogenously encoded pluripotency markers such as Oct4 and Nanog (Maherali and Hochedlinger, 2008). In the mouse system, researchers often take advantage of pluripotency reporter cell lines generated via knockin approaches in ESCs. In human reprogramming experiments, staining for pluripotency surface markers have been applied successfully to identify faithful reprogramming events (Lowry et al., 2008; Chan et al., 2009). Alternatively, it has been proposed that iPSC colony number can be assessed using morphological criteria or alkaline phosphatase staining when combined with inducible vectors coding for the reprogramming factors, as only cells that have entered the pluripotent state become independent of the ectopic factors while cells that have not progressed into pluripotency remain dependent on transgene expression and regress to their somatic state upon transgene silencing (Brambrink et al., 2008; Stadtfeld et al., 2008b; Wernig et al., 2008a).

Upon faithful reprogramming, mouse and human iPSCs are similar to their respective ESC counterparts in terms of gene expression and genome-wide distribution of epigenetic

marks (Maherali et al., 2007b; Okita et al., 2007; Takahashi et al., 2007; Wernig et al., 2007; Mikkelsen et al., 2008; Chin et al., 2009; Hawkins et al., 2010). In agreement with their molecular similarity to ESCs, reprogrammed cells also satisfy a range of functional pluripotency assays (Jaenisch and Young, 2008). These include differentiation into the three germlayers in vitro and in teratomas, and, for mouse iPSCs, contribution to chimera with germline transmission and the most stringent pluripotency assay of all, tetraploid (4N) complementation, which allows the derivation of adult mice solely from iPSCs (Jaenisch and Young, 2008; Boland et al., 2009; Kang et al., 2009; Zhao et al., 2009). However, not all mouse iPSC lines support tetraploid complementation and the inability to do so has recently been associated with inappropriate silencing of the imprinted *Dlk1–Dio3* gene cluster on mouse chromosome 12qF1 (Stadtfeld et al., 2010), indicating that reprogramming also results in aberrant epigenetic programming. Interestingly, germline competence of mouse iPSCs appears improved upon expression of additional transcription factors like Tbx3 during the reprogramming process (Han et al., 2010), although it remains unclear why this would be the case. Of course, the analysis of the developmental potential of human iPSCs, as with human ESCs, is limited to teratoma formation and in vitro differentiation.

DIFFERENTIATION POTENTIAL INFLUENCES REPROGRAMMING EFFICIENCY AND KINETICS

In the early days of reprogramming to induced pluripotency, it was thought that the low reprogramming efficiency seen within a couple of weeks after delivery of the reprogramming factors was due to the lack of expression of all four factors in the same cell, as four individual retroviruses each coding for one reprogramming factor were typically used. However, the current use of polycistronic lentiviral cassettes argues against the idea of heterogeneous transgene expression as the main cause of the low reprogramming efficiency as it only slightly increases the number of faithfully reprogrammed colonies over the system that uses individual

retroviruses (Chang et al., 2009; Sommer et al., 2009; Table 2.1).

Similar results were obtained with secondary reprogramming systems (Hockemeyer et al., 2008; Maherali et al., 2008; Wernig et al., 2008a; Carey et al., 2009; Stadtfeld et al., 2009; Woltjen et al., 2009; Table 2.1). These systems entail the generation of primary iPSC clones with inducible reprogramming cassettes and their subsequent differentiation in vitro, or via chimera formation upon blastocyst injection in vivo, in the absence of transgene expression to obtain genetically modified, homogeneous somatic cell populations. These cells can then be induced to re-express the reprogramming factors across the entire population to generate secondary iPSC clones. Even the establishment of the “reprogrammable” mouse with a single defined integration site for an inducible, polycistronic reprogramming factor cassette enabled only slightly more efficient reprogramming of fibroblasts compared to viral or transposon methods (Carey et al., 2009; Stadtfeld et al., 2009). However, the inducible secondary reprogramming system or reprogrammable mouse have several benefits as they allow reprogramming of cell types that are typically difficult to transduce and enable the comparison of reprogramming efficiencies of different somatic cell types from the same mouse.

Notably, experiments with cells from “reprogrammable” mice support the conclusion that insertional mutagenesis is not a key driver of reprogramming as predicted from non-integrative reprogramming studies and mapping of viral insertion sites (Aoi et al., 2008; Varas et al., 2009; Winkler et al., 2010). Consequently, one of the key questions in the reprogramming field has been whether only a particular subset of cells within the starting population, for example more undifferentiated cells such as adult stem cells or progenitors, possess the ability to reach the iPSC state.

To test whether fully differentiated cell types can be reprogrammed, the generation of

iPSC lines from mature mouse B cells was successfully achieved. Resulting iPSC lines contained specific immunoglobulin re-arrangements reflecting their origin from mature B cells and gave rise to mouse progeny with a monoclonal immune system (Hanna et al., 2008). However, upregulation of the myeloid transcription factor CCAAT/enhancer-binding proteins-α (CeBPα), which can reprogram B cells into macrophage-like cells (Xie et al., 2004), or downregulation of the transcription factor Pax5, an essential regulator of mature B cell development (Cobaleda et al., 2007), were necessary for iPSC induction from mature B cells. Subsequently, the establishment of iPS lines from mature human and mouse B and T lymphocytes without the need to modulate these blood-specific transcription factors was reported (Eminli et al., 2009; Loh et al., 2010; Seki et al., 2010; Staerk et al., 2010). The reasons behind the different requirements for mature B cell reprogramming in these studies remain unclear but may be related to the particular reprogramming system used, the expression levels of the reprogramming factors, and/or culture conditions.

In one of these studies, the authors then investigated how the developmental state affects reprogramming efficiency by isolating blood cells at various differentiation stages and found that progenitor and hematopoietic stem cells give rise to iPS colonies with a much higher efficiency and in less time compared to differentiated cells of the same lineage (10–28% vs. 0.03–0.5%; Eminli et al., 2009). As the enhancement in reprogramming was independent of cell division rate, these data argue against a model positing strictly that only few starting cell types, particularly more undifferentiated states, are susceptible to reprogramming, but rather suggest that the degree of differentiation influences reprogramming efficiency and kinetics. A similar correlation of differentiation state and reprogramming potential has been made in SCNT experiments (Hochedlinger and Jaenisch, 2006).

ALMOST ALL CELLS IN A POPULATION CAN GIVE RISE TO DAUGHTER CELLS THAT FORM IPSCS

From the results described above, it still remained unclear why only a select number of cells, even when expressing similar levels of the reprogramming factors, become iPSCs within 2–3 weeks. An interesting question was therefore to elucidate whether every cell in the starting population has the potential to eventually, that is, after longer exposure to the reprogramming factors, give rise to iPSCs. By clonally expanding individual pre-B cells or monocytes from a secondary reprogramming system in 96-well plates and screening the progeny derived from each cell for its ability to reprogram, it was shown that almost every starting cell ultimately gives rise to daughter cells that can reach the iPSC state (Hanna et al., 2009b). In agreement with the previously observed low reprogramming efficiency, the first reprogramming events could only be observed after 8–10 days in 3–5% of wells. However, prolonged exposure of doxycycline for up to 18 weeks and constant passaging led to around 92% of wells to become Nanog positive at variable times.

The finding that the timing of faithful reprogramming varies widely among cells indicates that at least one event driving the reprogramming process, if not more, is stochastic in nature (Hanna et al., 2009b; Figure 2.2A). Interestingly, an experimentally induced, increased proliferation rate of pre-B cells, through p53 or p21 inhibition or Lin28 overexpression, accelerates the reprogramming process (Hanna et al., 2009b). However, the authors also showed that Nanog overexpression in the same pre-B cells increased their reprogramming rate without altering cell cycle kinetics, indicating that the reprogramming rate per cell cycle can be enhanced. It is important to note though that, even by the end of these clonal reprogramming experiments at around 18 weeks, not all daughter cells within each clonal population were faithfully reprogrammed, but typically just a few, irrespective of an identical genetic background. These findings highlight that there are major epigenetic barriers that interfere with the

reprogramming of most cells in the culture and can be overcome by events that are stochastic in nature (Figure 2.2B).

Combining the fact that almost every pre-B cell in the culture has the potential to give rise to at least a few daughter cells that faithfully reprogram, with the result that, at least for the blood lineage, differentiation state influences both reprogramming efficiency and kinetics, all somatic cells may be amendable to reprogramming, but more undifferentiated cells in the population have a higher probability to overcome reprogramming barriers. However, experiments with clonal reprogramming assays of cells at different differentiation stages, which consider proliferation rate, plating efficiency, and transgene levels, need to be performed to eventually test this hypothesis. Such experiments may be complicated by the fact that different cell types require different cytokines in vitro that could directly alter the reprogramming process and many other variables.

MOLECULAR EVENTS DURING REPROGRAMMING

The detailed events occurring between the time of exogenous expression of the reprogramming factors and the establishment of the iPSC state are only slowly uncovered. This is primarily due to the low efficiency and slow kinetics of the process, and the fact that cells that will successfully complete the reprogramming process cannot be preselected. However, populations that give rise to iPSCs with higher efficiencies can be enriched from intermediate stages of reprogramming (Stadtfeld et al., 2008b).

Several groups have chronologically traced events that occur during the first 2–3 weeks upon induction of the reprogramming factors in mouse embryonic fibroblasts (MEFs; Figure 2.1). The first change in gene expression is the downregulation of somatic markers including key mesenchymal genes (Mikkelsen et al., 2008; Stadtfeld et al., 2008b; Li et al., 2010;

Samavarchi-Tehrani et al., 2010). Concomitantly, epithelial genes like E-cadherin become upregulated as cells undergo a mesenchymal-to-epithelial transition (MET) and start proliferating (Mikkelsen et al., 2008; Li et al., 2010; Samavarchi-Tehrani et al., 2010). Undergoing MET is an essential early step of reprogramming as activation of Tgfβ signaling, inhibition of BMP signaling, or depletion of MET genes such as E-cadherin interfere with reprogramming to induced pluripotency (Li et al., 2010; Samavarchi-Tehrani et al., 2010). Using high resolution time-lapse imaging to backtrack faithful reprogramming events, an increase in proliferation rate and a concomitant decrease in cell size were confirmed in all successful reprogramming cases as early events, followed by stereotypic colony formation 4–8 days later leading to iPSC clones (Smith et al., 2010).

It has also been reported that embryonic markers such as alkaline phosphatase (AP) and the stage-specific embryonic antigen-1 (SSEA-1) cell surface marker are induced relatively early in the reprogramming process in a subset of cells (Brambrink et al., 2008; Stadtfeld et al., 2008b). Some cells from the SSEA-1 positive subpopulation then give rise to faithfully reprogrammed cells and activate the expression of endogenously encoded Oct4, Sox2, and Nanog, which are considered the most stringent markers of complete reprogramming (Maherali et al., 2007b; Stadtfeld et al., 2008b). In faithfully reprogramming cells, many other pluripotency-related genes, that is genes highly expressed in ESCs and/or functionally important for the establishment and maintenance of the pluripotent state, are upregulated at around this point as well (Mikkelsen et al., 2008; Chan et al., 2009; Samavarchi-Tehrani et al., 2010). Meanwhile, SSEA-1 negative cells are depleted for iPSCs as judged at a parallel time point, suggesting a defined order of events in the reprogramming process.

When properly reprogrammed, cells can sustain the pluripotent state independently of ectopic reprogramming factor expression indicating a stable conversion of cell fate and the

establishment of the pluripotency network (Brambrink et al., 2008; Stadtfeld et al., 2008b). Depletion of the exogenous factors at earlier times of reprogramming leads the cells to revert back to a differentiated cell phenotype (Brambrink et al., 2008; Stadtfeld et al., 2008b). When using retroviruses, silencing of the integrated retroviral transgenes occurs rather efficiently (Maherali et al., 2007b), possibly due to the actions of Trim28, Zfp806, and histone and DNA methyltransferases as the pluripotent state is established (Lei et al., 1996; Wolf and Goff, 2009; Matsui et al., 2010; Rowe et al., 2010). Retroviral silencing is not absolutely necessary to establish an autonomous, self-renewing ESC state, since iPSCs can be generated using constitutively active lentiviral vectors to deliver the reprogramming factors (Brambrink et al., 2008; Sommer et al., 2010). However, differentiation of these cells can be severely impaired if the expression levels of the reprogramming factors remain high (Brambrink et al., 2008; Sommer et al., 2010). It should be noted that in female mouse cells, the somatically silenced X chromosome is reactivated during the reprogramming process (Maherali et al., 2007b), while in human iPSCs the inactive X chromosome is maintained from the somatic state (Tchieu et al., 2010), pointing at differences between the mouse and human reprogramming process (Figure 2.1).

Collectively, these data suggest that successful reprogramming to an iPSC clone within 2–3 weeks upon initial induction of the reprogramming factors follows a defined sequence of steps, each only taken successfully by few cells (Figures 2.2B and 2.2C). Why it typically takes at least 8–10 days to detect the first complete reprogramming event and whether the same path from one to the next reprogramming event is taken by each cell that ultimately will undergo faithful reprogramming remains unclear (Figure 2.2C). A finer resolution of intermediate cell states will address these question and reveal whether reprogramming simply reverses normal development and follows through a line of progenitor steps. Single cell studies will be key to address this problem (Chan et al., 2009; Smith et al., 2010). Studies of X chromosome

reactivation during mouse reprogramming may be particularly helpful as chromatin changes accompanying X chromosome silencing during development are well defined and because this process affects an entire chromosome which can be easily visualized in single cells.

Even though the first reprogramming events occur in 2–3 weeks, the pre-B cell experiment described above demonstrated that most reprogramming events will occur much later, even as late as 18 weeks post-induction of the reprogramming factors (Figure 2.2A). To explain this variable latency of induction of pluripotency, one can then propose a model in which one early, stochastically timed step determines the overall kinetics of each individual reprogramming event and all subsequent steps occur in a stereotypic fashion. Alternatively, each step may be stochastically timed, and transitions from one step to the next may not follow the same molecular path. Of course, these models need not to be mutually exclusive.

BARRIERS OF REPROGRAMMING

Exploring the pre-iPSC State

A key task during the last couple years has been to identify molecular barriers of reprogramming that explain why only few cells, even in a clonal population, reprogram. A few groups, including our own, have been able to isolate a relatively stable, intermediate population of cells coined partially induced or reprogrammed pluripotent stem (pre-iPS) cells. These cells arise from fibroblasts 2–3 weeks after induction of the reprogramming factors as SSEA1-positive colonies with an ESC-like morphology and are capable of clonal expansion (Mikkelsen et al., 2008; Silva et al., 2008; Sridharan et al., 2009). Comparison of the transcription profiles of pre-iPS and ES/iPSCs revealed that many endogenous pluripotency genes such as Oct4 and Nanog are not reactivated, while somatic markers are already efficiently silenced (Mikkelsen et al., 2008; Sridharan et al., 2009). In agreement with the notion that reactivation of the somatically silenced X chromosome during mouse cell reprogramming occurs with similar timing

as that of Oct4 and Nanog, the X is still inactive in mouse pre-iPSCs (Silva et al., 2008; Sridharan et al., 2009). Furthermore, pre-iPSCs are likely stabilized by ectopic expression of the four reprogramming factors and express a subset of genes that are neither active in MEFs or ES/iPSCs. Interestingly, pre-iPSCs obtained from neural precursor and B-cell reprogramming experiments appear stalled at a similar stage as those derived from fibroblasts suggesting that the reprogramming pathways from different somatic cells channel into comparable pre-iPS states (Mikkelsen et al., 2008).

While it is not absolutely clear that pre-iPSCs represent an intermediate that occurs transiently during the reprogramming process, they are not simply an aborted reprogramming artifact because pre-iPSCs can convert into iPSCs upon addition of ERK and GSK inhibitors (termed 2i), which also leads to the stabilization of Nanog protein levels (Silva et al., 2008; Sridharan et al., 2009). Similarly, the DNA demethylating agent 5-azacytidine (Mikkelsen et al., 2008), ascorbic acid addition (Esteban et al., 2010), or Tgfβ-inhibition (Ichida et al., 2009), each facilitate the conversion of pre-iPSCs into fully reprogrammed clones. Given that all these treatments also improve the efficiency and kinetics of the reprogramming process when starting from somatic cells and the fact that pre-iPSCs transcriptionally mirror a late intermediate of the reprogramming process, these findings support the use of pre-iPS as a useful platform for the identification of pathways that will allow the enhancement of final steps of reprogramming (Figure 2.1). Interestingly, various pre-iPSC clones react differently to a range of small molecule stimuli in their ability to convert to iPSCs (Ichida et al., 2009), suggesting that there are so far unappreciated molecular differences among them that will be interesting to uncover in the future.

Reprogramming Factor Binding in pre-iPSCs and iPSCs

An analysis of Oct4, Sox2, Klf4, and cMyc target genes by chromatin immunoprecipitation in combination with microarrays (ChIP-chip) in mouse iPS and pre-iPSCs provided significant insight into the action of the reprogramming factors (Sridharan et al., 2009). In iPSCs, just like in ESCs, cMyc and the trio of pluripotency transcription factors Oct4, Klf4, and Sox2 form largely separable transcriptional networks, based on the finding that they target largely non-overlapping sets of genes in these cells (Chen et al., 2008; Kim et al., 2008a; Sridharan et al., 2009). cMyc binds many genes involved in cellular metabolism, cell cycle regulation, and biosynthetic pathways, while the majority of Oct4, Sox2, and Klf4 targets encode developmental, transcriptional regulators. Interestingly, many ES/iPSC targets of Oct4, Sox2, and Klf4 are not bound by these transcription factors in pre-iPSCs (Sridharan et al., 2009). A correlation of binding with expression data suggested that the lack of Oct4, Sox2, and Klf4 binding to pluripotency-related genes in pre-iPSCs is responsible for the lack of reactivation of these genes at this state. On the contrary, cMyc targets of ES/iPSC are often already bound and strongly expressed in pre-iPSCs (Sridharan et al., 2009; Figure 2.1). Collectively, these findings suggest that the transcriptional network downstream of cMyc becomes induced early in the reprogramming process and that the activation of key pluripotency-associated genes occurs via binding of Oct4, Sox2, and Klf4 only at the end of the reprogramming process.

In agreement with this conclusion, expression of only cMyc in MEFs induces a gene expression profile most similar to that of ESCs when compared to the profiles of MEFs induced to express any of the other reprogramming factors individually (Sridharan et al., 2009). Additionally, upregulation of cMyc and Klf4 in fibroblasts before induction of Oct4 and Sox2 positively enhances reprogramming while pre-expression of Oct4 and Sox2 has no such effect (Markoulaki et al., 2009). Interestingly, cMyc is not essential as reprogramming factor but enhances the efficiency and kinetics of the process (Nakagawa et al., 2008; Wernig et al., 2008b). Thus, the main functions of ectopically expressed cMyc can, directly or indirectly, be

taken over by Oct4, Sox2, and Klf4, which may simply induce higher expression of endogenously encoded Myc genes. Even though ectopic cMyc function may have a powerful role early in the reprogramming process, it is likely that it is also essential for later steps as well as it needs to maintain expression of its downstream target network in the pluripotent state, shown by its role in ESC self-renewal.

Nanog Expression Lowers Barriers During Final Steps of Reprogramming

The inability of Oct4, Sox2, and Klf4 to bind to their iPS/ESC target gene promoters until late in the reprogramming process was proposed to present a major obstacle to faithful reprogramming to pluripotency (Sridharan et al., 2009) and it is interesting to speculate on the causes. One explanation may be that additional ESC-specific transcription factors are necessary to synergize with Oct4, Sox2, and Klf4 to allow binding and activation of pluripotency gene promoters. The key pluripotency regulator Nanog would be a great example for such a factor as it co-binds many of the Oct4, Sox2, and Klf4 targets in ESCs (Boyer et al., 2005; Loh et al., 2006), biochemically interacts with Oct4 and other transcription factors in ESCs (Wang et al., 2006), and is only upregulated during the pre-iPS to iPSC transition at the end of the reprogramming process (Mikkelsen et al., 2008; Silva et al., 2009; Sridharan et al., 2009). Nanog's absence in pre-iPSCs could therefore affect the binding pattern of Oct4, Sox2, and Klf4 in these cells. In agreement with this idea, Nanog positively affects reprogramming by cell fusion and facilitates human and mouse reprogramming (Silva et al., 2006; Yu et al., 2007; Hanna et al., 2009b). Similarly, Tgfβ-inhibition, which enhances the pre-iPS to iPSC transition, may also occur through upregulation of Nanog expression (Ichida et al., 2009). Importantly, Smith and colleagues demonstrated that Nanog is absolutely essential for the full induction of pluripotency and required only during the final stages of the reprogramming process (Silva et al., 2009). Similar to Nanog, other pluripotency transcription factors not yet strongly expressed in pre-iPSCs such as Sall4 also boost reprogramming efficiency (Tsubooka et al., 2009) and

biochemically interact with the Oct4 transcriptional network (Wang et al., 2006; Liang et al., 2008; Yang et al., 2008). Future studies aimed at understanding why the four reprogramming factors target different genes in pre-iPSCs in comparison to ES/iPSCs should yield insights into the transcriptional regulation of the pluripotency program, the role of chromatin (see below), and perhaps lead to new methods of increasing the conversion efficiency into iPSCs.

Chromatin State as a Reprogramming Barrier

Reprogramming to pluripotency is associated with a major resetting of the chromatin landscape, which is deemed necessary in order to activate the ESC-specific transcriptional program while silencing tissue-specific genes. Accordingly, genome-wide analyses of several histone methylation marks (H3K9me2, H3K4me3, H3K27me3) and DNA methylation in fibroblasts, pre-iPS and iPSCs indicated that substantial changes in these modifications occur during reprogramming and that they are reset to an ESC-like pattern (Maherali et al., 2007b; Mikkelsen et al., 2008; Chin et al., 2009; Doi et al., 2009; Sridharan et al., 2009; Hawkins et al., 2010). Likewise, asymmetric cytosine methylation that is only prevalent in ESCs is re-established upon faithful reprogramming based on candidate gene analysis (Lister et al., 2009).

A wide range of publications suggests that re-establishing ESC-like chromatin marks is critical for the reprogramming process. The analysis of promoter DNA methylation profiles of the Nanog and Oct4 loci revealed that these genes are hypermethylated in MEFs and pre-iPSCs in direct contrast to the unmethylated status observed in iPS and ESCs (Maherali et al., 2007a; Mikkelsen et al., 2008). Notably, this demethylation appears to occur very late in the reprogramming process and serves as a good indicator of faithful reprogramming events. In agreement with the notion that DNA methylation may interfere with efficient reprogramming, the addition of the DNA demethylating drug 5-azacytidine or depletion of maintenance methyl transferase Dnmt1 increases the efficiency with which iPSCs are generated from pre-iPSC lines

(Mikkelsen et al., 2008). When added at different times during the reprogramming process, only 5-azacytidine treatment during the final phase enhanced the process suggesting that it largely is advantageous at the late intermediate stage (Mikkelsen et al., 2008). However, one should be careful when interpreting such results given that inhibition of DNA methylation induces apoptosis in differentiated cells (Jackson-Grusby et al., 2001), but not in the pluripotent state (Li et al., 1992), and hence perhaps acts by changing population dynamics at earlier stages of reprogramming.

Interestingly, the DNA demethylation machinery centered on the AID protein has recently been implicated as an essential player in ESC–somatic cell fusion experiments (Bhutani et al., 2010), but direct targets and the question of how the enzyme is recruited still remain unclear, and its role in iPSC production has not yet been tested. Nevertheless, this finding suggests that active DNA demethylation is important to allow reprogramming. Alternatively, passive mechanisms may be in place to lower the DNA methylation content in pluripotency promoters, potentially acting in S-phase when the DNA methylation mark gets copied during DNA replication.

Similar to DNA methylation, it has been suggested that histone modifications impact the reprogramming process. For example, a subset of pluripotency genes lacks the typically activating histone H3 lysine 4 methylation in MEFs and pre-iPSCs but not in ES/iPSCs (Sridharan et al., 2009). Since histone H3 K4 methylation and DNA methylation are typically mutually exclusive (Meissner et al., 2008; Zhang et al., 2009), the lack of this modification may correlate strongly with the presence of DNA methylation at pluripotency promoters in pre-iPSCs and MEFs. This repressive chromatin environment could preclude Oct4, Sox2, and Klf4 binding at pluripotency loci in pre-iPSCs.

Further evidence for the role of chromatin in reprogramming comes from the result that histone deacetylase inhibitors including TSA, valproic acid (VPA), suberoylanilide hydroxamic acid (SAHA), and butyrate (Huangfu et al., 2008a; Liang et al., 2010), as well as the small molecule BIX (Shi et al., 2008), supposed to inhibit the histone H3 K9 methyltransferase G9A, enhance the production of iPSCs. It should be noted though, that all these small molecules that target either histone or DNA modifiers likely also act on the global chromatin status. Therefore, they could not only directly affect the chromatin state at key promoters or enhancers, but also be critical for the reprogramming process in a more indirect manner. However, addition of VPA enables replacement of cMyc in mouse and cMyc and Klf4 in human reprogramming experiments (Huangfu et al., 2008a,b), suggesting a functional overlap between these reprogramming factors and histone acetylation. The fact that cMyc interacts with histone acetyl transferases further supports this idea (McMahon et al., 2000; Frank et al., 2003).

Of course, chromatin-based regulation is not limited to the direct modification of DNA and histones. Chd1, a chromatin remodeling enzyme belonging to the chromodomain family with a SNF2-like helicase domain, is essential for the production of iPSCs (Gaspar-Maia et al., 2009). Similarly, overexpression of chromatin remodelers with an ESC-specific expression component, like the SWI/SNF-type BAF complex, enhances reprogramming efficiency and kinetics (Singhal et al., 2010), but again insights into the mechanisms involved are lacking at this point. Future studies of the localization of chromatin marks, of the enzymes that establish or remove these marks or remodel chromatin, and of reprogramming factor binding during the reprogramming process are necessary to reveal the detailed path by which chromatin governs reprogramming. One interesting question will be whether the chromatin state at pluripotency genes changes first to allow reprogramming factor binding or whether the reprogramming factors bind first to then change the chromatin state by recruiting the respective chromatin modifying machinery.

Nuclear Architecture as a Roadblock to Reprogramming

In the mouse system, distinct pluripotent stem cell populations have been generated from embryos at different developmental stages (Nichols and Smith, 2009). Epiblast stem cells (EpiSCs), obtained from the d5.5 egg cylinder express many pluripotency-associated genes also present in blastocyst-derived ESCs (e.g., Oct4 and Sox2) and, like ESCs, can differentiate into the three germ layers in vitro and in teratomas (Brons et al., 2007; Tesar et al., 2007). Notable differences between these two pluripotent stages, however, include differential signaling pathway dependence with EpiSCs relying on Activin A and bFGF and ESCs depending on Lif, the inability of EpiSCs to contribute to chimeras upon blastocyst injection, and silencing of one X chromosome in female EpiSCs but not in ESCs. EpiSCs therefore represent a developmentally more advanced, "primed" pluripotent state, while ESCs display "naïve" pluripotency (Nichols and Smith, 2009). Upon switching from bFGF/activin to Lif-containing media, ectopic expression of Nanog, Klf4, or cMyc, addition of small molecules that can replace these reprogramming factors, or enhancement of Lif signaling, mouse EpiSCs reprogram to the pluripotent capabilities of ESCs (Bao et al., 2009; Guo et al., 2009; Hanna et al., 2009a; Silva et al., 2009; Yang et al., 2010). Notably, for the field of induced pluripotency, the conversion of EpiSCs to the ES-like state is also typically characterized by a low efficiency, with only around 1% converting. While the optimal set of reprogramming factors or culture conditions for efficient EpiSC to ESC conversion may not yet be identified, this finding points to major barriers that even limit reprogramming from primed to naïve pluripotency and suggests that some aspects of the cellular identity of EpiSCs is more similar to somatic, differentiated cells than iPS/ESCs.

In agreement with this conclusion, the mapping of DNA replication timing has revealed that EpiSCs are more similar to committed cells types than to ES/iPSCs (Hiratani et al., 2009). Furthermore, within the nucleus of EpiSCs, there is an accumulation of compacted chromatin

near the periphery that is typically found in somatic cells but not in ESCs, and the Oct4 gene is localized more peripherally in EpiSCs than in ESCs (Hiratani et al., 2009). Together, these data suggest that a global event that can reorder replication timing profiles and subnuclear architecture is required to permit induction of naïve pluripotency. Interestingly, like EpiSCs, pre-iPSCs are more somatic cell-like in their genome organization and replication timing signature (Hiratani et al., 2009). Potentially, nuclear reorganization is difficult to achieve and represents a major barrier to the reactivation of pluripotency genes towards the end of the reprogramming process. Thus, EpiSCs, just like pre-iPSCs, are a useful tool to study the induction of naïve pluripotency, particularly since fewer factors are required for this conversion. As a side note, naïve human ESCs have so far not been derived from human blastocysts and the typical human ESC appears to be in the primed mouse EpiSCs state. Recently, Jaenisch and colleagues demonstrated that human ES/iPSCs can acquire a naïve pluripotent phenotype upon ectopic expression of KLF4 with OCT4 or KLF2 and culture in 2i media with LIF, but again, the efficiency of this conversion is low (Hanna et al., 2010).

Apoptosis and Senescence as Barriers of Fibroblast Reprogramming

In contrast to many somatic cells that often have a limited proliferative potential and undergo stress-induced senescence, ESCs self-renew indefinitely in culture and are therefore considered immortal. Intriguingly, senescence appears to be incompatible with reprogramming to pluripotency and could represent another barrier to the process. In support of this idea, serial passaging of fibroblasts is associated with a dramatic decrease in reprogramming efficiency (Utikal et al., 2009b). Several reports have now demonstrated that the downregulation of tumor suppressor components such as p53, p21 (Cdkn1a), p16 (Ink4a), and p19 (Arf) in fibroblasts enhances the efficiency and kinetics by which iPSCs are generated, either via inducing immortalization or interfering with apoptosis (Banito et al., 2009; Hong et al., 2009; Kawamura et al., 2009; Li et al., 2009a; Marion et al., 2009; Utikal et al., 2009b). In agreement with these

findings, iPSCs, like ESCs, normally express low levels of p53 and Arf. The fact that another study that used pre-B cells instead of fibroblasts argued that p53 depletion acts by enhancing cell cycle kinetics (Hanna et al., 2009b) may indicate that p53 acts differently in these cell types, or could be the result of the particular reprogramming scheme employed or differing reprogramming factor levels.

The use of knockout serum replacement (KSR) enhances the generation of mouse iPSCs in comparison to FBS-containing culturing media (Blelloch et al., 2007). Vitamin C (ascorbic acid), known for its antioxidant function, is contained in KSR and may be the key mediator of this effect since treatment of reprogramming cultures with vitamin C enhances both efficiency and kinetics of the process (Esteban et al., 2010). Surprisingly, this effect is largely independent of a modulation of reactive oxygen species, but may act through lowering p53 levels (Esteban et al., 2010). Given that reprogramming in the absence of p53, the guardian of genome integrity, enriches for damaged cells that are not desirable for clinical use (Marion et al., 2009), vitamin C treatment may be a safer choice to boost the reprogramming process, particularly when applying human cells for disease studies and cell replacement strategies. Similarly, hypoxic conditions can be applied during reprogramming to lower p53 levels in a subtle manner to enhance the reprogramming process (Utikal et al., 2009b; Yoshida et al., 2009).

Thus, the studies of pre-iPSCs and the reprogramming process have yielded considerable insight into the mechanism and barriers of the reprogramming process, albeit the detailed steps often remain unclear.

EPIGENETIC MEMORY AND CONTINUING CHANGES OF THE IPSC STATE

Recent evidence suggests that there are differences between ES and iPSCs that are indicative of an epigenetic memory of the starting cell in iPSCs and affect the differentiation potential of iPSCs (Kim et al., 2010; Polo et al., 2010). Specifically, when first derived, iPSCs still express genes and display a chromatin state resembling aspects of the starting cell and the cells appear to differentiate with higher efficiency into the originating lineage in comparison to other lineages.

These molecular and functional disparities between ES and iPSCs seem to disappear upon continued expansion of iPSCs (Chin et al., 2009; Chin et al., 2010; Polo et al., 2010), but whether this process selects for a more properly reprogrammed cell or whether reprogramming continues in culture is unclear. Similarly, the use of drugs that affect epigenetic function like TSA and 5-azacytidine erase the epigenetic memory (Kim et al., 2010). In comparison to SCNT-derived ESCs, iPSCs may carry a stronger epigenetic memory of their cell type of origin than SCNT-derived ESCs (Kim et al., 2010; Polo et al., 2010) suggesting that different mechanisms are responsible for those two reprogramming methods.

CONCLUSIONS

At present, the reprogramming field is paced by the synergistic relationship between studies of mechanisms and the development of improved reprogramming methods that are more conducive for clinical applications and disease modeling. While the exploration of intermediate states proves to be an invaluable tool for improving our understanding of the process, much work is needed in order to elucidate the exact mechanism of reprogramming and get a handle on whether each cell follows the same path to pluripotency, what controls the latency of reprogramming, and how different cellular origins affect the process. Recently, small molecule screens revealed modulators of the reprogramming process (Ichida et al., 2009; Lyssiotis et al., 2009) suggesting that reprogramming without protein factors may be feasible in

the near future. Another question that remains to be answered is whether reprogramming selects for cells that genetically inactivate the p53/Arf pathway or similar pathways, that when inactivated permanently are potentially detrimental to the iPSC state or their differentiated progeny. In any case, reprogramming to pluripotency via the Yamanaka approach has shown that a small set of transcription factors can dramatically modulate cell fate. We are beginning to see how new transcription factor combinations are being defined that can convert one somatic cell type to another without going through the pluripotent state (Zhou et al., 2008; Ieda et al., 2010; Vierbuchen et al., 2010) and it will be interesting to understand the similarities and differences of these processes.

Table 2.1. Summary of Reprogramming Methods and Efficiencies

Reprogramming efficiency is typically determined 2–4 weeks after reprogramming factor induction. By and large, efficiency is defined by the number of iPS clones (scored by different criteria) generated from a starting number of somatic cells. The method of efficiency calculation varies widely across labs.

Table 2.1

Reprogram scheme	Delivery/expression method	Advantage	Disadvantage	Donor cell type	Reprogramming efficiency (%)	Criteria used to score efficiency	References
Integrative	Retrovirus	• Good efficiency • Silenced transgenes in iPS cells	• Multiple genomic integrations	Various cell types	0.05-0.1	Drug resistance from pluripotency marker	Takahashi and Yamanaka 2006, Okita et al. 2007, Wernig et al. 2007
	Constitutive lentivirus	N/D	• Multiple genomic integrations • Differentiation potentially affected by incomplete transgene silencing		N/D	N/D	Brambrink et al. 2008
	Inducible lentivirus	• Good efficiency • Controlled expression	• Multiple genomic integrations (less when using polycistronic cassettes)		0.5	GFP reporter expression from pluripotency marker	Brambrink et al. 2008, Stadtfeld et al. 2008b, Sommer et al. 2009
	Secondary inducible lentiviral system	• Good efficiency • Controlled expression • No viral delivery necessary and expression of all factors in most cells • Allows reprogramming of cells that are typically difficult to transduce	• Multiple genomic integrations • Requires differentiation *in vivo* or *in vitro* and screening of several primary iPS cell lines		0.02 - 28* *myeloid progenitors		Wernig et al. 2008a, Hanna et al. 2008, Eminli et al. 2009
	Inducible, reprogrammable mouse		• Only applicable to mouse system • Requires targeting in ES cells and blastocyst injection	Mouse fibroblasts and blood	0.01-40* *hematopoietic stem cells	ES-like morphology, Nanog expression	Carey et al. 2009, Stadtfeld et al. 2010
Integrative/ excisable	LoxP-flanked lentiviral cassettes	• Good efficiency • No reprogramming factors left behind	• Analysis of Cre-mediated excision can be labor intensive • Viral and loxP sequences left behind	Mouse and human fibroblasts	0.005-0.01 (0.5 for mouse)	GFP reporter expression from pluripotency marker, Tra-1-60 and NANOG staining	Soldner et al. 2009, Sommer et al. 2010
	Transposons (piggyBac)	• Good efficiency • No vector sequence left behind	• Analysis of transposase-mediated excision can be labor intensive	Mouse and human fibroblasts	0.04 - 2.5	ES-like morphology, GFP reporter expression from pluripotency marker	Woltjen et al. 2009, Kaji et al. 2009
Non-integrative, DNA-based	Minicircle	• No genome modification	• Low efficiency • Possible integration • Screening for possible integration can be labor intensive	Human adult adipose stem cells	0.005	ES-like morphology	Jia et al. 2010
	Plasmid transfection			Mouse fibroblasts	0.000015	GFP reporter expression from pluripotency marker	Okita et al. 2008
	Episomal vector			Human foreskin fibroblasts	0.000006	ES-like morphology	Yu et al. 2009
	Adenovirus			Mouse hepatocytes	0.0005	GFP reporter expression from pluripotency marker	Stadtfeld et al. 2008
				Human embryonic fibroblasts	0.0002	ES-like morphology	Zhou and Freed, 2009
Non-integrative, DNA-free	ES protein extract	• No genome modification	• Low efficiency • Bacterial protein scheme requires VPA	Mouse cardiac and skin fibroblasts	0.001	ES-like morphology	Cho et al. 2010
	HEK293 expressed transducible protein			Human newborn fibroblasts	0.001	Alkaline phosphotase staining	Kim et al. 2009
	Bacterially expressed transducible protein			Mouse fibroblasts	0.006	GFP reporter expression from pluripotency marker	Zhou et al. 2009
	Sendai vectors (RNA)		• Requires continuous passaging to eliminate transgenes from iPS cells	Human foreskin fibroblasts	0.001-0.1	Alkaline phosphotase staining	Fusaki et al. 2009

Table 2.2. Reprogramming Factor Replacements

Listed are RNAs and proteins that are able to replace individual reprogramming factors. For a comprehensive review of small molecule replacers of reprogramming, see Feng et al. (2009b).

Original reprogramming factor(s)	Substitute factor	Proposed mechanism	References
c-Myc	n-Myc or l-Myc	Homologues that may be redundant in function. l-Myc has lower transformation activity.	Nakagawa et al. 2008, 2010
	Wnt3a	c-Myc is a target of Wnt signaling.	Marson et al. 2008
	miR-291-3p, -294, or -295	miR290 cluster is a c-Myc target.	Judson et al. 2009
Klf4	Klf1, Klf2, or Klf5	Homologues that may be redundant in function.	Nakagawa et al. 2008
	Esrrb or Esrrg	Esrrb regulates Klf4 expression, has similar binding targets as Oct4, Sox2, and Klf4, and forms a protein complex with Oct4.	Feng et al. 2009, van den Berg et al. 2008
Sox2	Sox1, Sox3, Sox15, or Sox18	Homologues that may be redundant in function.	Nakagawa et al. 2008
	Nanog	Nanog has similar binding targets as Sox2.	Ichida et al. 2009
Sox2 and c-Myc	Smad7	Inhibits Tgfβ signaling, promoting MET.	Li et al. 2010
Oct4	Nr5a1 or Nr5a2	Binds enhancer and promoter of Oct4 and regulates its expression.	Heng et al. 2010

Figure 2.1. Landmark Events on the Path to Induced Pluripotency

It is thought that expression of the four reprogramming factors Oct4 (O), Klf4 (K), Sox2 (S), and cMyc (M) triggers a cascade of events within two to three weeks that leads to the iPSC state. Stable and clonally expandable pre-iPSCs can also be generated as a by-product of this process and can enter the fully re-programmed state upon addition of indicated molecules.

Figure 2.1

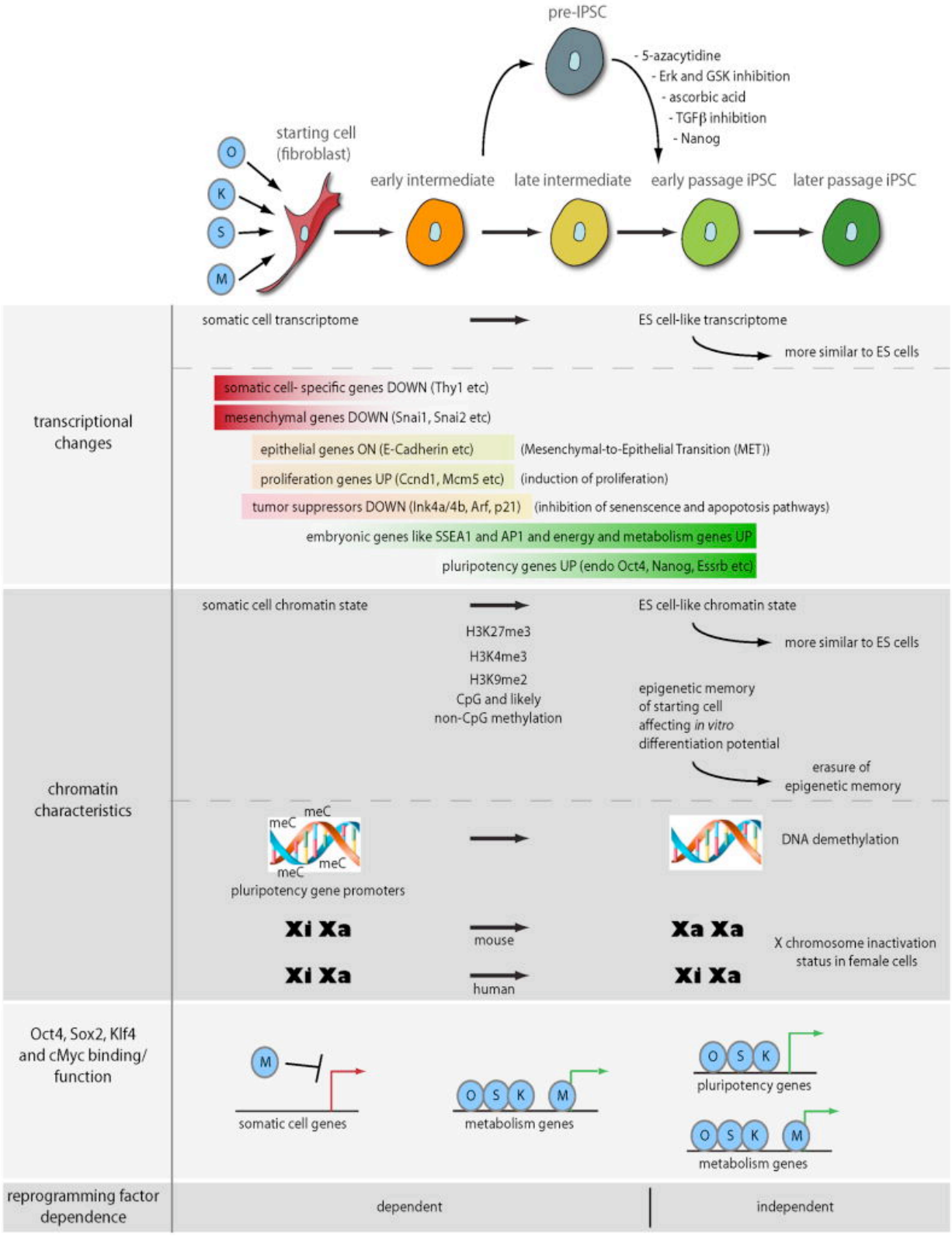

pre-IPSC
- 5-azacytidine
- Erk and GSK inhibition
- ascorbic acid
- TGFβ inhibition
- Nanog
starting cell (fibroblast)
O
K
S
M
early intermediate
late intermediate
early passage iPSC
later passage iPSC
transcriptional changes
somatic cell transcriptome
ES cell-like transcriptome
more similar to ES cells
somatic cell- specific genes DOWN (Thy1 etc)
mesenchymal genes DOWN (Snai1, Snai2 etc)
epithelial genes ON (E-Cadherin etc)
(Mesenchymal-to-Epithelial Transition (MET))
proliferation genes UP (Ccnd1, Mcm5 etc)
(induction of proliferation)
tumor suppressors DOWN (Ink4a/4b, Arf, p21)
(inhibition of senenscence and apopotosis pathways)
embryonic genes like SSEA1 and AP1 and energy and metabolism genes UP
pluripotency genes UP (endo Oct4, Nanog, Essrb etc)
chromatin characteristics
somatic cell chromatin state
H3K27me3
H3K4me3
H3K9me2
CpG and likely non-CpG methylation
ES cell-like chromatin state
more similar to ES cells
epigenetic memory of starting cell affecting in vitro differentiation potential
erasure of epigenetic memory
meC
pluripotency gene promoters
DNA demethylation
Xi Xa
mouse
Xa Xa
Xi Xa
human
Xi Xa
X chromosome inactivation status in female cells
Oct4, Sox2, Klf4 and cMyc binding/ function
somatic cell genes
metabolism genes
pluripotency genes
metabolism genes
reprogramming factor dependence
dependent
independent

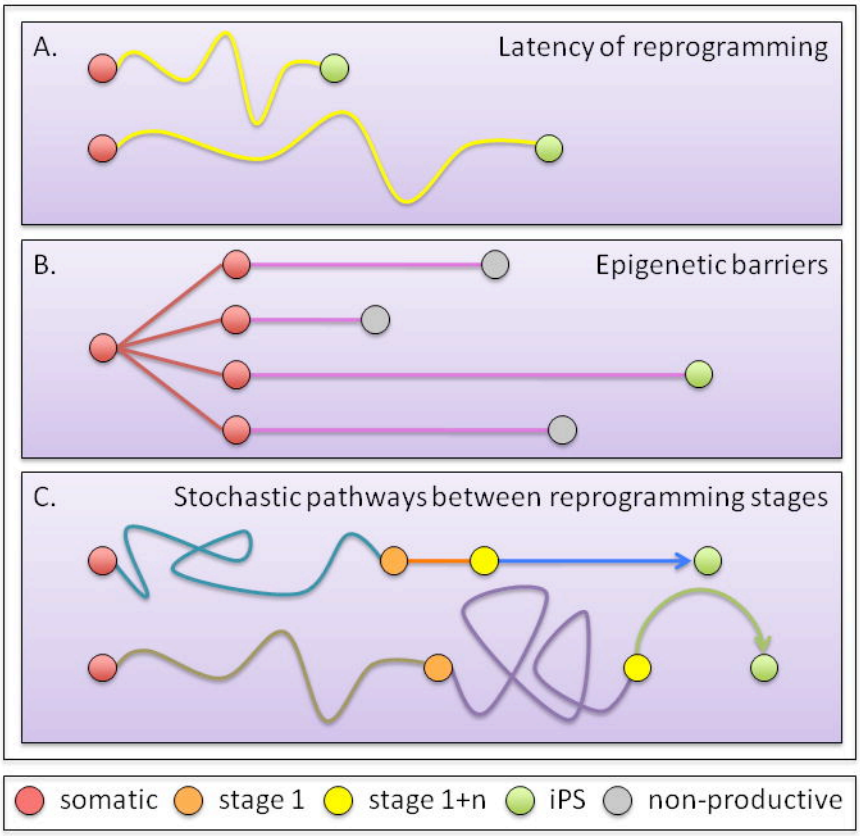
Figure 2.2. Considering Stochastic Events in Reprogramming

A. Latency of faithful reprogramming. All cells have the potential to give rise to daughter cells that can faithfully reprogramming, but conversion to iPSCs occurs at divergent times - as early as two weeks (top) to as late as 18 weeks for pre-B cells (bottom).

B. Epigenetic barriers. Within a clonal populations, only a few daughter cells reach the iPSC status due to epigenetic barriers that need to be overcome in a stochastic manner.

C. Pathways between reprogramming stages. It is thought that reprogramming occurs in defined steps (see Figure 2.1) but could use alternative pathways between these steps. The scenarios presented are not exhaustive, as other potential reprogramming profiles can exist.

Figure 2.2


A.
Latency of reprogramming
B.
Epigenetic barriers
C.
Stochastic pathways between reprogramming stages
somatic
stage 1
stage 1+n
iPS
non-productive

REFERENCES

Aasen T, Raya A, Barrero MJ, Garreta E, Consiglio A, Gonzalez F, Vassena R, Bilic J, Pekarik V, Tiscornia G, Edel M, Boue S, Izpisua Belmonte JC. 2008. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat Biotechnol 26:1276–1284.

Aoi T, Yae K, Nakagawa M, Ichisaka T, Okita K, Takahashi K, Chiba T, Yamanaka S. 2008. Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science 321:699–702.

Banito A, Rashid ST, Acosta JC, Li S, Pereira CF, Geti I, Pinho S, Silva JC, Azuara V, Walsh M, Vallier L, Gil J. 2009. Senescence impairs successful reprogramming to pluripotent stem cells. Genes Dev 23:2134–2139.

Bao S, Tang F, Li X, Hayashi K, Gillich A, Lao K, Surani MA. 2009. Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature 461:1292–1295.

Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. 2010. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature 463:1042–1047.

Blelloch R, Venere M, Yen J, Ramalho-Santos M. 2007. Generation of induced pluripotent stem cells in the absence of drug selection. Cell Stem Cell 1:245–247.

Boland MJ, Hazen JL, Nazor KL, Rodriguez AR, Gifford W, Martin G, Kupriyanov S, Baldwin KK. 2009. Adult mice generated from induced pluripotent stem cells. Nature 461:91–94.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122:947–956.

Brambrink T, Foreman R, Welstead GG, Lengner CG, Wernig M, Suh H, Jaenisch R. 2008. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell 2:151–159.

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448:191–195.

Carey BW, Markoulaki S, Beard C, Hanna J, Jaenisch R. 2009. Single-gene transgenic mouse strains for reprogramming adult somatic cells. Nat Method 7:56–59.

Chan EM, Ratanasirintrawoot S, Park IH, Manos PD, Loh YH, Huo H, Miller JD, Hartung O, Rho J, Ince TA, Daley GQ, Schlaeger TM. 2009. Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat Biotechnol 27:1033–1037.

Chang CW, Lai YS, Pawlik KM, Liu K, Sun CW, Li C, Schoeb TR, Townes TM. 2009. Polycistronic lentiviral vector for “hit and run” reprogramming of adult skin fibroblasts to induced pluripotent stem cells. Stem Cells (Dayton, Ohio) 27:1042–1049.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G,

Sung WK, Clarke ND, Wei CL, Ng HH. 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133:1106–1117.

Chin MH, Mason MJ, Xie W, Volinia S, Singer M, Peterson C, Ambartsumyan G, Aimiuwu O, Richter L, Zhang J, Khvorostov I, Ott V, Grunstein M, Lavon N, Benvenisty N, Croce CM, Clark AT, Baxter T, Pyle AD, Teitell MA, Pelegrini M, Plath K, Lowry WE. 2009. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell 5:111–123.

Chin MH, Pellegrini M, Plath K, Lowry WE. 2010. Molecular analyses of human induced pluripotent stem cells and embryonic stem cells. Cell Stem Cell 7:263–269.

Cho HJ, Lee CS, Kwon YW, Paek JS, Lee SH, Hur J, Lee EJ, Roh TY, Chu IS, Leem SH, Kim Y, Kang HJ, Park YB, Kim HS. 2010. Induction of pluripotent stem cells from adult somatic cells by protein-based reprogramming without genetic manipulation. Blood 116:386–395.

Cobaleda C, Jochum W, Busslinger M. 2007. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. Nature 449:473–477.

Doi A, Park IH, Wen B, Murakami P, Aryee MJ, Irizarry R, Herb B, Ladd-Acosta C, Rho J, Loewer S, Miller J, Schlaeger T, Daley GQ, Feinberg AP. 2009. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet 41:1350–1353.

Eminli S, Utikal J, Arnold K, Jaenisch R, Hochedlinger K. 2008. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. Stem Cell (Dayton, Ohio) 26:2467–2474.

Eminli S, Foudi A, Stadtfeld M, Maherali N, Ahfeldt T, Mostoslavsky G, Hock H, Hochedlinger K. 2009. Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. Nat Genet 41:968–976.

Esteban MA, Wang T, Qin B, Yang J, Qin D, Cai J, Li W, Weng Z, Chen J, Ni S, Chen K, Li Y, Liu X, Xu J, Zhang S, Li F, He W, Labuda K, Song Y, Peterbauer A, Wolbank S, Redl H, Zhong M, Cai D, Zeng L, Pei D. 2010. Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell 6:71–79.

Feng B, Jiang J, Kraus P, Ng JH, Heng JC, Chan YS, Yaw LP, Zhang W, Loh YH, Han J, Vega VB, Cacheux-Rataboul V, Lim B, Lufkin T, Ng HH. 2009a. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat Cell Biol 11:197–203.

Feng B, Ng JH, Heng JC, Ng HH. 2009b. Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. Cell Stem Cell 4:301–312.

Frank SR, Parisi T, Taubert S, Fernandez P, Fuchs M, Chan HM, Livingston DM, Amati B. 2003. MYC recruits the TIP60 histone acetyltransferase complex to chromatin. EMBO Rep 4:575–580.

Fusaki N, Ban H, Nishiyama A, Saeki K, Hasegawa M. 2009. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc Jpn Acad Ser B Phys Biol Sci 85:348–362.

Gaspar-Maia A, Alajem A, Polesso F, Sridharan R, Mason MJ, Heidersbach A, Ramalho-Santos J, McManus MT, Plath K, Meshorer E, Ramalho-Santos M. 2009. Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature 460:863–868.

Guo G, Yang J, Nichols J, Hall JS, Eyres I, Mansfield W, Smith A. 2009. Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development (Cambridge, England) 136:1063–1069.

Han J, Yuan P, Yang H, Zhang J, Soh BS, Li P, Lim SL, Cao S, Tay J, Orlov YL, Lufkin T, Ng HH, Tam WL, Lim B. 2010. Tbx3 improves the germ-line competency of induced pluripotent stem cells. Nature 463:1096–1100.

Hanna J, Markoulaki S, Schorderet P, Carey BW, Beard C, Wernig M, Creyghton MP, Steine EJ, Cassady JP, Foreman R, Lengner CJ, Dausman JA, Jaenisch R. 2008. Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. Cell 133:250–264.

Hanna J, Markoulaki S, Mitalipova M, Cheng AW, Cassady JP, Staerk J, Carey BW, Lengner CJ, Foreman R, Love J, Gao Q, Kim J, Jaenisch R. 2009a. Metastable pluripotent states in NOD-mouse-derived ESCs. Cell Stem Cell 4:513–524.

Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, Creyghton MP, van Oudenaarden A, Jaenisch R. 2009b. Direct cell reprogramming is a stochastic process amenable to acceleration. Nature 462:595–601.

Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, Soldner F, Cassady JP, Muffat J, Carey BW, Jaenisch R. 2010. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl Acad Sci USA 107:9222–9227.

Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, Pelizzola M, Edsall LE, Kuan S, Luu Y, Klugman S, Antosiewicz-Bourget J, Ye Z, Espinoza C, Agarwahl S, Shen L, Ruotti V, Wang W, Stewart R, Thomson JA, Ecker JR, Ren B. 2010. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell 6:479–491.

Heng JC, Feng B, Han J, Jiang J, Kraus P, Ng JH, Orlov YL, Huss M, Yang L, Lufkin T, Lim B, Ng HH. 2010. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. Cell Stem Cell 6:167–174.

Hiratani I, Ryba T, Itoh M, Rathjen J, Kulik M, Papp B, Fussner E, Bazett-Jones DP, Plath K, Dalton S, Rathjen PD, Gilbert DM. 2009. Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. Genome Res 20:155–169.

Hochedlinger K, Jaenisch R. 2006. Nuclear reprogramming and pluripotency. Nature 441:1061–1067.

Hochedlinger K, Yamada Y, Beard C, Jaenisch R. 2005. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell 121:465–477.

Hockemeyer D, Soldner F, Cook EG, Gao Q, Mitalipova M, Jaenisch R. 2008. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell Stem Cell 3:346–353.

Hong H, Takahashi K, Ichisaka T, Aoi T, Kanagawa O, Nakagawa M, Okita K, Yamanaka S. 2009. Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature 460:1132–1135.

Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, Melton DA. 2008a. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol 26:795–797.

Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, Melton DA. 2008b. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 26:1269–1275.

Ichida JK, Blanchard J, Lam K, Son EY, Chung JE, Egli D, Loh KM, Carter AC, Di Giorgio FP, Koszka K, Huangfu D, Akutsu H, Liu DR, Rubin LL, Eggan K. 2009. A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing Nanog. Cell Stem Cell 5:491–503.

Ieda M, Fu JD, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, Srivastava D. 2010. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell 142:375–386.

Jackson-Grusby L, Beard C, Possemato R, Tudor M, Fambrough D, Csankovszki G, Dausman J, Lee P, Wilson C, Lander E, Jaenisch R. 2001. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet 27:31–39.

Jaenisch R, Young R. 2008. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 132:567–582.

Jia F, Wilson KD, Sun N, Gupta DM, Huang M, Li Z, Panetta NJ, Chen ZY, Robbins RC, Kay MA, Longaker MT, Wu JC. 2010. A nonviral minicircle vector for deriving human iPS cells. Nat Method 7:197–199.

Judson RL, Babiarz JE, Venere M, Blelloch R. 2009. Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol 27:459–461.

Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K. 2009. Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature 458:771–775.

Kang L, Wang J, Zhang Y, Kou Z, Gao S. 2009. iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. Cell Stem Cell 5:135–138.

Kawamura T, Suzuki J, Wang YV, Menendez S, Morera LB, Raya A, Wahl GM, Belmonte JC. 2009. Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature 460:1140–1144.

Kim J, Chu J, Shen X, Wang J, Orkin SH. 2008a. An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132:1049–1061.

Kim JB, Zaehres H, Wu G, Gentile L, Ko K, Sebastiano V, Arauzo-Bravo MJ, Ruau D, Han DW, Zenke M, Scholer HR. 2008b. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. Nature 454:646–650.

Kim D, Kim CH, Moon JI, Chung YG, Chang MY, Han BS, Ko S, Yang E, Cha KY, Lanza R, Kim KS. 2009a. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell Stem Cell 4:472–476.

Kim JB, Sebastiano V, Wu G, Arauzo-Bravo MJ, Sasse P, Gentile L, Ko K, Ruau D, Ehrich M, van den Boom D, Meyer J, Hubner K, Bernemann C, Ortmeier C, Zenke M, Fleischmann BK, Zaehres H, Scholer HR. 2009b. Oct4-induced pluripotency in adult neural stem cells. Cell 136:411–419.

Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LI, Yabuuchi A, Takeuchi A, Cunniff KC, Hongguang H, McKinney-Freeman S, Naveiras O, Yoon TJ, Irizarry RA, Jung N, Seita J, Hanna J, Murakami P, Jaenisch R, Weissleder R, Orkin SH, Weissman IL, Feinberg AP, Daley GQ. 2010. Epigenetic memory in induced pluripotent stem cells. Nature 467:285–290.

Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, Li E. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development (Cambridge, England) 122:3195–3205.

Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69:915–926.

Li H, Collado M, Villasante A, Strati K, Ortega S, Canamero M, Blasco MA, Serrano M. 2009a. The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature 460:1136–1139.

Li W, Wei W, Zhu S, Zhu J, Shi Y, Lin T, Hao E, Hayek A, Deng H, Ding S. 2009b. Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. Cell Stem Cell 4:16–19.

Li R, Liang J, Ni S, Zhou T, Qing X, Li H, He W, Chen J, Li F, Zhuang Q, Qin B, Xu J, Li W, Yang J, Gan Y, Qin D, Feng S, Song H, Yang D, Zhang B, Zeng L, Lai L, Esteban MA, Pei D. 2010. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell 7:51–63.

Liang J, Wan M, Zhang Y, Gu P, Xin H, Jung SY, Qin J, Wong J, Cooney AJ, Liu D, Songyang Z. 2008. Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat Cell Biol 10:731–739.

Liang G, Taranova O, Xia K, Zhang Y. 2010. Butyrate promotes induced pluripotent stem cell generation. J Biol Chem 285:25516–25521.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462:315–322.

Liu H, Zhu F, Yong J, Zhang P, Hou P, Li H, Jiang W, Cai J, Liu M, Cui K, Qu X, Xiang T, Lu D, Chi X, Gao G, Ji W, Ding M, Deng H. 2008. Generation of induced pluripotent stem cells from adult rhesus monkey fibroblasts. Cell Stem Cell 3:587–590.

Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong KY, Sung KW, Lee CW, Zhao XD, Chiu KP, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei CL, Ruan Y, Lim B, Ng HH. 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet 38:431–440.

Loh YH, Hartung O, Li H, Guo C, Sahalie JM, Manos PD, Urbach A, Heffner GC, Grskovic M, Vigneault F, Lensch MW, Park IH, Agarwal S, Church GM, Collins JJ, Irion S, Daley GQ. 2010. Reprogramming of T cells from human peripheral blood. Cell Stem Cell 7:15–19.

Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, Sridharan R, Clark AT, Plath K. 2008. Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci USA 105:2883–2888.

Lyssiotis CA, Foreman RK, Staerk J, Garcia M, Mathur D, Markoulaki S, Hanna J, Lairson LL, Charette BD, Bouchez LC, Bollong M, Kunick C, Brinker A, Cho CY, Schultz PG, Jaenisch R. 2009. Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4. Proc Natl Acad Sci USA 106:8912–8917.

Maherali N, Hochedlinger K. 2008. Guidelines and techniques for the generation of induced pluripotent stem cells. Cell Stem Cell 3:595–605.

Maherali N, Hochedlinger K. 2009. Tgfbeta signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. Curr Biol 19:1718–1723.

Maherali N, Sridharan R, Xie W, Utikal J, Eminli S, Arnold K, Stadtfeld M, Yachechko R, Tchieu J, Jaenisch R, Plath K, Hochedlinger K. 2007. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell 1:55–70.

Maherali N, Ahfeldt T, Rigamonti A, Utikal J, Cowan C, Hochedlinger K. 2008. A high-efficiency system for the generation and study of human induced pluripotent stem cells. Cell Stem Cell 3:340–345.

Marion RM, Strati K, Li H, Murga M, Blanco R, Ortega S, Fernandez-Capetillo O, Serrano M, Blasco MA. 2009. A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature 460:1149–1153.

Markoulaki S, Hanna J, Beard C, Carey BW, Cheng AW, Lengner CJ, Dausman JA, Fu D, Gao Q, Wu S, Cassady JP, Jaenisch R. 2009. Transgenic mice with defined combinations of drug-inducible reprogramming factors. Nat Biotechnol 27:169–171.

Marson A, Foreman R, Chevalier B, Bilodeau S, Kahn M, Young RA, Jaenisch R. 2008. Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell 3:132–135.

Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, Tachibana M, Lorincz MC, Shinkai Y. 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464:927–931.

McMahon SB, Wood MA, Cole MD. 2000. The essential cofactor TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc. Mol Cell Biol 20:556–562.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454:766–770.

Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, Bernstein BE, Jaenisch R, Lander ES, Meissner A. 2008. Dissecting direct reprogramming through integrative genomic analysis. Nature 454:49–55.

Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 26:101–106.

Nakagawa M, Takizawa N, Narita M, Ichisaka T, Yamanaka S. 2010. Promotion of direct reprogramming by transformation-deficient Myc. Proc Natl Acad Sci USA 107:14152–14157.

Nichols J, Smith A. 2009. Naive and primed pluripotent states. Cell Stem Cell 4:487–492.

Okita K, Ichisaka T, Yamanaka S. 2007. Generation of germline-competent induced pluripotent stem cells. Nature 448:313–317.

Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S. 2008. Generation of mouse induced pluripotent stem cells without viral vectors. Science (New York, NY) 322:949–953.

Park IH, Zhao R, West JA, Yabuuchi A, Huo H, Ince TA, Lerou PH, Lensch MW, Daley GQ. 2008. Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451:141–146.

Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, Tan KY, Apostolou E, Stadtfeld M, Li Y, Shioda T, Natesan S, Wagers AJ, Melnick A, Evans T, Hochedlinger K. 2010. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat Biotechnol 28:848–855.

Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, Maillard PV, Layard-Liesching H, Verp S, Marquis J, Spitz F, Constam DB, Trono D. 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. Nature 463:237–240.

Samavarchi-Tehrani P, Golipour A, David L, Sung HK, Beyer TA, Datti A, Woltjen K, Nagy A, Wrana JL. 2010. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell 7:64–77.

Seki T, Yuasa S, Oda M, Egashira T, Yae K, Kusumoto D, Nakata H, Tohyama S, Hashimoto H, Kodaira M, Okada Y, Seimiya H, Fusaki N, Hasegawa M, Fukuda K. 2010. Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. Cell Stem Cell 7:11–14.

Shi Y, Do JT, Desponts C, Hahm HS, Scholer HR, Ding S. 2008. A combined chemical and genetic approach for the generation of induced pluripotent stem cells. Cell Stem Cell 2:525–528.

Silva J, Chambers I, Pollard S, Smith A. 2006. Nanog promotes transfer of pluripotency after cell fusion. Nature 441:997–1001.

Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A. 2008. Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol 6:e253.

Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A. 2009. Nanog is the gateway to the pluripotent ground state. Cell 138:722–737.

Singhal N, Graumann J, Wu G, Arauzo-Bravo MJ, Han DW, Greber B, Gentile L, Mann M, Scholer HR. 2010. Chromatin-remodeling components of the BAF complex facilitate reprogramming. Cell 141:943–955.

Smith ZD, Nachman I, Regev A, Meissner A. 2010. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. Nat Biotechnol 28:521–526.

Soldner F, Hockemeyer D, Beard C, Gao Q, Bell GW, Cook EG, Hargus G, Blak A, Cooper O, Mitalipova M, Isacson O, Jaenisch R. 2009. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell 136:964–977.

Sommer CA, Stadtfeld M, Murphy GJ, Hochedlinger K, Kotton DN, Mostoslavsky G. 2009. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cell (Dayton, Ohio) 27:543–549.

Sommer CA, Sommer AG, Longmire TA, Christodoulou C, Thomas DD, Gostissa M, Alt FW, Murphy GJ, Kotton DN, Mostoslavsky G. 2010. Excision of reprogramming transgenes improves the differentiation potential of iPS cells generated with a single excisable vector. Stem Cell (Dayton, Ohio) 28:64–74.

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

Stadtfeld M, Brennand K, Hochedlinger K. 2008a. Reprogramming of pancreatic beta cells into induced pluripotent stem cells. Curr Biol 18:890–894.

Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. 2008b. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell 2:230–240.

Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K. 2008c. Induced pluripotent stem cells generated without viral integration. Science (New York, NY) 322:945–949.

Stadtfeld M, Maherali N, Borkent M, Hochedlinger K. 2009. A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nat Method 7:53–55.

Stadtfeld M, Apostolou E, Akutsu H, Fukuda A, Follett P, Natesan S, Kono T, Shioda T, Hochedlinger K. 2010. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. Nature 465:175–181.

Staerk J, Dawlaty MM, Gao Q, Maetzel D, Hanna J, Sommer CA, Mostoslavsky G, Jaenisch R. 2010. Reprogramming of human peripheral blood cells to induced pluripotent stem cells. Cell Stem Cell 7:20–24.

Takahashi K, 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, Yamanaka S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872.

Tchieu J, Kuoy E, Chin MH, Trinh H, Patterson M, Sherman SP, Aimiuwu O, Lindgren A, Hakimian S, Zack JA, Clark AT, Pyle AD, Lowry WE, Plath K. 2010. Female human iPSCs retain an inactive X chromosome. Cell Stem Cell 7:329–342.

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448:196–199.

Tsubooka N, Ichisaka T, Okita K, Takahashi K, Nakagawa M, Yamanaka S. 2009. Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. Genes Cell 14:683–694.

Utikal J, Maherali N, Kulalert W, Hochedlinger K. 2009a. Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. J Cell Sci 122:3502–3510.

Utikal J, Polo JM, Stadtfeld M, Maherali N, Kulalert W, Walsh RM, Khalil A, Rheinwald JG, Hochedlinger K. 2009b. Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature 460:1145–1148.

van den Berg DL, Zhang W, Yates A, Engelen E, Takacs K, Bezstarosti K, Demmers J, Chambers I, Poot RA. 2008. Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol Cell Biol 28:5986–5995.

Varas F, Stadtfeld M, de Andres-Aguayo L, Maherali N, di Tullio A, Pantano L, Notredame C, Hochedlinger K, Graf T. 2009. Fibroblast-derived induced pluripotent stem cells show no common retroviral vector insertions. Stem Cell (Dayton, Ohio) 27:300–306.

Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Sudhof TC, Wernig M. 2010. Direct conversion of fibroblasts to functional neurons by defined factors. Nature 463:1035–1041.

Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, Orkin SH. 2006. A protein interaction network for pluripotency of embryonic stem cells. Nature 444:364–368.

Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R. 2007. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448:318–324.

Wernig M, Lengner CJ, Hanna J, Lodato MA, Steine E, Foreman R, Staerk J, Markoulaki S, Jaenisch R. 2008a. A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nat Biotechnol 26:916–924.

Wernig M, Meissner A, Cassady JP, Jaenisch R. 2008b. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. Cell Stem Cell 2:10–12.

Winkler T, Cantilena A, Metais JY, Xu X, Nguyen AD, Borate B, Antosiewicz-Bourget JE, Wolfsberg TG, Thomson JA, Dunbar CE. 2010. No evidence for clonal selection due to lentiviral integration sites in human induced pluripotent stem cells. Stem Cell (Dayton, Ohio) 28:687–694.

Wolf D, Goff SP. 2009. Embryonic stem cells use ZFP809 to silence retroviral DNAs. Nature 458:1201–1204.

Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hamalainen R, Cowling R, Wang W, Liu P, Gertsenstein M, Kaji K, Sung HK, Nagy A. 2009. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 458:766–770.

Wu Z, Chen J, Ren J, Bao L, Liao J, Cui C, Rao L, Li H, Gu Y, Dai H, Zhu H, Teng X, Cheng L, Xiao L. 2009. Generation of pig induced pluripotent stem cells with a drug-inducible system. J Mol Cell Biol 1:46–54.

Xie H, Ye M, Feng R, Graf T. 2004. Stepwise reprogramming of B cells into macrophages. Cell 117:663–676.

Yang J, Chai L, Fowles TC, Alipio Z, Xu D, Fink LM, Ward DC, Ma Y. 2008. Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proc Natl Acad Sci USA 105:19756–19761.

Yang J, van Oosten AL, Theunissen TW, Guo G, Silva JC, Smith A. 2010. Stat3 activation is limiting for reprogramming to ground state pluripotency. Cell Stem Cell 7:319–328.

Yoshida Y, Takahashi K, Okita K, Ichisaka T, Yamanaka S. 2009. Hypoxia enhances the generation of induced pluripotent stem cells. Cell Stem Cell 5:237–241.

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA. 2007. Induced pluripotent stem cell lines derived from human somatic cells. Science (New York, NY) 318:1917–1920.

Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, Slukvin II, Thomson JA. 2009. Human induced pluripotent stem cells free of vector and transgene sequences. Science (New York, NY) 324:797–801.

Yusa K, Rad R, Takeda J, Bradley A. 2009. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. Nat Method 6:363–369.

Zhang Y, Rohde C, Tierling S, Jurkowski TP, Bock C, Santacruz D, Ragozin S, Reinhardt R, Groth M, Walter J, Jeltsch A. 2009. DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. PLoS Genet 5:e1000438.

Zhao XY, Li W, Lv Z, Liu L, Tong M, Hai T, Hao J, Guo CL, Ma QW, Wang L, Zeng F, Zhou Q. 2009. iPS cells produce viable mice through tetraploid complementation. Nature 461:86–90.

Zhou W, Freed CR. 2009. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. Stem Cell (Dayton, Ohio) 27:2667–2674.

Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. 2008. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. Nature 455:627–632.

Zhou H, Wu S, Joo JY, Zhu S, Han DW, Lin T, Trauger S, Bien G, Yao S, Zhu Y, Siuzdak G, Scholer HR, Duan L, Ding S. 2009. Generation of induced pluripotent stem cells using recombinant proteins. Cell Stem Cell 4:381–384.

CHAPTER 3

Stage-Specific Regulation of Reprogramming to Induced Pluripotent Stem Cells by Wnt Signaling and T Cell Factor Proteins

ABSTRACT

Wnt signaling is intrinsic to mouse embryonic stem cell self-renewal. Therefore, it is surprising that reprogramming of somatic cells to induced pluripotent stem cells (iPSCs) is not strongly enhanced by Wnt signaling. Here, we demonstrate that active Wnt signaling inhibits the early stage of reprogramming to iPSCs, whereas it is required and even stimulating during the late stage. Mechanistically, this biphasic effect of Wnt signaling is accompanied by a change in the requirement of all four of its transcriptional effectors: T cell factor 1 (Tcf1), Lef1, Tcf3, and Tcf4. For example, Tcf3 and Tcf4 are stimulatory early but inhibitory late in the reprogramming process. Accordingly, ectopic expression of Tcf3 early in reprogramming combined with its loss of function late enables efficient reprogramming in the absence of ectopic Sox2. Together, our data indicate that the stepwise process of reprogramming to iPSCs is critically dependent on the stage-specific control and action of all four Tcfs and Wnt signaling.

INTRODUCTION

The generation of induced pluripotent stem cells (iPSCs) from fibroblasts by ectopic expression of Oct4, Sox2, cMyc, and Klf4 established a major landmark in the field of stem cell biology, as it allowed the generation of patient-specific pluripotent cells (Maherali et al., 2007; Okita et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007). The reprogramming process is quite robust, in that ectopic expression of the reprogramming factors works on a wide range of differentiated cells to produce iPSCs (Stadtfeld and Hochedlinger, 2010). However, reprogramming to iPSCs is inefficient, in that only a few somatic cells of the starting population transition to pluripotency after a latency period of around 2 weeks (Papp and Plath, 2013). Thus, events that are currently largely unknown need to occur to achieve reprogramming to the pluripotent state. Indeed, the starting cell type, the reprogramming factor combination used, the method of overexpression, and the culture conditions all have major effects on the activation of the endogenous pluripotency gene regulatory network and even the epigenetic state of the

reprogrammed cells (Papp and Plath, 2013). In this study, we focus on the role of Wnt signaling in reprogramming to iPSCs.

The Wnt/β-catenin signaling pathway is intricately linked to the pluripotent state (Clevers and Nusse, 2012). For instance, mouse embryonic stem cells (ESCs) secrete active Wnt ligands and autocrine Wnt activity is required to prevent their differentiation (ten Berge et al., 2011), indicating that Wnt signaling is both necessary and sufficient for the self-renewal of these cells. Mouse ESCs can even self-renew efficiently in the absence of serum and extrinsic signals as long as Wnt/β-catenin signaling is stimulated and ERK kinases are inhibited ("2i" culture condition) (Ying et al., 2008).

Canonical Wnt signaling is classically described as functioning in two states. In the absence of a Wnt ligand, a complex of proteins, including Axin, Apc, Ck1, and Gsk3, stimulates the ubiquitin-mediated destruction of β-catenin (Clevers and Nusse, 2012). In the absence of stable β-catenin, T cell factor (Tcf) proteins (Tcf1, Lef1, Tcf3, and Tcf4 in mammals) transcriptionally repress Wnt target genes by interacting with corepressor proteins, such as Groucho or the C-terminal binding protein (Ctbp) and recruiting them to their DNA recognition sites through their high-mobility group (HMG) domain, which is nearly identical in all Tcfs (Clevers and Nusse, 2012). When a Wnt ligand activates the pathway, the β-catenin destruction complex is inhibited, enabling β-catenin to translocate to the nucleus, where it can bind to a conserved domain present near the amino terminus of all Tcfs (Clevers and Nusse, 2012). Upon binding to a Tcf, β-catenin can switch the activity of Tcfs from transcriptional repression to activation by recruiting coactivators, such as CBP (Takemaru and Moon, 2000). Although Tcfs share homologous HMG and β-catenin interaction domains, differences among individual Tcfs cause them to function uniquely within the Wnt pathway. For example, the effect of β-catenin binding can differ, either inducing the classic conversion from a repressor to transactivator for

Tcf1 and Lef1 or only inactivating the repressor activity of Tcf3 (B.J.M., unpublished data). Thus, individual Tcfs can cause overlapping or diverse effects, depending on how their conserved and unique elements are regulated.

An important understanding of how Wnt signaling affects ESCs has been attained through the appreciation of the diverse effects of Tcfs. Together with core pluripotency transcription factors, Oct4, Sox2, and Nanog, Tcf3 co-occupies many pluripotency genes, including *Nanog* and *Esrrb* (Cole et al., 2008; Marson et al., 2008a; Martello et al., 2012; Tam et al., 2008; Yi et al., 2008). Ablation of Tcf3 stimulates *Nanog* and *Esrrb* expression, similar to the activation of Wnt/β-catenin signaling (Cole et al., 2008; Martello et al., 2012; Pereira et al., 2006; Yi et al., 2011), and allows self-renewal of ESCs in serum-free conditions without Wnt pathway stimulation (Yi et al., 2011). It is therefore thought that Tcf3 acts exclusively as a transcriptional repressor in ESCs, even in the presence of stable β-catenin. Tcf4 mainly displays similar transcriptional repressor activity as Tcf3, but it is expressed at low levels in ESCs (Pereira et al., 2006; Yi et al., 2011). By contrast, Tcf1 and Lef1 display β-catenin-dependent transcriptional activator activity in ESCs, and endogenous Tcf1 activity counteracts some, but not all, transcriptional repression by Tcf3 (Yi et al., 2011).

The central importance of Wnt signaling and inhibition of Tcf3 for self-renewal of mouse ESC has stimulated investigations into the effects of Wnt signaling on reprogramming to pluripotency. In experiments where somatic cell nuclei are reprogrammed to the pluripotent state upon fusion with ESCs, treating ESCs with exogenous Wnt3a, stabilized β-catenin, or downregulation of Tcf3 each stimulates the efficiency by which somatic cells are reprogrammed (Han et al., 2010; Lluis et al., 2008, Lluis et al., 2011). The effects of Wnt3a or Tcf3 ablation on fusion-mediated reprogramming are substantial, increasing reprogramming efficiency up to 1,000-fold (Lluis et al., 2008, Lluis et al., 2011). By contrast, Wnt/β-catenin stimulation or Tcf3

depletion cause only a weak enhancement of reprogramming to iPSCs (Lluis et al., 2011; Marson et al., 2008b). In addition, β-catenin was among the original 24 factors screened by Takahashi and Yamanaka, and its overexpression was found to have no significant effect on iPSC formation (Takahashi and Yamanaka, 2006). Currently, it remains unclear how Wnt signaling has such a substantial impact on the self-renewal of mouse ESCs and reprogramming by fusion with ESCs, yet causes relatively minor effects on the outcome of iPSC-based reprogramming experiments.

To elucidate how Wnt/β-catenin signaling affects reprogramming to iPSCs, we determined the effects of inhibiting or stimulating Wnt signaling and the requirement for the four Tcfs during different stages of the reprogramming process. Our results demonstrate that early events in reprogramming are stimulated by inhibiting Wnt signaling, whereas late events are stimulated by activating the pathway. These effects are mediated by differential activities of the four Tcfs, and dynamic manipulation of Tcf3 levels allows for the efficient formation of iPSCs without exogenous Sox2. Our findings showcase that the poor efficiency of reprogramming is at least partially caused by changing molecular requirements in the process, where events promoting one phase are inhibitory for a subsequent phase, calling for further optimization of the iPSC technology.

RESULTS

Wnt Signaling Is Essential for Late Stages of Reprogramming to iPSCs

We began elucidating the role of Wnt signaling in reprogramming to iPSCs by determining whether endogenous Wnt signaling is necessary for the process. In these experiments, we employed mouse embryonic fibroblasts (MEFs) carrying a single tetracycline-inducible polycistronic cassette encoding *Oct4*, *Sox2*,*cMyc*, and *Klf4* (inducible OSCK [iOSCK]) and a reverse tetracycline transactivator (*M2rtTA*) transgene (Figure S3.1A); induced

reprogramming by addition of doxycycline; and assessed reprogramming efficiency under various treatments by quantifying colonies positive for expression of the pluripotency factor Nanog. To inhibit endogenous Wnt signaling, iOSCK MEFs were transduced with a retrovirus expressing Dickkopf1 (Dkk1), a secreted ligand and natural antagonist to the Wnt coreceptor LRP5/LRP6 (Mao et al., 2001). Reduction of transcript levels of the Wnt target gene *Axin2* and of TOPflash luciferase reporter activity confirmed that ectopic Dkk1 expression efficiently inhibited Wnt signaling (Biechele et al., 2009; Figures S3.1B and S3.1C). Notably, Dkk1 expression greatly reduced the numbers of Nanog-positive colonies (Figure 3.1A), suggesting that endogenous Wnt signaling is essential for the formation of iPSCs. To confirm that the reduction of Nanog-positive colonies was due to effects on Wnt signaling, IWP2, a potent small-molecule inhibitor of Porcupine, which is necessary for the processing and secretion of all Wnt ligands (Chen et al., 2009), was continuously added throughout the reprogramming process. This independent method of inhibiting endogenous Wnt signaling strongly blocked the formation of Nanog-positive colonies (Figure 3.1B), demonstrating that the production of active Wnt ligands is essential for reprogramming.

The substantial negative effects of inhibiting endogenous Wnt signaling (Figures 3.1A and 3.1B) contrasted with the minimal positive effects of exogenously stimulating Wnt reported previously, and there was even no effect when cMyc was included in the reprogramming cocktail with Oct4, Sox2, and Klf4 (Marson et al., 2008b). To determine if an effect of exogenous Wnt signaling could be observed with our experimental set-up, Nanog-expressing colonies were measured in reprogramming experiments continuously treated with purified recombinant Wnt3a protein, which strongly activated a TOPflash luciferase reporter (Figure S3.1C). Consistent with previous reports, we observed no increase in the reprogramming of iOSCK MEFs to Nanog-expressing colonies upon Wnt3a treatment (Figure 3.1C). Moreover, recombinant Wnt3a reduced the number of Nanog-expressing colonies by half, indicating that constitutive

exogenous Wnt signaling is even inhibitory for the induction of Nanog when reprogramming is performed with our inducible OSCK polycistronic expression cassette.

Because both stimulating and blocking Wnt signaling resulted in an inhibition of iOSCK reprogramming, albeit to different extents, we examined the possibility of different responses to Wnt signaling during the progression to pluripotency. To test this, we divided the reprogramming process into early, mid, and late stages and observed the effects of Wnt inhibition and stimulation on reprogramming (Figure 3.1D). Early treatment with IWP2 yielded a 2-fold increase of Nanog-positive colonies, whereas late treatment strongly reduced reprogramming (Figure 3.1E). Conversely, Wnt3a addition early resulted in a dramatic inhibition of the formation of Nanog-positive colonies, whereas treatment late increased the number of Nanog-positive colonies over 2-fold (Figure 3.1F). The efficiency of reprogramming was not affected when either IWP2 or Wnt3a were added in the middle phase of the process (Figures 3.1E and 3.1F). Together, these data demonstrate a biphasic response of the iPSC reprogramming process to Wnt signaling.

The possibility that this biphasic response could be caused by different responses to graded levels of Wnt activity, as previously described by Cosma and colleagues for cell fusion experiments (Lluis et al., 2008), was ruled out by the dose-dependent manner by which IWP2 affected Nanog-positive colony formation (Figures S3.1D–F). Moreover, effects on levels of well-established Wnt targets, such as *Axin2* and *Tcf1*, confirmed that Wnt3a and IWP2 stimulated and inhibited Wnt signaling, respectively, as expected (Figures S3.1E and S3.1F).

We conclude that our experimental approach resolves reprogramming effects that were previously overlooked by continuous treatment of reprogramming cultures with Wnt pathway effectors. Two distinct phases of Wnt response affect two stages of reprogramming, which can

be temporally defined as early and late stages. The conclusion that the stimulation of Wnt signaling establishes a strong barrier for early reprogramming events is further supported by the findings that preincubation of MEFs with Wnt3a before the start of reprogramming dramatically impaired reprogramming and preincubation with IWP2 enhanced reprogramming (Figure S3.1G). In contrast, the late phase not only depends on endogenously active Wnt signaling but also benefits from the ectopic stimulation of the pathway.

Tcf1/Lef1 and Tcf3/Tcf4 Have Opposing Roles in the Biphasic Response to Wnt Signaling during Reprogramming

Wnt ligands can elicit multiple downstream effects, some of which are independent of the canonical Tcf-β-catenin regulation of target genes. To determine whether the biphasic effects of Wnt signaling on reprogramming to iPSCs were mediated by Tcfs, we depleted each Tcf (*Tcf1, Lef1, Tcf3,* and *Tcf4*) during reprogramming by small interfering RNA (siRNA)-mediated knockdown. We also tested the effect of depleting all possible pairs of Tcfs to address the potential for redundancy between family members. The requirement for the Tcfs was first examined during the early stage of reprogramming (Figures 3.2A and S3.2A–D). The strongest effects due to loss-of-function of a single Tcf early in the process were seen for *Lef1* and *Tcf4*, respectively (Figure 3.2B). Their knockdown had opposing effects on the induction Nanog-positive colonies; depletion of *Lef1* increased and *Tcf4* knockdown decreased colony numbers. These effects were magnified when the knockdown of *Lef1* was combined with that of *Tcf1* or when *Tcf4* was depleted together with *Tcf3* (Figure 3.2B). These findings reveal (1) redundancies among Tcf family members and (2) an antagonistic effect between two groups of Tcfs early in reprogramming: endogenous Tcf1 and Lef1 are inhibitors, whereas Tcf3 and Tcf4 are enhancers in this phase of reprogramming.

Based on the effects of Wnt signaling on the early reprogramming phase, one would predict that Tcf1 and Lef1 mediate Wnt effects, whereas Tcf3 and Tcf4 counteract Wnt effects during this stage. Consistent with this hypothesis, the transcript levels of Wnt target genes *Tcf1, Lef1,* and *Axin2* were significantly increased upon *Tcf3* and *Tcf4* knockdown early in reprogramming (Figure 3.2C). To test this hypothesis further, we combined IWP2 treatment with knockdown of *Tcf1/Lef1* or *Tcf3/Tcf4* early in reprogramming (Figures 3.2D, S3.2E, and S3.2F). The combined knockdown of *Tcf1/Lef1* increased the number of Nanog-positive colonies in the absence of IWP2 but failed to further increase colony numbers when endogenous Wnt signaling was blocked by IWP2 (Figure 3.2D). Thus, the effects of Tcf1 and Lef1 early in reprogramming overlap with those seen by inhibiting endogenous Wnt signaling, which is consistent with Wnt/β-catenin-dependent transcriptional activator activities for Tcf1 and Lef1 during this reprogramming phase. Conversely, knockdown of *Tcf3/Tcf4* inhibited reprogramming, regardless of the presence or absence of IWP2 (Figure 3.2D). The Wnt target genes *Tcf1*, *Lef1*, and *Axin2* were upregulated upon *Tcf3/Tcf4*knockdown, even in the presence of IWP2 (i.e., without active Wnt signaling) (Figure S3.2G). Therefore, the inhibitory effect of *Tcf3/Tcf4* depletion on early reprogramming does not require active Wnt signaling, which is most consistent with transcriptional repressor activities for Tcf3 and Tcf4. Notably, the simultaneous knockdown of all four Tcfs reduced the number of Nanog-positive colonies compared to control (Figures 3.2E, S3.2H, and S3.2I). This result indicates that the mediators of active Wnt signaling, *Tcf1* and *Lef1,* are not the critical targets of Tcf3 and Tcf4 repression during the early stage of reprogramming, as reducing the aberrant *Tcf1/Lef1* upregulation observed upon *Tcf3/Tcf4* depletion did not rescue the reprogramming efficiency.

To test the role of Tcfs during the late phase of reprogramming, we transfected siRNAs once on day 6 of reprogramming and assessed the formation of Nanog-positive colonies 3 days later (Figures 3.3A and S3.3A–D). Among all siRNA treatments, only the knockdown of *Tcf3* and

Tcf4, individually or together, consistently enhanced reprogramming (Figures 3.3B and S3.3E), indicating that Tcf3 and Tcf4 inhibit late reprogramming events. Because the effect of the *Tcf3*and *Tcf4* double knockdown was not additive compared to their respective single knockdowns, Tcf3 and Tcf4 likely act in the same pathway. Importantly, the reprogramming enhancement due to *Tcf3* or *Tcf4* knockdown was nullified when *Tcf1* or *Lef1* were concurrently depleted (Figures 3.3B and S3.3F). Because these data demonstrated that the loss of Tcf3 or Tcf4 requires Tcf1 of Lef1 for a positive effect late in reprogramming, which are typically the mediators of active Wnt signaling, we next tested the requirement of Wnt signaling in this context more directly by combining the *Tcf3/Tcf4* knockdown with IWP2 treatment. Our results show that IWP2 prevented the enhancing effect of *Tcf3* or *Tcf4* depletion late in reprogramming (Figure 3.3C, green bars). This effect does not appear to be due to a dramatic change in *Tcf1* and *Lef1* levels (Figure S3.3G). Together, these findings indicate that depletion of *Tcf3* and *Tcf4* promotes the late stage of reprogramming through a mechanism that requires Tcf1 or Lef1 as mediators of active Wnt signaling.

Although depletion of *Tcf1* and/or *Lef1* during the late phase did not inhibit reprogramming (Figure 3.3B), their depletion mitigated the inhibitory effect of Wnt inhibition (i.e., IWP2 treatment) during the late phase of reprogramming, even when combined with depletion of *Tcf3* and *Tcf4* (Figure 3.3C, orange bars). These data are most consistent with the interpretation that the activity of endogenous Wnt signaling during the late phase is necessary to prevent Tcf1 and Lef1 from becoming potent inhibitors of reprogramming. We speculate that, in the absence of active Wnt signaling late in reprogramming, Tcf1/Lef1 are transcriptional repressors at target genes that are essential for the induction of pluripotency. Depleting *Tcf1/Lef1* under “no Wnt” conditions would relieve the repressive effect and allow other, alternative pathways to activate these critical genes. Such alternative pathways may also explain why depletion of *Tcf1* and*Lef1* alone did not inhibit reprogramming as seen in Figure

3.3B. Although we favor this explanation, it is also possible that residual activity of Tcf1 or Lef1 after siRNA knockdown is enough to fulfill a critical function, which could be addressed in the future by using genetic knockout models.

Taken together, our data uncover different requirements of the Tcfs early and late in reprogramming, which is consistent with the changing role of Wnt signaling between the early and late stages. Early in reprogramming, Tcf3 and Tcf4 stimulate reprogramming and are inhibited by active Wnt signaling mediated by Tcf1 and Lef1. Late in reprogramming, Tcf3 and Tcf4 are inhibitory and regulate the activity of the Wnt signaling pathway. We posit that the distinct activities of individual Tcf factors are responsible for the biphasic effects of Wnt signaling on iPSC reprogramming.

Biphasic Effects of Tcf3 Affect the Requirement for Exogenous Reprogramming Factors

Given the fundamental role of Tcf3 in regulating pluripotency in ESCs, we reasoned that further elucidating how Tcf3 contributes to the biphasic Wnt signaling effect during reprogramming to iPSCs would provide the greatest insights into mechanisms of the process. First, we determined if overexpression of Tcf3 would affect the dynamics of reprogramming (Figure 3.4A). Constitutive Tcf3 expression throughout reprogramming reduced the number of cells positive for the surface marker SSEA1, which marks late reprogramming intermediates (Stadtfeld et al., 2008), and also decreased the formation of Nanog and *Oct4*-GFP-positive colonies in a dose-dependent manner (Figures 3.4B, 3.4C, S3.4A, and S3.4B). Proliferation of the reprogramming culture was not affected by Tcf3 expression (data not shown), and quantitative PCR (qPCR) confirmed the reduction of *Nanog* and *Esrrb* transcripts in Tcf3-expressing reprogramming cultures (Figures S3.4C–E), confirming that Tcf3 overexpression is incompatible with late stages of iPSC formation. However, the expression of an early marker of reprogramming, E-cadherin (Cdh1), which marks the mesenchymal-to-epithelial transition

(Samavarchi-Tehrani et al., 2010), was increased when Tcf3 was overexpressed (Figure 3.4D). Similarly, Tcf3 overexpression resulted in a dramatic increase in alkaline phosphatase (AP)-positive colonies, which normally arise at a midpoint of reprogramming (Figure 3.4E). Together, these data suggest that Tcf3 overexpression stimulates early reprogramming events and colony formation but inhibits later events, including pluripotency gene induction. These data are in agreement with the observation that depletion of endogenous *Tcf3/Tcf4* early in reprogramming is inhibitory (Figure 3.2), whereas depletion late promotes reprogramming (Figure 3.3).

Tcf3 has been described to function in mice exclusively as a transcriptional repressor, whereas the other Tcfs have been shown to be able to switch between repressor and activator states (Wu et al., 2012; B.J.M., unpublished data). The effects of Tcf3 overexpression on reprogramming enabled mutational analysis of the domains of Tcf3 required for stimulation of AP-positive colony formation using previously characterized mutants. Tcf3 mutants that lack the domain responsible for the interaction with β-catenin (ΔN) or Ctbp (ΔC) repress Tcf-β-catenin target genes similarly to wild-type Tcf3. Those that lack the groucho-interaction region (ΔG) or carry point mutations in the HMG DNA-binding domain (ΔH) do not repress Tcf-β-catenin target genes (Merrill et al., 2001; Figure 3.4Fi). During reprogramming, all Tcf3 mutants were expressed at similar levels and localized to the nucleus, ruling out the possibility that differences between mutants could be due to lack of expression or different subcellular localization (Figures S3.4F and S3.4G). Similar to wild-type Tcf3, expression of ΔC and ΔN mutants increased the number of AP-positive colonies (Figure 3.4Fii). By contrast, the ΔG and ΔH mutants that lacked repressor activity also lacked the ability to stimulate AP colony formation (Figure 3.4Fii). Therefore, direct binding of Tcf3 to DNA and Tcf3's repressor activity are important for stimulating the early phase of reprogramming.

To determine whether the effects of endogenous Tcf3 were modulated by the reprogramming factors, iPSC reprogramming was examined using all possible combinations of reprogramming factors. For these experiments, we established the genetic ablation of *Tcf3* during reprogramming by employing MEFs homozygous for a conditional *Tcf3* allele (Merrill et al., 2004) that also carry an estrogen receptor-tagged *Cre* recombinase transgene. These MEFs were initially transduced with separate retroviruses to express the reprogramming factors Oct4, Sox2, cMyc, and Klf4, and after splitting, half of the reprogramming culture was treated with tamoxifen (Tam) to induce *Tcf3* ablation. Upon 24 hr of exposure to Tam, excision of the loxp-flanked cassette (Figure 3.5A) and elimination of Tcf3 protein occurred efficiently (Figure 3.5B). Deletion of *Tcf3* increased the number of Nanog-positive colonies consistently but less than 2-fold without enhancing the kinetics of the process (Figure 3.5Ci). A similar effect due to *Tcf3* deletion was also observed when cMyc was omitted from the reprogramming factor cocktail (Figures 3.5Cii, 3.5D, and S3.5A). The enhancement of OSCK and OSK reprogramming by *Tcf3* loss was observed in media containing fetal bovine serum or knockout serum replacement, which is known to enhance reprogramming (Esteban et al., 2010; Figure S3.5B) and was not simply a consequence of an increased proliferation rate (Figures S3.5C and S3.5D). Deletion of *Tcf3* at the very beginning of the reprogramming process reduced the enhancing effect and yielded only a few more Nanog-positive colonies than control (Figure S3.5E), indicating that constitutive ablation of *Tcf3* throughout the entire reprogramming process causes only a minor increase in the overall efficiency. These data are consistent with our findings that the timing of Tcf3 activity is critical, due to the biphasic nature of Wnt effects on iPSC reprogramming.

Of all the possible combinations of reprogramming factors, ablation of *Tcf3* caused the strongest effect on OCK reprogramming. Previous studies have reported that reprogramming in the absence of ectopic Sox2 results in the generation of partially reprogrammed ESC-like colonies, in which the pluripotency network is not activated (Takahashi and Yamanaka, 2006).

Initially, we found that a very small number of these ESC-like colonies obtained upon OCK-induced reprogramming expressed Nanog in the complete absence of *Tcf3* (Figures 3.5D and S3.5F), indicating that constitutive *Tcf3* deletion enabled OCK reprogramming but at an extremely low rate and with dramatically delayed kinetics compared to OSK or OSCK reprogramming. However, passaging-dependent mechanisms magnified this effect. Specifically, we observed that ESC-like colonies isolated and expanded from a *Tcf3−/−* OCK reprogramming culture at day 30 induced Nanog expression with high efficiency, whereas Nanog remained largely undetectable when clones from a parallel *Tcf3+/+* OCK reprogramming culture were expanded (Figure S3.5G). Similarly, splitting *Tcf3−/−* OCK reprogramming cultures resulted in the induction of Nanog expression in many colonies (Figures 3.6C and 3.6D).

These Nanog-positive *Tcf3*−/− OCK-reprogrammed cell lines displayed silencing of retroviral reprogramming factor expression and lacked *Tcf3* and retroviral*Sox2* integration (Figures S3.5H and S3.5I). Hierarchical clustering and Pearson correlation of genome-wide gene expression data showed that OSK and OCK*Tcf3*−/− iPSC lines were similar to wild-type ESCs and iPSCs and clearly different from MEFs and a line of partially reprogrammed OCK pre-iPSCs (Figures 3.5E and S3.5J). *Tcf3*−/− reprogrammed lines also produced teratomas with three embryonic germ layers (Figure S3.5K) and upregulated markers of each germ layer during embryoid body differentiation, albeit with delayed kinetics relative to wild-type iPSCs (Figure 3.5F), which is a characteristic of *Tcf3*−/− ESCs compared to wild-type ESCs (Yi et al., 2008). Furthermore, *Tcf3*−/− iPSC lines bear similar expression differences as *Tcf3*−/− ESCs when compared to their wild-type counterparts (Figure S3.5L), further indicating that they closely resemble *Tcf3−/−* ESCs. Together, these data demonstrate that reprogramming in the absence of Tcf3 and ectopic Sox2 yields bona fide iPSCs.

Given that *Tcf3* deletion is advantageous for the late stage of OSCK reprogramming and enabled completion of OCK reprogramming, we tested whether partially reprogrammed colonies that normally are the end-product of OCK reprogramming (OCK pre-iPSCs), characterized by an ESC-like morphology and lack of pluripotency network expression (Sridharan et al., 2009; Takahashi and Yamanaka, 2006), are blocked from reaching pluripotency by Tcf repressor activity. Notably, knockdown of *Tcf3* and/or *Tcf4* yielded a large number of *Nanog*-GFP-positive colonies as early as 72 hr post-siRNA transduction, whereas *Tcf1* knockdown did not induce *Nanog*-GFP expression (Figures 3.6A, 3.6B, and S3.6A–C). *Tcf3* and *Tcf4* knockdown in OCK pre-iPSCs induced the Wnt signaling target genes *Lef1*, *Tcf1*, and *Axin2* (Figure S3.6D), and the concurrent knockdown of *Tcf1* dramatically inhibited the appearance of *Nanog*-GFP-positive colonies without affecting overall colony morphology or cell number (Figures 3.6B, S3.6E, and S3.6F). *Lef1* siRNA knockdown did not affect the OCK pre-iPSC to iPSC transition (data not shown). Together, these data indicate that *Tcf3* and *Tcf4* knockdown can rapidly trigger induction of pluripotency in OCK pre-iPSCs. Furthermore, the transition from OCK pre-iPSCs to pluripotency appears to require a similar mechanism as the late stage of OSCK reprogramming (i.e., a Tcf1/Lef1-dependent pathway, likely requiring active Wnt signaling).

Expression Changes due to Tcf3 Ablation Differ Early and Late in Reprogramming

To determine downstream genes mediating the effects of Tcf3, we analyzed the gene expression changes in OCK reprogramming cultures in the presence and absence of *Tcf3*. Parallel *Tcf3*$^{+/+}$ and *Tcf3*$^{-/-}$ OCK reprogramming cultures were split at day 21 to enhance the Tcf3-mediated reprogramming effect, and RNA samples were collected at several time points throughout the reprogramming experiment (Figures 3.6C and 3.6D). qPCR confirmed the decrease of *Tcf3* messenger RNA (mRNA) levels upon activation of *Cre* and the increase in *Nanog* transcript levels in the *Tcf3*$^{-/-}$ OCK reprogramming culture at late time points (Figure S3.6G). None of the tested endogenous Sox family members were precociously upregulated in

the absence of *Tcf3* (Figure S3.6G), thereby discounting a simple mechanism by which *Tcf3* ablation could enable the induction of pluripotency in the absence of ectopic Sox2 (Nakagawa et al., 2008).

We next combined our genome-wide gene expression data with unsupervised short time-series expression miner (STEM) analysis (Ernst and Bar-Joseph, 2006) to capture expression differences and groups of coregulated genes between the *Tcf3*+/+ and *Tcf3*−/− OCK reprogramming cultures. The three most significant groups of coregulated genes are depicted in Figure 3.6E. Group 1 genes are more highly expressed in ESCs than MEFs, initially (at day 15) expressed at lower levels in the *Tcf3*−/− reprogramming culture compared to the *Tcf3*+/+ culture but slightly surpassed the levels of the *Tcf3*+/+ culture by day 22. Based on gene ontology (GO) analysis, these genes function in the regulation of cell proliferation (Figure 3.6E). Group 2 genes are strongly induced in the *Tcf3*+/+ reprogramming culture but not in the *Tcf3*−/− culture at day 26 and are implicated in morphogenesis and neuronal differentiation. Group 3 genes are more highly expressed in the *Tcf3*−/− reprogramming culture at day 26 and include several pluripotency-related genes, such as *Zfp42*, *Dppa3, Esrrb,* and *Tcfcp2l1*, consistent with the induction of faithful reprogramming specifically in the absence of *Tcf3*. These data indicate that OCK-transduced MEFs progress faster into an intermediate reprogramming state in the presence of *Tcf3* but then upregulate various lineage regulators later in the reprogramming process. Given that the expression of developmental genes has been suggested to be a barrier to reprogramming (Mikkelsen et al., 2008), these genes could block the entry into pluripotency. In the absence of *Tcf3*, the upregulation of a large number of developmental genes appears to be efficiently suppressed, which could overcome the pluripotency blockade.

To confirm that the suppression of developmental genes late in reprogramming is not simply a consequence of expression changes that occurred earlier in the process due to

continuous *Tcf3* deletion, we determined direct expression changes upon *Tcf3* depletion in a late reprogramming stage. Depletion of *Tcf3*in OCK-pre-iPSCs led to the downregulation of a similar group of developmental genes as defined by group 2 (Figure S3.6H). Interestingly, active Wnt signaling is known as a negative regulator of neural genes (Aubert et al., 2002; Yoshikawa et al., 1997). Because Wnt target genes, such as *Tcf1*, *Lef1*, *Cdx1*, and *Brachyury*, were upregulated both in late *Tcf3*−/− reprogramming cultures and *Tcf3*-depleted OCK-pre-iPSCs (Figures S3.6D and S3.6I) and active Wnt signaling is required for the enhancing effects of *Tcf3* deletion late in reprogramming (Figures 3.3 and 3.6B), the induction of Wnt signaling upon *Tcf3* deletion may therefore be directly responsible for the suppression of neural genes late in reprogramming. Taken together, these data demonstrate that Tcf3 has different targets in the early and late stages of the process, which is consistent with the biphasic role of Wnt signaling during reprogramming.

Stage-Specific Modulation of Tcf3 Levels Enables Efficient OCK Reprogramming

The biphasic response to Wnt signaling and stage-specific effects of Tcfs indicate that, to arrive at the pluripotent state, individual cells progress through a Wnt “low” (or Tcf3 high) state followed by progression through a Wnt “high” (or Tcf3 low) activity state. To test this idea directly, we established a system that allowed us to manipulate *Tcf3* levels in a stage-dependent manner, where each cell expressed high *Tcf3* levels at the early stage and reduced *Tcf3* levels at the late stage of reprogramming (Figure 3.7A). Based on our data, we reasoned that elevated *Tcf3* should promote early reprogramming events and subsequent depletion of *Tcf3* would then promote late events. This hypothesis was tested in the context of OCK reprogramming, the best system to observe reprogramming enhancement in a *Tcf3*-dependent manner. We expressed *Tcf3* at a range of levels early in OCK reprogramming, from day 1 to day 8, taking advantage of a doxycycline-inducible expression system (Figure 3.7A). At 0 and 0.002 μg/ml of dox, representing MEF and ESC-like mRNA levels of *Tcf3*, respectively (Figure

3.7B), OCK reprogramming cultures appeared similar at day 8 of reprogramming, displaying nascent colonies (Figure S3.7A). At much higher *Tcf3* levels induced by 0.02 µg/ml dox (Figure 3.7B), more and bigger colonies emerged (Figure S3.7A). On day 8, dox was withdrawn to stop *Tcf3* overexpression and siRNAs targeting *Tcf3* were added to ensure the reduction of *Tcf3* in the late phase (Figure 3.7A). Reprogramming cultures were monitored daily for *Oct4*-GFP-positive colonies, prompting the following conclusions (Figure 3.7C): (1) *Tcf3* overexpression early in OCK reprogramming is not sufficient for the induction of reprogrammed cells. (2) *Tcf3* knockdown late, without prior overexpression of *Tcf3,* only yielded rare *Oct4*-GFP-positive colonies similar to our findings described in Figures 3.5 and S3.5. (3) Induction of ESC-like transcript levels of *Tcf3* early (0.002 µg/ml dox) followed by *Tcf3* knockdown late resulted in large numbers of *Oct4*-GFP-positive colonies. (4) Very high levels of *Tcf3* (0.02 µg/ml dox) early in reprogramming eventually gave rise to some *Oct4*-GFP-positive colonies when combined with *Tcf3* knockdown late, albeit with lower efficiency, even though this condition resulted in the most promising induction of ESC-like colonies at day 8, indicating that the exact levels of *Tcf3* early in reprogramming are critical.

Three *Oct4*-GFP-positive OCK colonies, treated with 0.002 µg/ml dox, and subsequent *Tcf3* siRNA knockdown were stably expanded from this experiment and confirmed to lack the *Sox2* reprogramming vector (Figure S3.7B). These cell lines exhibited typical characteristics of pluripotent stem cells; in addition to their ESC-like morphology, they have silenced the retroviral expression of the reprogramming factors (Figure S3.7C), expressed the endogenous pluripotency genes *Sox2* and *Nanog,* and displayed ESC-like *Tcf3* transcript levels (Figures S3.7D and S3.7E). Upon blastocyst injection of two clones, we received pups with contribution of iPSCs to various tissues, as tested by PCR for the retroviral *Tcf3* transgene (Figures 3.7D and S3.7F). Taken together, this experiment provides the proof of principle that, during reprogramming, cells transition through stages in which the activity of the Wnt/Tcf machinery

dramatically differs and where precise levels of *Tcf3* are critical for achieving successful reprogramming.

DISCUSSION

Somatic cells en route to the pluripotent state undergo specific events, starting with the loss of somatic cell identity and culminating in the expression of the full pluripotency network (Papp and Plath, 2013). In this study, we performed a comprehensive analysis of the role of Wnt signaling and the requirement of its transcriptional effectors Tcf1, Lef1, Tcf3, and Tcf4 in this process. Our work shows that reprogramming of mouse fibroblasts is biphasic with respect to its dependence on endogenous Wnt signaling, Tcf proteins, and the consequences of ectopic Wnt stimulation (summarized in Figure 3.7E).

Two phases of Wnt signaling could be temporally separated into early and late stages of reprogramming, which enabled us to study the molecular roles for Wnt and Tcfs during each phase. In the early stage, the activation of Wnt signaling leads to a reprogramming block via Tcf1 and Lef1, likely due to induction of Wnt target genes that interfere with reprogramming events. In contrast, Tcf3 and Tcf4 promote early reprogramming events by repressing Wnt pathway target genes, including *Tcf1* and *Lef1*, and likely other targets not stimulated by Tcf1/Lef1 and active Wnt signaling. The targets of Tcf3/Tcf4 repression interfere with efficient reprogramming when expressed during the early stage. In the late stage, Wnt signaling promotes reprogramming. Interestingly, Tcf1/Lef1 and Tcf3/Tcf4 have opposing roles, as they did in the early stage; however, the relationship between Tcf1/Lef1 and Tcf3/Tcf4 is different compared to the early stage. Our data suggest that Tcf3 and Tcf4 repress the expression of Tcf1 and Lef1 late in reprogramming, thereby limiting the activity of Wnt signaling. Accordingly, deletion of *Tcf3/Tcf4* late in reprogramming enhances iPSC formation through a mechanism that requires Tcf1 or Lef1 and active Wnt signaling. Thus, Tcf1 and Lef1 appear to be critical target

genes of Tcf3 and Tcf4 late in reprogramming. We propose that Wnt3a addition stimulates the late stage of reprogramming primarily by making Tcf1/Lef1 strong activators of key target genes and preventing Tcf1/Lef1 from acting as transcriptional repressors. Although we do not exclude a direct effect of Wnt3a on Tcf3 or Tcf4 activity or levels, our results suggest that Wnt stimulation acts upstream of Tcf1/Lef1 to enhance the late reprogramming phase. The late stage of reprogramming is likely unaffected by *Lef1* or *Tcf1* depletion, because alternative pathways are active that can act on a similar set of target genes. One such pathway may be the leukemia inhibitory factor (Lif)/Jak/Stat signaling pathway. Notably, in the presence of Lif, there is no consequence on ESC self-renewal upon *Tcf1* depletion (Yi et al., 2011). However, the ability of Wnt3a to sustain ESC self-renewal upon Lif withdrawal is stimulated by Tcf1 (Yi et al., 2011), indicating a redundancy between distinct signaling pathways in maintaining the pluripotent state, which may extend to a redundancy in acquiring pluripotency.

Throughout reprogramming, we suggest that the grouping of Tcf1/Lef1 versus Tcf3/Tcf4 reflects predominant Wnt-dependent activator functions of Tcf1/Lef1 and repressor functions of Tcf3/Tcf4. The observation that the four Tcfs fall into two distinct groups for their effect on reprogramming to iPSCs provides further insight into the roles of the factors as mediators of Wnt signaling. The grouping of the factors supports the diversification of the Tcf family into isoforms with specialized and distinct activities (Cadigan and Waterman, 2012). This contrasts the switch model pertaining to invertebrates, where a single Tcf gene product performs both activation and repression. The activator effect attributed to Tcf1/Lef1 during reprogramming is consistent with analysis of *Lef1−/−;Tcf1−/−* double mutant mice, which display a *Wnt3a−/−* like phenotype (Galceran et al., 1999). The repressor activity of Tcf3/Tcf4 is consistent with the β-catenin independent effects caused by conditional *Tcf3* ablation in the skin of *Tcf4−/−* mice (Nguyen et al., 2009).

We made the striking discovery that solely manipulating the levels of Tcf3, from slight overexpression early to depletion late in the process, allows efficient and faithful reprogramming in the absence of ectopic Sox2. On a molecular level, this finding highlights a function of Sox2 that can be complemented by regulators of the Wnt pathway. The recently described competition between Tcf3 and Sox2 for binding at Oct-Sox DNA sites provides a possible mechanistic explanation for the effects of *Tcf3* ablation during late reprogramming (Zhang et al., 2013). Notably, our data highlight that the degree to which Wnt signaling activation and inhibition affect the early and late stages of reprogramming is dependent on the reprogramming factor combination used.

The duality of effects of Wnt during reprogramming provides a strong example of a factor being necessary at one step but being a barrier at a different step of the long reprogramming process. A priori, it is likely that many factors could cause similar biphasic or context-specific effects during reprogramming. Reprogramming methods that account for dynamic changes in signaling requirements, perhaps in other pathways, will more efficiently guide somatic cells into the desired pluripotent state. Moving forward, determining the reprogramming stage-specific target genes of Tcf1/Lef1 and Tcf3/Tcf4 under Wnt “on” and “off” conditions, along with different reprogramming factor combinations, will be a key question to answer to further understand the biphasic action of Wnt signaling in reprogramming to iPSCs.

EXPERIMENTAL PROCEDURES

Expression Constructs and Cell Lines

For reprogramming with retroviral factors, Oct4, Sox2, Klf4, and cMyc were expressed from pMX retroviruses, as previously described (Maherali et al., 2007). For overexpression, the complementary DNAs (cDNAs) encoding full-length *Tcf3* or its domain mutants (Merrill et al., 2001), *Dkk1*, or *Tomato* fluorescent protein (used as control) were also cloned into the pMX

vector. For inducible *Tcf3* overexpression experiments, the *Tcf3* cDNA was cloned into the pRetroX-Tight-Hyg vector, allowing doxycycline-inducible expression in MEFs carrying the *M2rtTA* transgene in the *Rosa26* (*R26*) locus. For reprogramming experiments utilizing tet-iOSCK reprogramming factors, MEFs harboring the *R26-M2rtTA* and a single, doxycycline-inducible, polycistronic cassette coding for OSCK (Sommer et al., 2009) in the *Col1A* locus were generated from mice similarly to a published report (Stadtfeld et al., 2010). Some of the MEFs used for reprogramming carried the *Oct4-GFP* transgene (Szabó et al., 2002) or the *GFP* knockin in the endogenous *Nanog* locus (Maherali et al., 2007), as indicated.

Reprogramming Experiments

For reprogramming, MEFs cultured in mouse ESC media containing 15% Fetal bovine serum (FBS) were infected with respective retroviruses for 12-18 hr and split three days post-transduction onto coverslips that were pre-seeded with irradiated mouse feeders. Typically, mouse ESC media containing Leukemia inhibitory factor (Lif) and 15% Knockout serum replacement (KSR) was added on day five post-transduction and changed every three days thereafter until the indicated point of analysis. Reprogramming experiments described as being performed in FBS did not receive KSR-containing media. Reprogramming efficiency was determined by quantifying the number of GFP-reporter positive colonies or by immunostaining for Nanog on fixed reprogramming cultures. All iOSCK MEF reprogramming experiments followed the same FBS to KSR schedule with media containing 2 µg/ml doxycycline. When iOSCK reprogramming was combined with pMX-transgene overexpression (for Tcf3 or Dkk1), cells were cultured in media containing 2 µg/ml doxycycline to induce reprogramming 12 hr after infection with the respective pMX virus. For OSK and OCK reprogramming experiments using dox-inducible Tcf3, cells were selected in 100 µg/ml hygromycin for the first eight days of reprogramming. In all Tcf3 knockout reprogramming experiments, Tcf3 conditional MEFs (Merrill et al., 2004) carrying a *Cre-ERT2* transgene driven by the human ubiquitin C promoter (JAX

007001) were induced to reprogram by transduction of retroviruses expressing the indicated reprogramming factors. Treatment with 1 μM 4-Hydroxytamoxifen (Sigma H7904) for 24 hr induced efficient ablation of *Tcf3*, and was performed at indicated time points. Ethanol was used as a vehicle control.

Recombinant Wnt3a and IWP2 were purchased from Peprotech (315-20) and Stemgent (04-0034), respectively, and used at 80ng/ml (Wnt3a) and 2 μM (IWP2), unless otherwise indicated. Water and DMSO were used as vehicle controls for recombinant Wnt3a and IWP2, respectively. For siRNA knockdown experiments, 20nM siRNAs with the following sequences were transfected with RNAi-MAX: Tcf1 (CTACGAACATTTCAGCAGT), Lef1 (AATGAGAGCGAATGTCGTA), Tcf3 (GAGAAGAAACCTCACGTGA or GGCACAACCTGTCAAGAG (used in the OCK pre-iPSCs experiments)), and Tcf4 (TCACGCCTCTCATCACGTA). siRNA targeting Luciferase was purchased from Dharmacon RNA Technologies and used for control knockdown, and was added to all siRNA conditions targeting individual Tcfs to obtain identical siRNA load relative to double or triple knockdown conditions. Unless otherwise stated, a single round of siRNA transfection was used in all reprogramming experiments at the indicated time points.

To isolate and expand iPSC lines, colonies from primary reprogramming cultures were transferred with a micropipette to 96 well plates, trypsinized, plated into 24 well plates, and subsequently passaged into 6 well plates with continuous culture in ESC media containing Lif and 15% of either FBS or KSR. Culture vessels were always gelatinized and seeded with irradiated feeders. For passaging of all iPSC lines derived from KSR reprogramming conditions, trypsin inhibitor (Invitrogen 17075-029) was used as instructed. Embryoid body differentiation was done as described (Yi et al., 2008). For teratoma formation, 500,000 cells for each iPSC line were injected subcutaneously into the leg muscle of SCID mice. Teratomas were recovered

three weeks post-injection, fixed overnight in 4% formaldehyde, paraffin embedded, and processed with hematoxylin and eosin. For chimera formation, iPSCs were injected into host blastocysts derived from black C57Bl/6 mice at the UCLA Transgenic Facility.

OCK pre-iPSCs were generated from MEFs carrying a GFP reporter in the endogenous *Nanog* locus as described previously (Sridharan et al., 2009) and cultured on feeders in mouse ESC media containing Lif and 15% KSR. OCK pre-iPSCs for the siRNA experiment described in Figures 3.6A and 3.6B were sorted out by flow cytometry to remove the very low percentage of Nanog GFP positive typically observed in these cultures. The resulting GFP negative population was confirmed to stay GFP negative by microscopy during the length of the experiment. FACS sorted, GFP negative OCK pre-iPSCs were grown on feeders in ESC media containing Lif with 1.5% FBS and 13.5% KSR for 24hrs to allow recovery, 500,000 cells were plated on gelatinized dishes overnight in the same media, and media was then changed to the ESC with Lif and 15% KSR medium, and transfected with siRNAs using the PepMute reagent from SignaGen Laboratories.

RT- and Genotyping PCRs

Total RNA was isolated from cells using the RNeasy kit (QIAGEN) and cDNA generated with Superscript III (Invitrogen). qPCR values were generated using the ddCT method normalized to U6, unless otherwise indicated. Primers used for detecting expression of pMX transgenes and in the embryoid body assays have been previously described (Maherali et al., 2007; Yi et al., 2008). qPCR primers used for detecting mRNA expression are as follows: Axin2*:* GAGAGTGAGCGGCAGAGC, CGGCTGACTCGTTCTCCT; Brachyury: GCTTCAAGGAGCTAACTAACGAG, CCAGCAAGAAAGAGTACATGGC; Cdh1: FW: GACGCTGAGCATGTGAAGAA, RV: CAGGACCAGGAGAAGAGTGC, Cdx1: FW: ACGCCCTACGAATGGATG, RV: CCTTGGTTCGGGTCTTACC, Esrrb: FW:

CTCTCATTTGGGCCTAGCAG, RV: CCCTCCTGTCTCCTTGTCAC, Gapdh: FW: GTGTCCGTCGTGGATCTGA, RV: CCTGCTTCACCACCTTCTTG, Lef1: FW: TGAGTGCACGCTAAAGGAGA, RV: ATAATTGTCTCGCGCTGACC, Nanog: FW: AGGGTCTGCTACTGAGATGCTCTG, RV: CAACCACTGGTTTTTCTGCCACCG, Sox1: FW: CACAACTCGGAGATCAGCAA, RV: CTTGAGCAGCGTCTTGGTCT, Sox2: FW: TAGAGCTAGACTCCGGGCGATGA, RV: TTGCCTTAAACAAGACCACGAAA, Sox3: FW: CACAACTCCGAGATCAGCAA, RV: CTTGAGCAGCGTCTTGGTCT, Sox4: FW: CCGACATGCACAACGCCGAGAT, RV: TTGCCCGACTTCACCTTCTTTCG, Sox11: FW: GACTGGTGCAAGACGGCGTCGG, RV: AGCCTCTTGGAGATCTCGGCGT, Sox15: FW: AGCACCGGGTCTGTCCCCTT, RV: GCAGTGGGAAGAGGTGTAACTGCC, Sox18: FW: ATGGCTTTGGCCGCGGAGAG, RV: CCTTCCACGCTTTGCCCAGCA, Tcf1: FW: ATCTGCTCATGCCCTACCC, RV: GGTGTGGACTGCTGAAATGTT, Tcf3: FW: AGCTCGGACTCCGAGGCGGAGA, RV: GGTACCCAGGATACGCAGGTCC, Tcf4: FW: CACCCGGCCATCGTCACAC, RV: GCCACCTGCGCCCGAGAAT, U6: FW: CGCTTCGGCAGCACATATAC, RV: TCACGAATTTGCGTGTCATC. PCR primers used for genotyping are as follows: Tcf3: FW: AGTCGTCCCTGGTCAACGAATCGG, RV: ACAGAGTAGCTATCTGGAGCTCGG; pRetro-Tcf3: FW: AGATCGCCTGGAGAAGGATC, RV: GGTACCCAGGATACGCAGGTCC.

Gene Expression Data Analysis

RNA expression profiling was performed on the Affymetrix Gene Chip Mouse Genome 430 2.0 arrays at the UCLA microarray core facility. All microarray data from this platform were normalized together with Robust Multichip Analysis (RMA) in R using the Bioconductor package. Our analyses are based on the high confidence probes ending in "_at" and "a_at" and for all probes, any expression value below 50 was set to 50 before further analysis. Pearson correlations were performed in R and hierarchical clustering in Cluster 3.0 and visualized by

Java Treeview. Gene ontology analysis was performed using GATHER (Chang and Nevins, 2006), and a Bayes factor of 6 was used as the significance cut off. STEM analysis was done as described (Ernst and Bar-Joseph, 2006), without specifying the number of profiles generated, thereby obtaining 'unsupervised' profiles that best fit the data. Five significant group of co-regulated genes were obtained, and the top three groups with the highest number of probes are discussed. Gene Set Enrichment Analysis (GSEA) was performed following recommendations detailed on the GSEA website (Mootha et al., 2003; Subramanian et al., 2005) with 1000 permutations of the gene sets and a family-wise error rate p-value of less than 0.05 was accepted as significant. Gene sets for *Tcf3* knockout in ESCs were derived from previously published expression data (Yi et al., 2008).

Western Blotting, Immunostaining, AP Detection, Cytometry, and Luciferase Assay

The antibodies used for Western blotting were rabbit anti-Tcf3 (generated by the Merrill lab), and mouse anti-Gapdh (Fitzgerald RD1-TRK5G4-6C5). For immunostaining, cell lines or reprogramming cultures were grown on glass coverslips, fixed with 4% paraformaldehyde (PFA) in PBS for 10 min, permeabilized with 0.5% Triton 100-X in PBS for 5 min, and immunostained with rabbit anti-Nanog (Cosmo Bio RCAB0002P-F) or goat anti-Tcf3 (Santa Cruz Biotechnology sc-8635). Clusters of cells with at least six nuclei staining positive for Nanog were scored as a Nanog positive colony. For AP detection, reprogramming cultures were fixed with 4% PFA in PBS for 20 min, washed with water, and incubated with a mixture of 1mg/ml Fast Red TR Salt (Sigma) and 4% Napthol AS-MX Solution (Sigma) for 10 min. For Crystal Violet staining, fixed cultures were incubated in 0.05% Crystal Violet (Sigma C3886) solution for 30 min and washed three times with distilled water before scanning. For the FACS experiments involving SSEA1 detection, trypsinized cells were stained for 30 min in PBS with the Alexa Flour 647-labeled SSEA1 antibody from Biolegends (1:100 dilution). FACSDiva (BD Biosciences) was used to quantify the percentage of *Nanog*-GFP positive cells from at least 100,000 cells, and FACS

plots were analyzed using Flowjo software. TOPFLASH-Luciferase assays were performed using a lentiviral BAR and fuBAR luciferase reporter (Biechele et al., 2009), Promega Dual-Glo reagents and GLOMAX Luminometer.

Figure 3.1. Biphasic Role of Wnt Signaling in Reprogramming to iPSCs

A. iOSCK MEFs were transduced with a *Tomato* (Ctrl) and *Dkk1*-encoding retrovirus, respectively, and treated with dox to express the reprogramming factors, and Nanog-positive colonies were quantified at day 13.

B. iOSCK MEFs were treated with dox and IWP2 or vehicle, respectively, continuously throughout reprogramming, and Nanog positive colonies were counted at day 9.

C. As in (B), except that Wnt3a was added continuously throughout reprogramming.

D. Schematic of the reprogramming experiments to determine the effect of Wnt3a or IWP2 during different stages of reprogramming.

E. Nanog colony count for IWP2-treated reprogramming cultures as described in (D).

F. Nanog colony count for Wnt3a-treated reprogramming cultures as described in (D). All reprogramming counts represent the average of two representative experiments, and error bars depict standard deviation.

See Figure S3.1 for additional information.

Figure 3.1

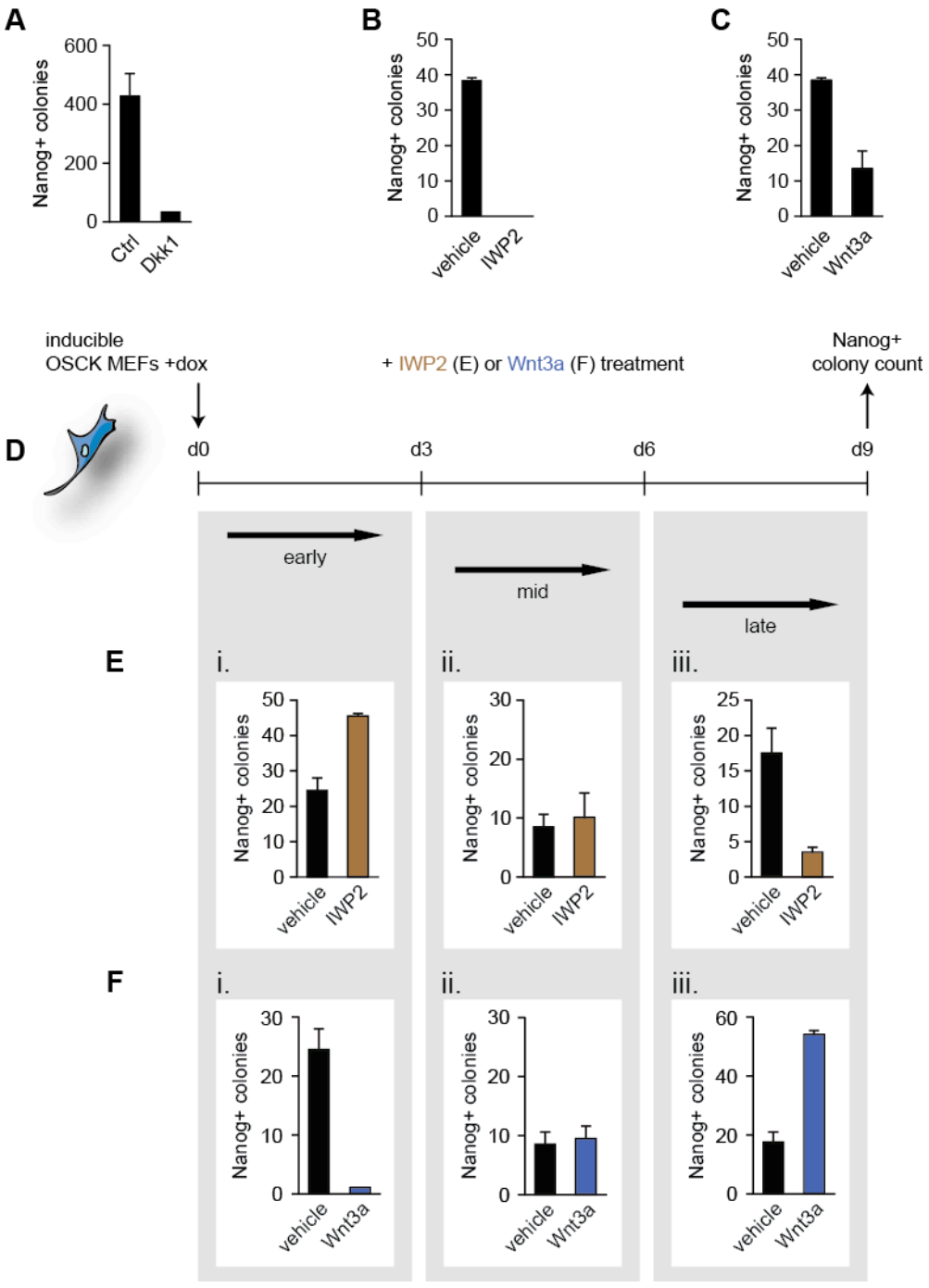

A
Nanog+ colonies
600
400
200
0
Ctrl
Dkk1
B
Nanog+ colonies
50
40
30
20
10
0
vehicle
IWP2
C
Nanog+ colonies
50
40
30
20
10
0
vehicle
Wnt3a
inducible
OSCK MEFs +dox
+ IWP2 (E) or Wnt3a (F) treatment
Nanog+
colony count
D
d0
d3
d6
d9
early
mid
late
E
i.
ii.
iii.
Nanog+ colonies
vehicle
IWP2
F
i.
ii.
iii.
Nanog+ colonies
vehicle
Wnt3a

Figure 3.2. Tcf1 and Lef1 Inhibit whereas Tcf3 and Tcf4 Promote the Early Phase of Reprogramming

A. Schematic of the reprogramming experiment testing the requirement of Tcfs in the early phase of reprogramming. iOSCK MEFs were transfected with siRNAs targeting Tcfs individually, in combination or with siCtrl twice: first 12 hr prior to the induction of OSCK factors and again together with dox addition. Knockdown was confirmed on day 3 of reprogramming, and Nanog-positive colonies were counted at day 10 (experiment 1) or day 9 (experiment 2).

B. Nanog-positive colony count from two independent experiments (Exp1 and Exp2), each with two technical replicates (A and B).

C. qPCR for *Lef1*, *Tcf1*, and *Axin2* transcript levels relative to siCtrl upon *Tcf3*/*Tcf4* double knockdown at day 3 of reprogramming. Values represent the average of duplicate sampling, and error bars represent standard deviation.

D. As in (B), except that reprogramming cultures were treated with IWP2 or vehicle from day 0 to day 3 in addition to indicated siRNAs. Nanog-positive colony count from two independent experiments (Exp1 and Exp2), each with two technical replicates (A and B) is given.

E. Number of Nanog-positive colonies upon simultaneous knockdown of all four Tcfs early in reprogramming. Knockdown was performed as part of experiment 1 shown in (B) and experimental conditions shared with (B) are indicated by asterisks.

See Figure S3.2 for additional information.

Figure 3.2

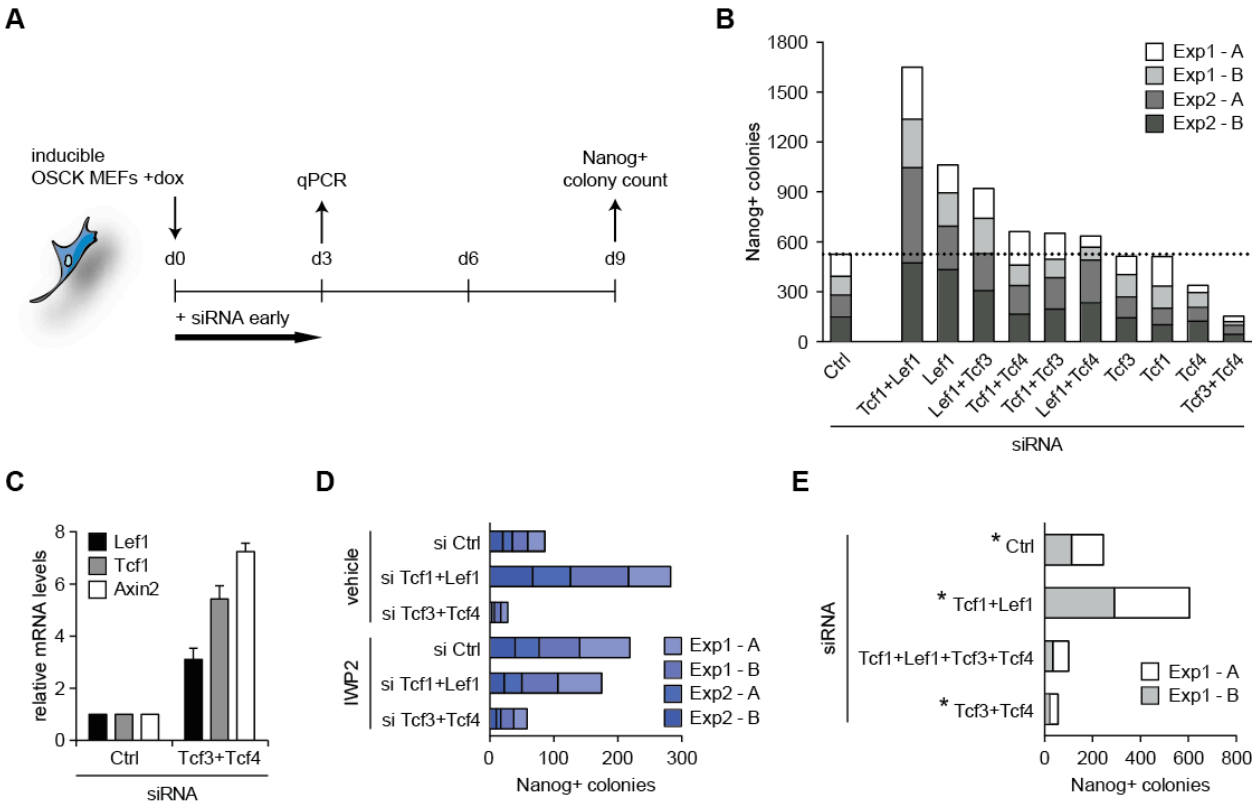

A
inducible
OSCK MEFs +dox
qPCR
Nanog+
colony count
d0
d3
d6
d9
+ siRNA early
B
Nanog+ colonies
Exp1 - A
Exp1 - B
Exp2 - A
Exp2 - B
Ctrl
Tcf1+Lef1
Lef1
Lef1+Tcf3
Tcf1+Tcf4
Tcf1+Tcf3
Lef1+Tcf4
Tcf3
Tcf1
Tcf4
Tcf3+Tcf4
siRNA
C
relative mRNA levels
Lef1
Tcf1
Axin2
Ctrl
Tcf3+Tcf4
siRNA
D
vehicle
IWP2
si Ctrl
si Tcf1+Lef1
si Tcf3+Tcf4
Exp1 - A
Exp1 - B
Exp2 - A
Exp2 - B
Nanog+ colonies
E
siRNA
Ctrl
Tcf1+Lef1
Tcf1+Lef1+Tcf3+Tcf4
Tcf3+Tcf4
Exp1 - A
Exp1 - B
Nanog+ colonies

Figure 3.3. *Tcf3/Tcf4* Depletion Enhances the Late Phase of Reprogramming in a *Tcf1/Lef1*-Dependent Manner

A. Schematic of the reprogramming experiment testing the requirement of Tcfs in the late phase of reprogramming. iOSCK MEFs were transfected with siRNAs targeting Tcfs individually or in combination with siCtrl once at day 6 postinduction of OSCK. Transcript levels and Nanog-positive colonies were quantified on day 10 (experiment 1) or day 9 (experiment 2).

B. Nanog-positive colony count from two independent experiments (Exp1 and Exp2), each with two technical replicates (A and B).

C. As in (B), except that the cultures were treated with IWP2 or vehicle from day 6 to day 9 in addition to indicated siRNAs. Nanog-positive colony count from technical replicates of a representative experiment is given.

See Figure S3.3 for additional information.

Figure 3.3

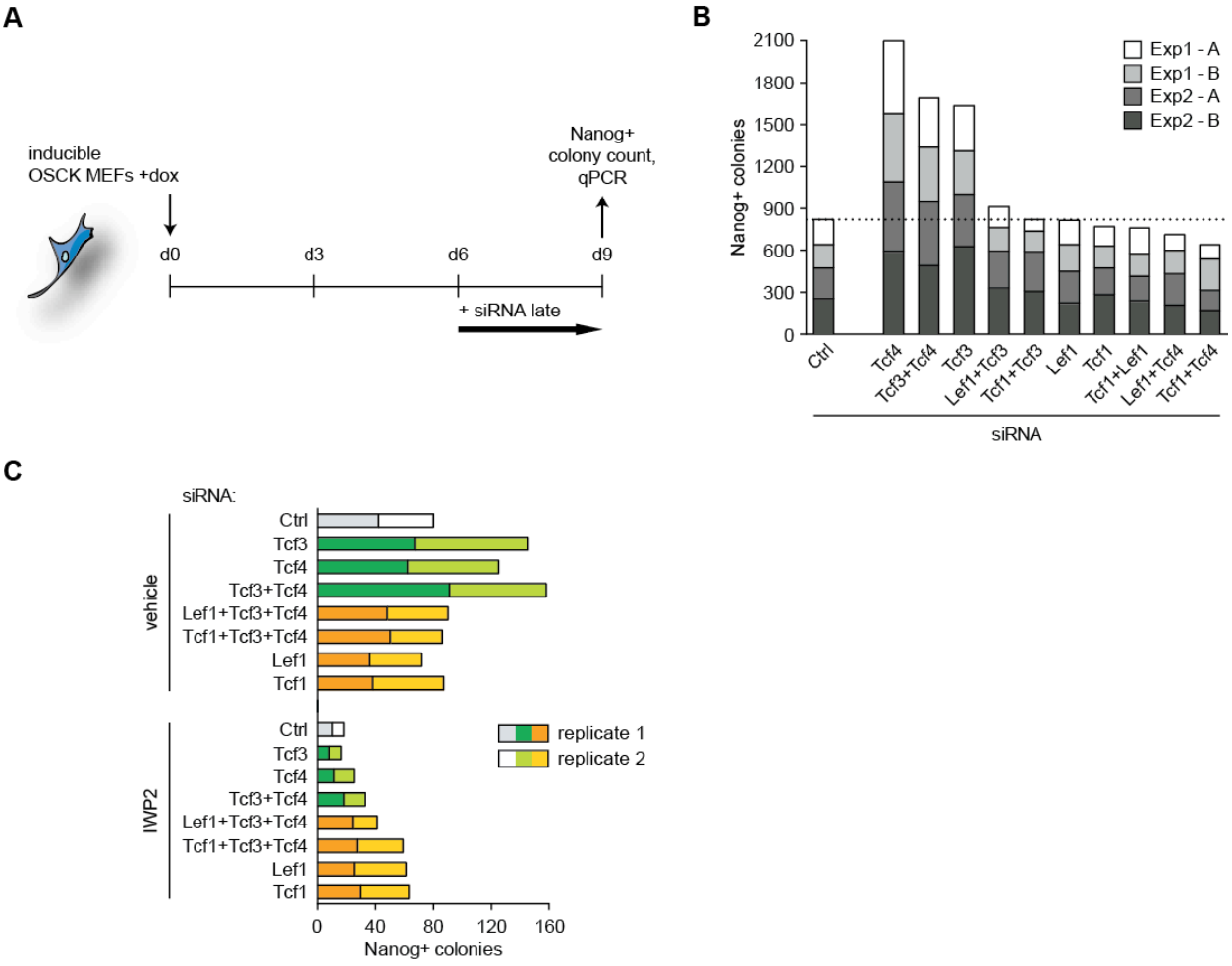

A
inducible
OSCK MEFs +dox
d0
d3
d6
d9
Nanog+
colony count,
qPCR
+ siRNA late
B
Nanog+ colonies
2100
1800
1500
1200
900
600
300
0
Exp1 - A
Exp1 - B
Exp2 - A
Exp2 - B
Ctrl
Tcf4
Tcf3+Tcf4
Tcf3
Lef1+Tcf3
Tcf1+Tcf3
Lef1
Tcf1
Tcf1+Lef1
Lef1+Tcf4
Tcf1+Tcf4
siRNA
C
siRNA:
vehicle
IWP2
Ctrl
Tcf3
Tcf4
Tcf3+Tcf4
Lef1+Tcf3+Tcf4
Tcf1+Tcf3+Tcf4
Lef1
Tcf1
replicate 1
replicate 2
0
40
80
120
160
Nanog+ colonies

Figure 3.4. Ectopic Expression of Tcf3 Promotes Early Reprogramming Events but Inhibits the Induction of Pluripotency

A. Experimental design for continuous Tcf3 overexpression from the pMX-retroviral vector during reprogramming of iOSCK MEFs. Tomato expression served as control overexpression (Ctrl).

B. Fluorescence-activated cell sorting (FACS) analysis showing the percentage of SSEA1 positive cells at indicated times of a representative reprogramming experiment.

C. Nanog colony count representing the average of two independent experiments, each with two technical replicates: one counted on day 11 and one on day 15. Error bars indicate standard deviation of the averaged values.

D. Transcript levels of *E-cadherin* during reprogramming relative to MEFs. Values represent the average of duplicate sampling, and error bars represent standard deviation.

E. Staining for AP activity at day 10 of reprogramming.

F. Schematic of the Tcf3 domain structure and Tcf3 mutants used in this reprogramming experiment (i). Reprogramming was performed as in (A), with Tcf3 variants retrovirally expressed throughout reprogramming. AP colony count of a representative reprogramming experiment at day 11 is given (ii).

See Figure S3.4 for additional information.

Figure 3.4

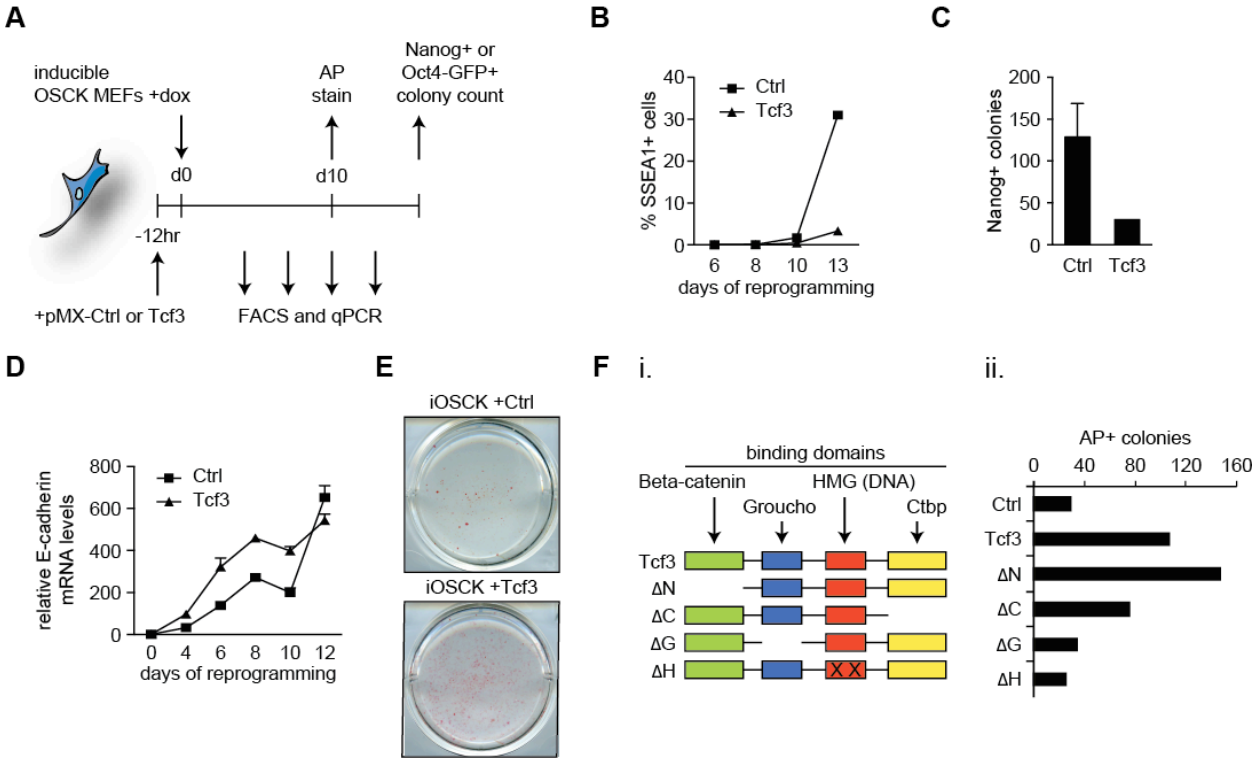

A
inducible
OSCK MEFs +dox
AP
stain
Nanog+ or
Oct4-GFP+
colony count
d0
d10
-12hr
+pMX-Ctrl or Tcf3
FACS and qPCR
B
% SSEA1+ cells
Ctrl
Tcf3
days of reprogramming
C
Nanog+ colonies
Ctrl
Tcf3
D
relative E-cadherin mRNA levels
Ctrl
Tcf3
days of reprogramming
E
iOSCK +Ctrl
iOSCK +Tcf3
F
i.
binding domains
Beta-catenin
HMG (DNA)
Groucho
Ctbp
Tcf3
ΔN
ΔC
ΔG
ΔH
ii.
AP+ colonies
Ctrl
Tcf3
ΔN
ΔC
ΔG
ΔH

Figure 3.5. *Tcf3* Ablation Allows Reprogramming in the Absence of Sox2

A. PCR genotyping of *Tcf32loxp/2loxp*;Tg(*UBC-Cre-ERT2*) MEFs after treatment with 1 μM tamoxifen (tam) or vehicle for 24 hr. Cells with homozygous 1loxp alleles are referred to as KO or *Tcf3−/−*, and cells with homozygous 2loxP alleles as WT or *Tcf3+/+*.

B. Western blot for Tcf3 on MEFs described in (A) 24–72 hr after tam addition. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) served as a loading control.

C. *Tcf32loxp/2loxp*;Tg(*UBC-Cre-ERT2*) MEFs were transduced with separate retroviruses encoding OSCK (i) or OSK (ii) and treated with and without tamoxifen, respectively, at day 4 to delete *Tcf3*. Nanog-positive colonies were quantified at the indicated days of reprogramming by fixing parallel reprogramming wells and immunostaining for Nanog.

D. Immunostaining for Nanog and Tcf3 in iPSC lines isolated from OSK or OCK reprogramming cultures in which *Tcf3* was deleted at day 4. DAPI staining marks nuclei. ESCs and OSK WT iPSCs serve as controls.

E. Hierarchical clustering of log2 expression ratios of indicated cell lines relative to the average intensity of each probe across all arrays. Only probes 2-fold differentially expressed between ESCs and MEFs were included.

F. *Tcf3* WT and *Tcf3* KO OSK and OCK iPSC lines were differentiated by embryoid bodies. RNA was harvested at indicated time points and analyzed by semiquantitative RT-PCR for expression of representative genes of each of the three embryonic germ layers. GAPDH serves as a loading control.

See Figure S3.5 for additional information.

Figure 3.5

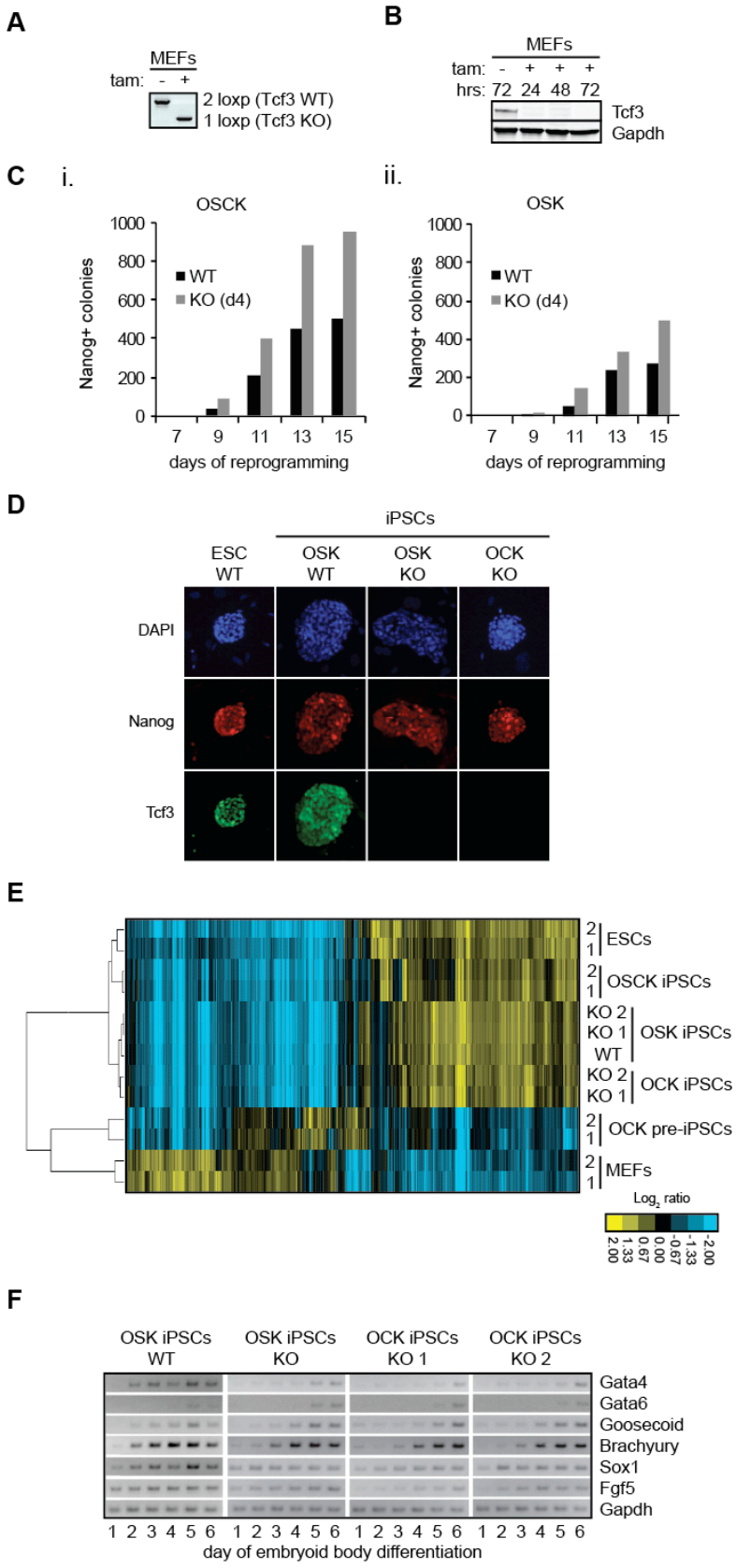

A
MEFs
tam: - +
2 loxp (Tcf3 WT)
1 loxp (Tcf3 KO)
B
MEFs
tam: - + + +
hrs: 72 24 48 72
Tcf3
Gapdh
C
i.
OSCK
Nanog+ colonies
WT
KO (d4)
days of reprogramming
ii.
OSK
Nanog+ colonies
WT
KO (d4)
days of reprogramming
D
iPSCs
ESC WT
OSK WT
OSK KO
OCK KO
DAPI
Nanog
Tcf3
E
ESCs
OSCK iPSCs
KO 2
KO 1
WT
OSK iPSCs
KO 2
KO 1
OCK iPSCs
OCK pre-iPSCs
MEFs
Log2 ratio
F
OSK iPSCs WT
OSK iPSCs KO
OCK iPSCs KO 1
OCK iPSCs KO 2
Gata4
Gata6
Goosecoid
Brachyury
Sox1
Fgf5
Gapdh
day of embryoid body differentiation

Figure 3.6. Regulation of OCK Reprogramming by Tcf3

A. Schematic of the OCK pre-iPSC reprogramming experiment. Different siRNAs were transfected twice into wild-type OCK pre-iPSCs containing the *Nanog*-GFP reporter. Knockdown efficiency was assessed by qPCR 3 days after the initial transfection, and GFP-positive colonies were quantified 4 days after the initial siRNA transfection.

B. *Nanog*-GFP positive colony count for a representative experiment with four technical replicates per condition is shown.

C. Schematic of the OCK reprogramming experiment used for gene expression analysis.*Tcf32loxp/2loxp*;Tg(*UBC-Cre-ERT2*) MEFs were infected retrovirally with OCK and split on day 3. Half of the culture was treated with tamoxifen (tam) at day 4 to generate the Tcf3 KO condition, and the other half was exposed to vehicle control (WT). Both WT and KO reprogramming cultures were split again on day 21 to enhance the effect of *Tcf3* deletion on the induction of pluripotency. RNA from KO and WT reprogramming cultures was harvested at indicated time points from parallel reprogramming wells and analyzed for gene expression.

D. Quantification of Nanog-positive colonies at day 26 of the OCK reprogramming experiment described in (C).

E. STEM analysis for all transcripts that are at least 2-fold differentially expressed between *Tcf3* KO and WT OCK reprogramming cultures at any of the profiled time points during OCK reprogramming. The top three groups with significant patterns of coregulated gene expression changes are shown, and the number of probes and genes belonging to each group is indicated. Left: log2 expression ratio between *Tcf3* KO and WT cultures for the probes in each group; middle: heat maps displaying the log2 expression ratios of probes belonging to these groups for both WT and KO reprogramming cultures, respectively, and of ESCs, all relative to MEFs; and right: significantly enriched GO terms for each group. Example genes from each group are also given.

See Figure S3.6 for additional information.

Figure 3.6

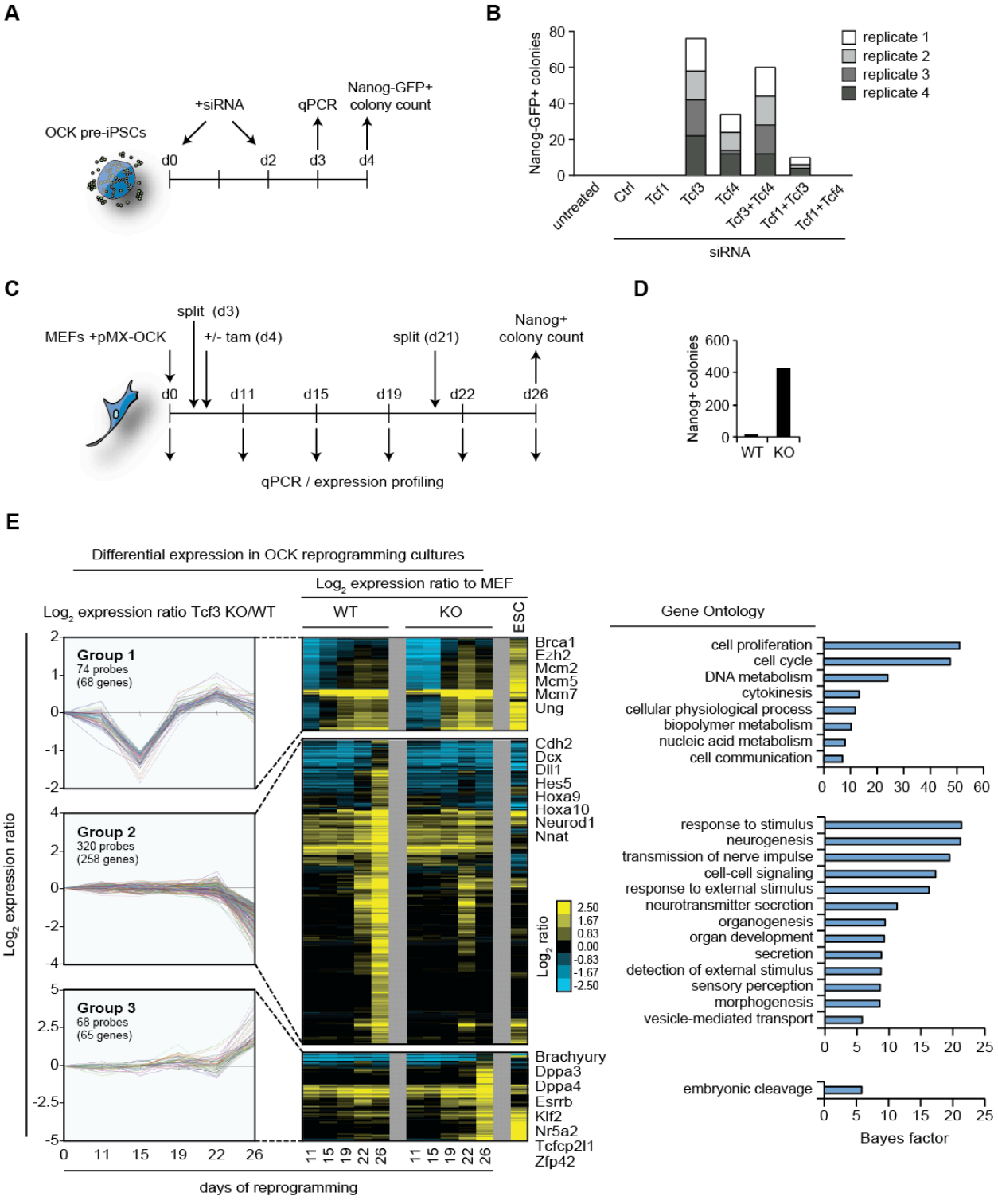

A
OCK pre-iPSCs
+siRNA
qPCR
Nanog-GFP+ colony count
d0
d2
d3
d4
B
Nanog-GFP+ colonies
replicate 1
replicate 2
replicate 3
replicate 4
untreated
Ctrl
Tcf1
Tcf3
Tcf4
Tcf3+Tcf4
Tcf1+Tcf3
Tcf1+Tcf4
siRNA
C
MEFs +pMX-OCK
split (d3)
+/- tam (d4)
split (d21)
Nanog+ colony count
d0
d11
d15
d19
d22
d26
qPCR / expression profiling
D
Nanog+ colonies
WT
KO
E
Differential expression in OCK reprogramming cultures
Log2 expression ratio to MEF
Log2 expression ratio Tcf3 KO/WT
WT
KO
ESC
Group 1
74 probes
(68 genes)
Group 2
320 probes
(258 genes)
Group 3
68 probes
(65 genes)
Log2 expression ratio
Brca1
Ezh2
Mcm2
Mcm5
Mcm7
Ung
Cdh2
Dcx
Dll1
Hes5
Hoxa9
Hoxa10
Neurod1
Nnat
Brachyury
Dppa3
Dppa4
Esrrb
Klf2
Nr5a2
Tcfcp2l1
Zfp42
Log2 ratio
2.50
1.67
0.83
0.00
-0.83
-1.67
-2.50
days of reprogramming
Gene Ontology
cell proliferation
cell cycle
DNA metabolism
cytokinesis
cellular physiological process
biopolymer metabolism
nucleic acid metabolism
cell communication
response to stimulus
neurogenesis
transmission of nerve impulse
cell-cell signaling
response to external stimulus
neurotransmitter secretion
organogenesis
organ development
secretion
detection of external stimulus
sensory perception
morphogenesis
vesicle-mediated transport
embryonic cleavage
Bayes factor

Figure 3.7. Stepwise Modulation of *Tcf3* Levels Enables Efficient Reprogramming in the Absence of Sox2

A. Scheme for the OCK reprogramming experiment with *Tcf3* level modulation. MEFs carrying the *M2rtTA* and*Oct4-GFP* transgenes were infected with a dox-inducible retrovirus (pRetro) encoding *Tcf3* and subsequently with constitutive retroviruses (pMX) encoding *OCK*. Dox was added during the first 8 days of reprogramming at different concentrations, and subsequently, *Tcf3* was depleted by repeated siRNA transfection every 3 days beginning on day 8 until day 29. *Oct4*-GFP-positive colonies were quantified at indicated days. Hygromycin was added from d0–d8 of reprogramming to ensure that all cells carry the pRetro-Hygro-*Tcf3*-expressing vector.

B. Titration of *Tcf3* transcript levels is achieved by varying dox concentrations in OCK reprogramming. *Tcf3*transcript levels were measured in the reprogramming culture at day 4 and are presented relative to MEF levels. Note that 0.002 μg/ml dox induces ESC-like levels of *Tcf3*. Values represent the average of triplicate sampling, and error bars represent standard deviation. Expression of Tcf3 in virtually all cells was confirmed by immunostaining (data not shown).

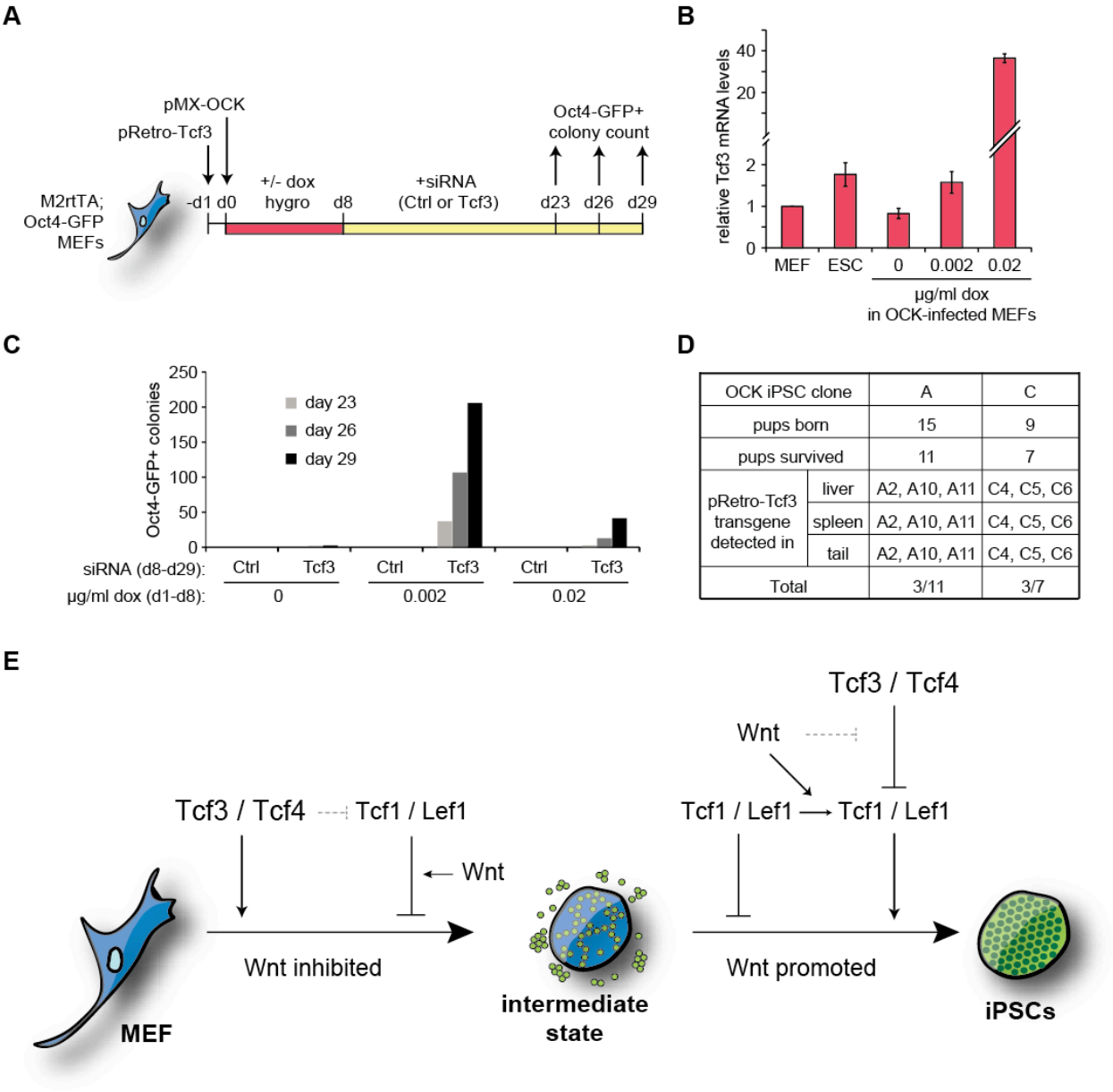
C. Quantification of *Oct4*-GFP-positive colonies at indicated days of reprogramming. *Tcf3* was induced at different levels using different amounts of dox, and siRNA was added as indicated.

D. Characterization of mice obtained upon blastocyst injection of OCK iPSC clones A and C, which were expanded from reprogramming cultures treated with 0.002 μg/ml dox and subsequently depleted for *Tcf3*, as described in (C). Result of PCRs with primers specifically amplifying proviral integrations of pRetro-*Tcf3* on genomic DNA extracted from liver, spleen, and tails are summarized.

E. Model summarizing the stage-specific roles of Tcfs and Wnt signaling in reprogramming to iPSCs, where the switch in Wnt response is associated with changing requirements for Tcfs and two pairs of Tcf family members, Tcf1/Lef1 and Tcf3/Tcf4, have opposing functions early and late in reprogramming.

See Figure S3.7 for additional information.

Figure 3.7



OCK iPSC clone		A	C
pups born		15	9
pups survived		11	7
pRetro-Tcf3 transgene detected in	liver	A2, A10, A11	C4, C5, C6
	spleen	A2, A10, A11	C4, C5, C6
	tail	A2, A10, A11	C4, C5, C6
Total		3/11	3/7

Figure S3.1. Characterization of the Effect of Wnt Inhibition and Stimulation on Reprogramming, Related to Figure 3.1.

A. Description of iOSCK MEFs. MEFs carrying the reverse tetracycline transactivator (*M2rtTA*) in the *R26*locus and a doxycycline (dox)-inducible polycistronic *OSCK* cassette in the *Col1A* locus (iOSCK) were derived from mice carrying the respective alleles. Upon addition of dox, OSCKs become expressed in virtually all cells (data not shown).

B. Characterization of reprogramming in the presence of *Dkk1* overexpression. qPCR analysis showing the transcript levels of the known Wnt target gene *Axin2* at the indicated time points of iOSCK reprogramming in the presence of retroviral *Tomato* (Ctrl) or *Dkk1* expression. The data are shown relative to levels in *Tomato*-infected iOSCK MEFs at day four of reprogramming. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation. We conclude that *Dkk1* expression reduces*Axin2* levels throughout reprogramming.

C. Confirmation of Wnt3a and Dkk1 activity. iOSCK MEFs were lentivirally infected with a TOPflash-driven luciferase construct or a control construct (TOPflash and unresponsive control) (Biechele et al., 2009) allowing us to assess the activity of the Wnt signaling pathway (TOPflash assay). Luciferase activity was assayed at 72 hr of iOSCK reprogramming in which *Tomato* (Ctrl) or *Dkk1* were retrovirally overexpressed and recombinant Wnt3a was added as indicated. The average luciferase activity from technical replicates of a representative experiment is shown with error bars indicating standard deviation. The data show that treatment with recombinant Wnt3a parallel to iOSCK expression induces Tcf-reporter activity significantly, but does not affect expression of the unresponsive control construct. Concomitant overexpression of *Dkk1* counteracts the Wnt3a-dependent induction of the reporter. *Dkk1* overexpression does not reduce basal TOPflash activity, which is close to background levels without ectopic Wnt3a treatment, as commonly observed, likely caused by the lack of sensitivity of the reporter system.

D. Schematic of the reprogramming experiment determining the effect of endogenous Wnt signaling inhibition by IWP2 treatment and of ectopic Wnt stimulation by Wnt3a during the indicated phases of reprogramming, related to Figures 3.1D–F. Compared to Figure 3.1D, an additional treatment period is indicated, where MEFs were preincubated for 3 days with IWP2 or Wnt3a before induction of reprogramming factor expression. Experimental results are described in (E) through (G).

E. (i) IWP2 or vehicle were added early in reprogramming at the indicated concentrations, and Nanog colony number was assessed at day nine. A representative experiment is shown with average counts of two technical replicates with error bars indicating standard deviation, indicating that IWP2 treatment enhances the formation of Nanog positive colonies in a dose-dependent manner. (ii) 2 μM IWP2 were added early in reprogramming and the expression of Wnt-regulated genes *Axin2* and *Tcf1* was determined at day 3. Transcript levels are relative to vehicle conditions. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation. As expected, blockage of Wnt ligands results in decreased expression of Wnt targets. (iii) qPCR showing that *Lef1*, *Tcf1*, and *Axin2* transcript levels are increased when recombinant Wnt3a is added early in reprogramming. Expression was analyzed 72 hr after dox and Wnt3a addition. Transcript levels are relative to

vehicle. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

F. (i) As in (E), except that IWP2 was added late in reprogramming, demonstrating that IWP2 late in reprogramming inhibits the formation of Nanog positive colonies in a dose-dependent manner. A representative experiment is shown with average counts of two technical replicates with error bars indicating standard deviation. (ii) IWP2 treatment reduces the expression of Wnt-regulated genes *Axin2* and *Tcf1*. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation. Note that DMSO (vehicle) is consistently observed to enhance at higher concentrations in late reprogramming phases.

G. iOSCK MEFs were treated with IWP2 or Wnt3a for 72 hr prior to dox-induction and the number of Nanog positive colonies was determined on day nine, demonstrating that IWP2 pretreatment enhances and Wnt3a preaddition inhibits iOSCK reprogramming. A representative experiment is shown with average counts of two technical replicates with error bars indicating standard deviation.

Figure S3.1

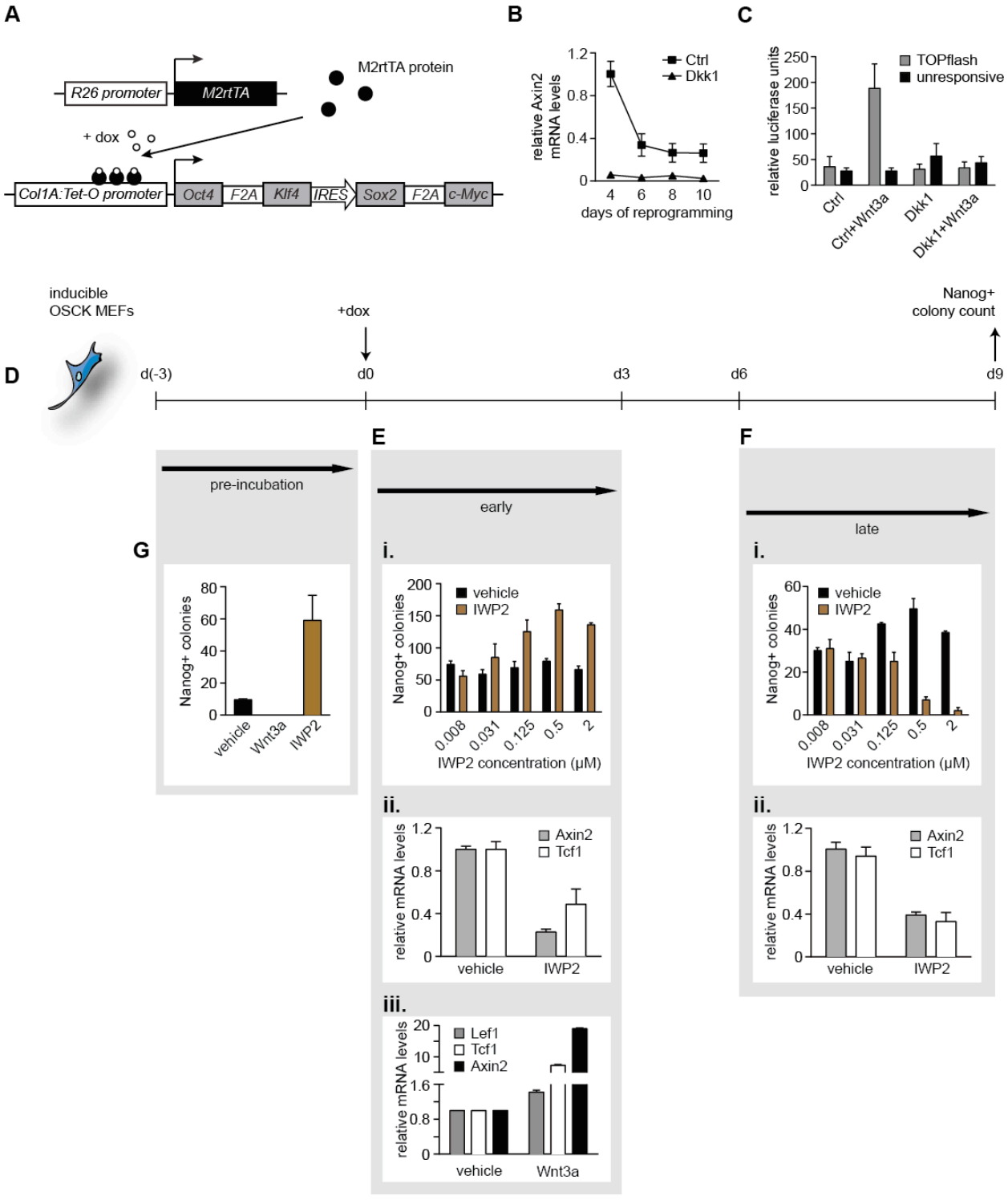

A
R26 promoter
M2rtTA
M2rtTA protein
+ dox
Col1A:Tet-O promoter
Oct4
F2A
Klf4
IRES
Sox2
F2A
c-Myc
B
relative Axin2 mRNA levels
Ctrl
Dkk1
days of reprogramming
C
relative luciferase units
TOPflash
unresponsive
Ctrl
Ctrl+Wnt3a
Dkk1
Dkk1+Wnt3a
D
inducible OSCK MEFs
+dox
Nanog+ colony count
d(-3)
d0
d3
d6
d9
E
pre-incubation
early
late
F
G
i.
ii.
iii.
Nanog+ colonies
vehicle
Wnt3a
IWP2
IWP2 concentration (μM)
relative mRNA levels
Axin2
Tcf1
Lef1

Figure S3.2. Characterization of the Requirement of Tcfs Early in Reprogramming, Related to Figure 3.2

A. qPCR showing the mRNA levels of *Tcf1* relative to siCtrl in reprogramming cultures that were transfected with siRNAs targeting *Tcf1* for the experiment shown in Figure 3.2B, where Tcfs were depleted early in the process. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

B. As in (A), except for *Tcf3*.

C. As in (A), except for *Tcf4*.

D. As in (A), except for *Lef1*.

E. qPCR showing the transcript levels of *Tcf1* and *Lef1* relative to siCtrl upon knockdown of *Tcf1/Lef1* early in reprogramming in the presence or absence of IWP2 for the experiment described in Figure 3.2D. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

F. As in (E), except for *Tcf3* and *Tcf4* upon *Tcf3/Tcf4* knockdown.

G. As in (E), except for transcript levels of the Wnt-regulated genes *Tcf1*, *Lef1*, and *Axin2* upon *Tcf3/Tcf4*knockdown.

H. qPCR showing the mRNA levels of *Tcf1* and *Lef1* relative to siCtrl upon knockdown of *Tcf1/Lef1*, all four Tcfs, or *Tcf3/Tcf4*, respectively, early in reprogramming for the experiment described in Figure 3.2E. Note the cross-regulation among the Tcf family members. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation

I. As in (H), except for *Tcf3* and *Tcf4*.

Figure S3.2

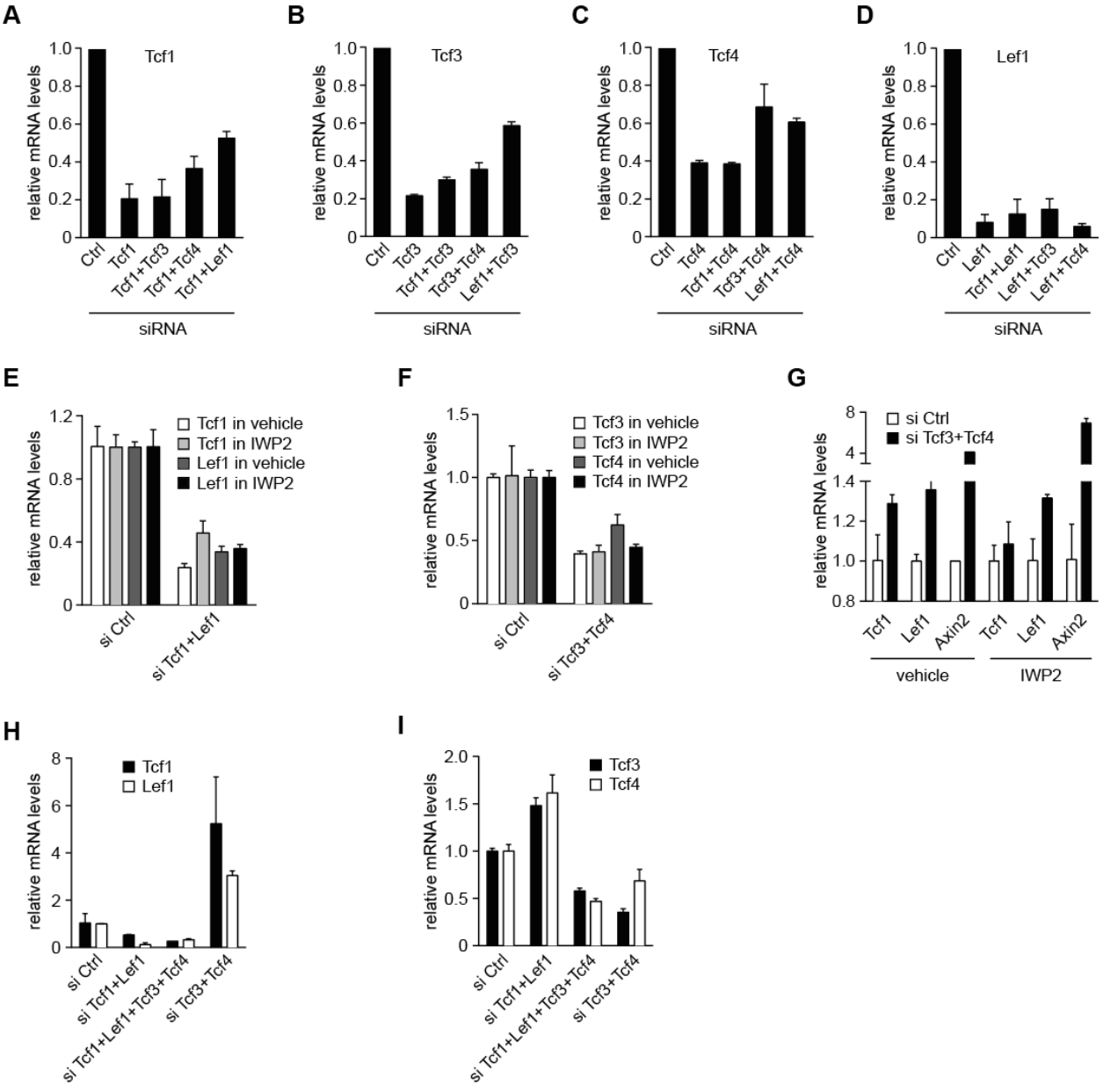

A
Tcf1
relative mRNA levels
Ctrl
Tcf1
Tcf1+Tcf3
Tcf1+Tcf4
Tcf1+Lef1
siRNA
B
Tcf3
Ctrl
Tcf3
Tcf1+Tcf3
Tcf3+Tcf4
Lef1+Tcf3
siRNA
C
Tcf4
Ctrl
Tcf4
Tcf1+Tcf4
Tcf3+Tcf4
Lef1+Tcf4
siRNA
D
Lef1
Ctrl
Lef1
Tcf1+Lef1
Lef1+Tcf3
Lef1+Tcf4
siRNA
E
Tcf1 in vehicle
Tcf1 in IWP2
Lef1 in vehicle
Lef1 in IWP2
si Ctrl
si Tcf1+Lef1
F
Tcf3 in vehicle
Tcf3 in IWP2
Tcf4 in vehicle
Tcf4 in IWP2
si Ctrl
si Tcf3+Tcf4
G
si Ctrl
si Tcf3+Tcf4
Tcf1
Lef1
Axin2
vehicle
IWP2
H
Tcf1
Lef1
si Ctrl
si Tcf1+Lef1
si Tcf1+Lef1+Tcf3+Tcf4
si Tcf3+Tcf4
I
Tcf3
Tcf4
si Ctrl
si Tcf1+Lef1
si Tcf1+Lef1+Tcf3+Tcf4
si Tcf3+Tcf4

Figure S3.3. Characterization of the Requirement of Tcfs Late in Reprogramming, Related to Figure 3.3

A. qPCR showing the transcript levels of *Tcf1* relative to siCtrl in reprogramming cultures transfected with siRNAs targeting *Tcf1* for the experiment shown in Figure 3.3B, where Tcfs were depleted late in the reprogramming process. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

B. As in (A), except for *Tcf3*.

C. As in (A), except for *Tcf4*.

D. As in (A), except for *Lef1*.

E. An iOSCK reprogramming time course experiment was performed in which the indicated siRNA combinations were transfected late in reprogramming, at day six. Reprogramming cultures were analyzed every day from day seven to ten for the presence of Nanog positive colonies.

F. Count of Nanog positive colonies on day nine of reprogramming upon treatment with indicated siRNAs at day six. These knockdowns were performed as part of experiment 2 shown in Figure 3.3B and the siCtrl and siTcf3/Tcf4 data are the same as those shown in Figure 3.3B.

G. qPCR showing the transcript levels of the Wnt-regulated genes *Tcf1*, *Lef1*, and *Axin2* relative to siCtrl in the vehicle condition upon knockdown of *Tcf3/Tcf4* late in reprogramming in the presence and absence of IWP2 for the experiment described in Figure 3.3C. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

Figure S3.3

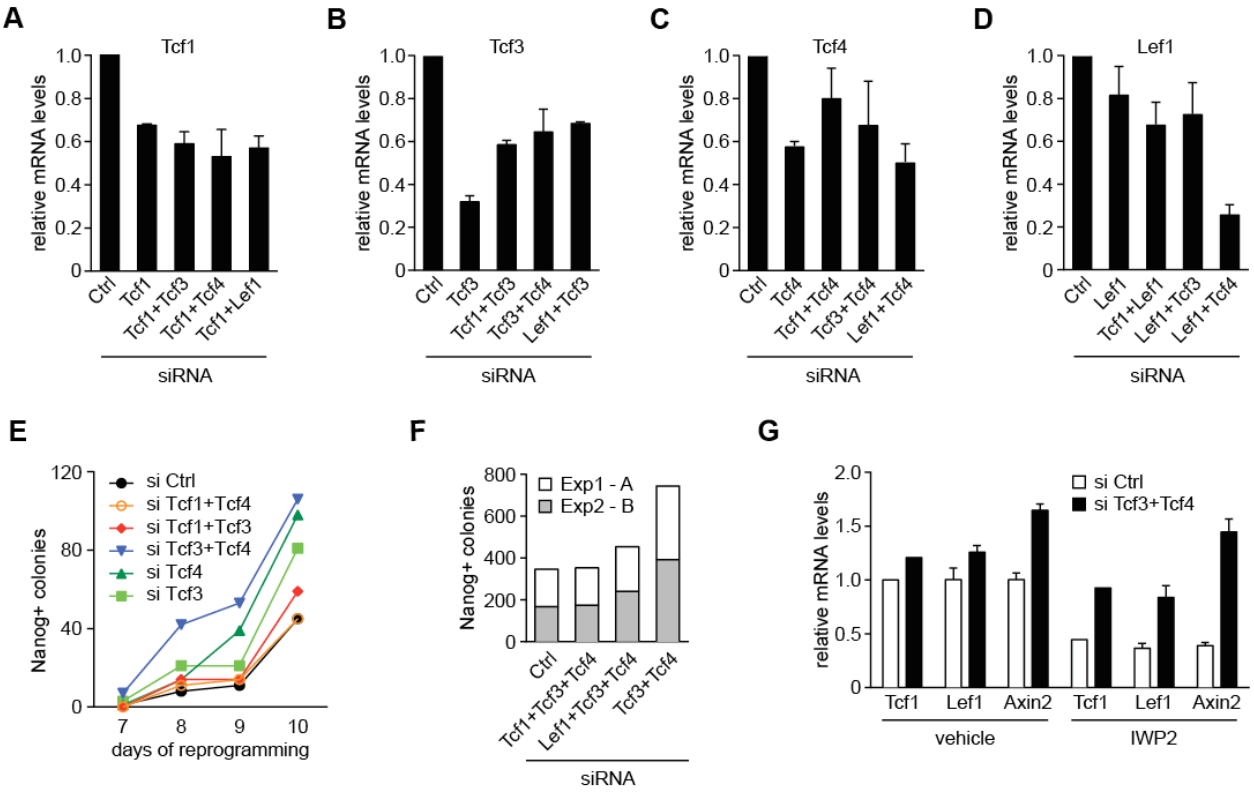

A
Tcf1
relative mRNA levels
1.0
0.8
0.6
0.4
0.2
0
Ctrl
Tcf1
Tcf1+Tcf3
Tcf1+Tcf4
Tcf1+Lef1
siRNA
B
Tcf3
relative mRNA levels
Ctrl
Tcf3
Tcf1+Tcf3
Tcf3+Tcf4
Lef1+Tcf3
siRNA
C
Tcf4
relative mRNA levels
Ctrl
Tcf4
Tcf1+Tcf4
Tcf3+Tcf4
Lef1+Tcf4
siRNA
D
Lef1
relative mRNA levels
Ctrl
Lef1
Tcf1+Lef1
Lef1+Tcf3
Lef1+Tcf4
siRNA
E
Nanog+ colonies
120
80
40
0
si Ctrl
si Tcf1+Tcf4
si Tcf1+Tcf3
si Tcf3+Tcf4
si Tcf4
si Tcf3
7
8
9
10
days of reprogramming
F
Nanog+ colonies
800
600
400
200
0
Exp1 - A
Exp2 - B
Ctrl
Tcf1+Tcf3+Tcf4
Lef1+Tcf3+Tcf4
Tcf3+Tcf4
siRNA
G
relative mRNA levels
2.0
1.5
1.0
0.5
0
si Ctrl
si Tcf3+Tcf4
Tcf1
Lef1
Axin2
Tcf1
Lef1
Axin2
vehicle
IWP2

Figure S3.4. Characterization of the Tcf3 Overexpression Effect on Reprogramming, Related to Figure 3.4

A. MEFs carrying the *Oct4-GFP* transgene and *M2rtTA* in the *R26* locus were infected with a dox-inducible retrovirus (pRetro) encoding *Tcf3* and subsequently with separate constitutive retroviruses (pMX) encoding*OSK*. Indicated amounts of dox were added 1 day postinfection to titrate *Tcf3* transcript levels. qPCR for *Tcf3*transcript levels relative to MEFs was performed 4 days post pMX-OSK infection and dox induction. Note that 0.002 µg/ml dox induces ESC-like levels of *Tcf3* and the *Tcf3* level in the 0 µg/ml dox sample is increased compared to MEFs, probably because *Tcf3* levels rise slightly early in reprogramming (data not shown). Values represent the average of triplicate sampling and error bars indicate standard deviation. The asterisks indicate the same data displayed in Figure 3.7B.

B. The reprogramming experiment described in (A) was quantified for *Oct4*-GFP positive colonies at day 14, demonstrating that Tcf3 inhibits the induction of iPSCs in a dose-dependent manner. A representative experiment is shown with average counts of three technical replicates with error bars indicating standard deviation.

C. iOSCK MEFs were transduced retrovirally (with the pMX virus) to express *Tomato* (Ctrl) or *Tcf3* and treated with dox to express the reprogramming factors, as described in Figure 3.4A. qPCR analysis showing the transcript levels of Tcf3 relative to MEFs throughout the reprogramming time course. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

D. As in (C), except that *Nanog* transcript levels were determined.

E. As in (C), expect that *Esrrb* transcript levels were determined.

F. All Tcf3 mutants used in Figure 3.4F were expressed well in iOSCK MEFs and showed nuclear localization as detected by immunostaining for Tcf3 (green) at day four of reprogramming. DAPI staining marks nuclei.

G. As in (F), except that cells were analyzed by western blotting for Tcf3 (top) or Oct4 (bottom) expression. Ectopic Oct4 is induced in all cells and serves as a loading control. Tcf3 transgenes migrate differently in the gel due to their varying sizes upon domain deletion.

Figure S3.4

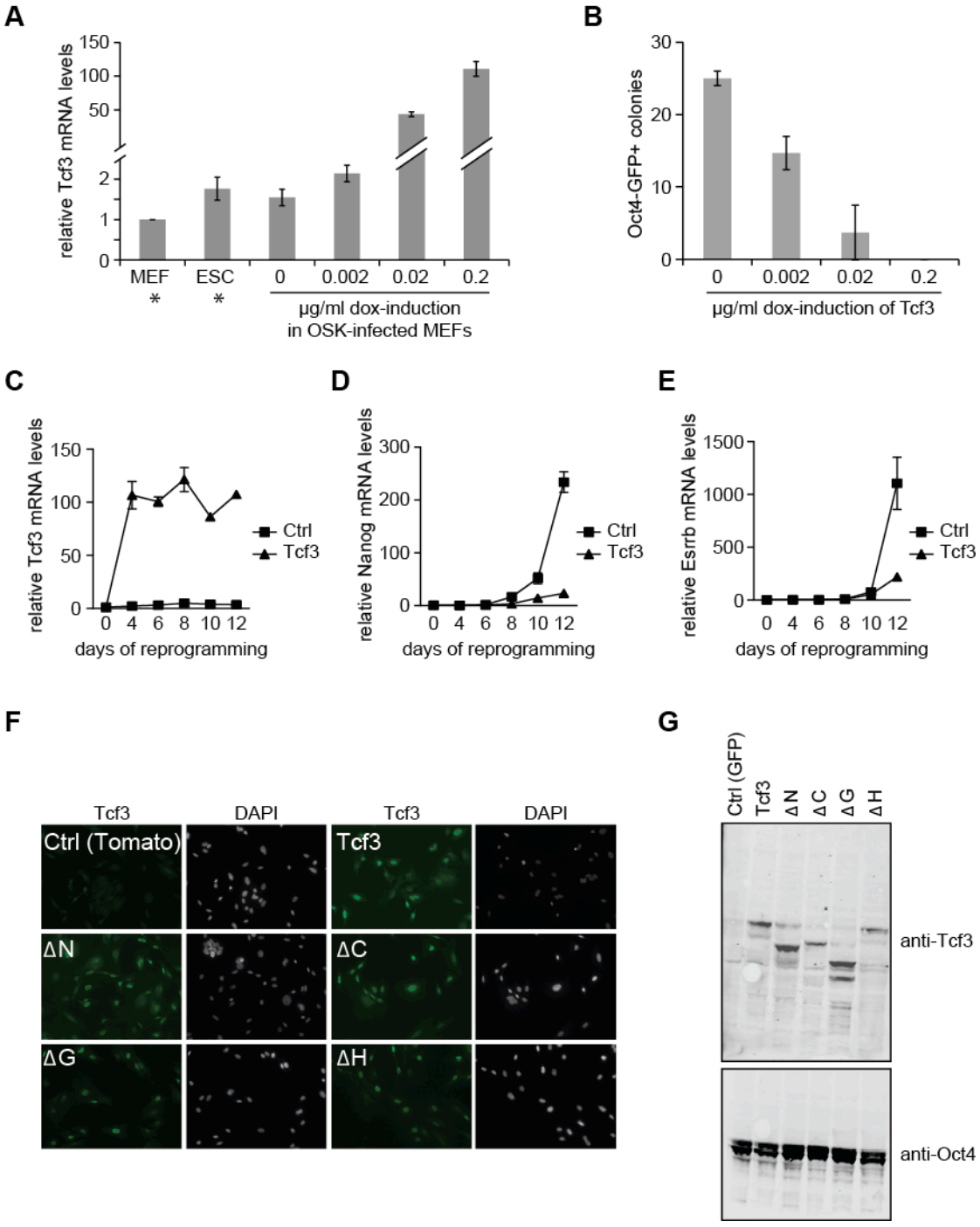

A
relative Tcf3 mRNA levels
150
100
50
2
1
0
MEF
ESC
0
0.002
0.02
0.2
*
*
μg/ml dox-induction
in OSK-infected MEFs
B
Oct4-GFP+ colonies
30
20
10
0
0
0.002
0.02
0.2
μg/ml dox-induction of Tcf3
C
relative Tcf3 mRNA levels
150
100
50
0
0 4 6 8 10 12
days of reprogramming
Ctrl
Tcf3
D
relative Nanog mRNA levels
300
200
100
0
0 4 6 8 10 12
days of reprogramming
Ctrl
Tcf3
E
relative Esrrb mRNA levels
1500
1000
500
0
0 4 6 8 10 12
days of reprogramming
Ctrl
Tcf3
F
Tcf3
DAPI
Tcf3
DAPI
Ctrl (Tomato)
Tcf3
ΔN
ΔC
ΔG
ΔH
G
Ctrl (GFP)
Tcf3
ΔN
ΔC
ΔG
ΔH
anti-Tcf3
anti-Oct4

Figure S3.5. Characterization of Reprogramming in the Complete Absence of *Tcf3*, Related to Figure 3.5

A. *Tcf32loxp/2loxp*;Tg(*UBC-Cre-ERT2*) MEFs also carrying the *Oct4-GFP* transgene were transduced with separate retroviruses encoding OSK and treated with and without tamoxifen, respectively, at day four to delete*Tcf3*. *Oct4*-GFP positive colonies were quantified at day 16 of reprogramming. A representative experiment is shown with average counts of three technical replicates with error bars indicating standard deviation.

B. Summary of Nanog counts from different reprogramming experiments that determined the effect of *Tcf3*ablation in OSK or OSCK reprogramming in different ESC media condition (FBS or KSR). In all these experiments, Tcf3 was always deleted at day four by tamoxifen addition. These data highlight a consistent 2-fold enhancement of reprogramming by Tcf3 ablation with and without cMyc and in FBS or KSR.

C. MEFs were treated with tamoxifen (+tam) to delete *Tcf3* or with vehicle (-tam) and cell counts were performed up to 144 hr after treatment for two independent experiments. Note that different amounts of cells were initially plated in the two experiments, which move the curves along the *y* axis. These data demonstrate that *Tcf3* ablation does not alter the proliferation rate of fibroblasts.

D. As in (C), except that cell counts were performed during OSK reprogramming. Again, there was no difference in proliferation rate due to *Tcf3* deletion.

E. *Tcf32loxp/2loxp*;Tg(*UBC-Cre-ERT2*) MEFs were transduced with separate retroviruses encoding OSK and treated with and without tamoxifen, respectively, at day zero or day four to delete *Tcf3*. Nanog positive colonies were quantified at day 17. The enhancing effect of *Tcf3* ablation is less pronounced when deletion is performed at day zero.

F. Summary of two independent reprogramming experiments that determined the effect of *Tcf3* ablation on OCK reprogramming. The knockout of *Tcf3* was performed at the indicated days and Nanog positive colonies were assessed by immunostaining for Nanog (IF) in experiment 1 or based on *Oct4*-GFP reporter expression in experiment 2 on the indicated day of reprogramming. These data demonstrate that *Tcf3* ablation causes a subtle enhancement of OCK reprogramming.

G. *Tcf32loxp/2loxp*;Tg(*UBC-Cre-ERT2*) MEFs were transduced with separate retroviruses encoding OCK and treated with and without tamoxifen, respectively, at day four of reprogramming to delete *Tcf3*. This experiment was performed both in FBS-containing and KSR-containing media. 24 colonies were isolated from the indicated reprogramming cultures at day 30, expanded, and immunostained for Nanog. The graph indicates the percentage of picked clones expressing Nanog, and inset numbers indicate the quantity of clones analyzed. This experiment demonstrated that expanding and passaging of OCK colonies results in efficient conversion to the Nanog positive stage but only in the absence of *Tcf3*. All cell lines obtained from tamoxifen-treated cultures were confirmed to be *Tcf3−/−* by PCR genotyping (data not shown).

H. qPCR for ectopic expression levels from retroviral integrations of the *Oct4* and *Klf4* transgenes in *Tcf3* KO OCK iPSC clones, normalized to *Gapdh*. OSCK pre-iPSCs serve as the reference sample as they express high levels of retroviral reprogramming factors (Sridharan et al., 2009). OCK *Tcf3* KO iPSCs have silenced the reprogramming factors, consistent with the idea that they have faithfully established the self-renewing,

pluripotent state. Values represent the average of triplicate analysis and error bars standard deviation.

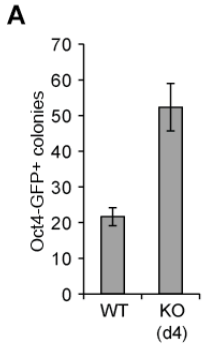
I. Confirmation of *Tcf3* deletion and absence of retroviral *Sox2* in *Tcf3* KO-OCK iPSC lines. All *Tcf3* KO OCK iPSC lines analyzed were deleted for *Tcf3* (and therefore are 1loxp) and carried integration of retroviral *Oct4*,*Klf4*, and *cMyc* but not *Sox2*. An OSCK iPSC line obtained from *Tcf32loxp/2loxp*;Tg(*UBC-Cre-ERT2*) MEFs without tamoxifen treatment served as a positive control for the retroviral *Sox2* PCR and the 2loxp *Tcf3* alleles.

J. Pearson correlation coefficients for log2 expression values of all probes between MEFs, *Tcf3* WT and *Tcf3*KO OSK and OCK iPSC lines, OCK-pre-iPSCs, and WT iPSCs and ESCs. The correlation values indicate the similarity between ESCs and all iPSCs. For these expression profiles, ESCs, previously generated iPSCs, and MEFs were cultures in FBS-containing media, while the Tcf3 KO and WT iPSCs were cultured in KSR containing media, likely contributing to some expression differences seen between ESCs and Tcf3 WT and KO iPSCs.

K. Subcutaneous injections of *Tcf3* KO and WT iPSCs into NOD/SCID mice produced teratomas exhibiting histological structures representative of all three germ layers.

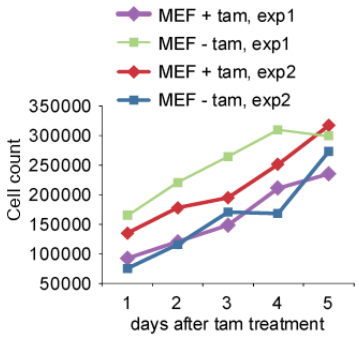
L. Expression differences between *Tcf3* KO and *Tcf3* WT OSK iPSCs are strongly enriched for genes upregulated (top panel) and downregulated (bottom panel) upon *Tcf3* deletion in ESCs, as determined by GSEA analysis (Mootha et al., 2003; Subramanian et al., 2005). The heatmap from red (upregulated) to blue (downregulated) depicts the ranked expression changes between *Tcf3* KO and WT iPSCs. Black lines above the heat map mark all genes that are at least 2-fold upregulated (top) or downregulated (bottom) in response to*Tcf3* knockout in ESCs (Yi et al., 2008). The green lines indicate the enrichment score and the p value measures the family-wise error rate (FWER).

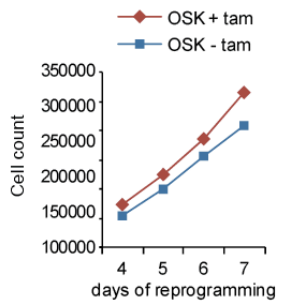
Figure S3.5

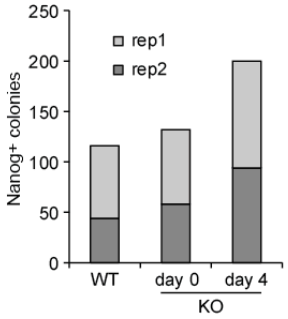


B

Media	Experiment	Day	# Nanog+ colonies		
			Tcf3 wt	Tcf3 ko	ko/wt
FBS	OSK-1	20	118	238	2.02
	OSK-1	26	308	512	1.66
	OSK-2	19	11	21	1.91
	OSK-2	25	33	69	2.09
	OSCK-2	19	46	94	2.04
	OSCK-2	25	118	198	1.68
KSR	OSK-3	13	47	118	2.51
	OSK-3	17	98	180	1.84
	OSCK-3	13	158	330	2.09
	OSCK-3	17	180	356	1.98

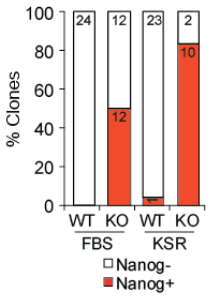


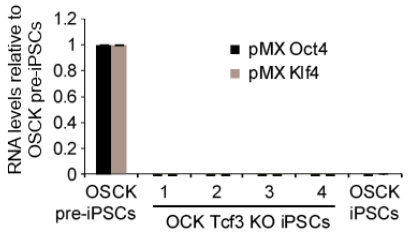


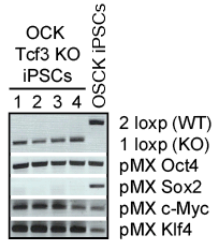


F

Exp.	OCK + Tcf3 WT or KO	# Nanog+ colonies	Day
1	WT	0 (IF)	32
	KO day 4	0 (IF)	32
	KO day 8	0 (IF)	32
	KO day 11	1 (IF)	32
	KO day 18	3 (IF)	32
2	WT	0 (GFP)	35
	KO day 4	2 (GFP)	35
	KO day 19	2 (GFP)	35

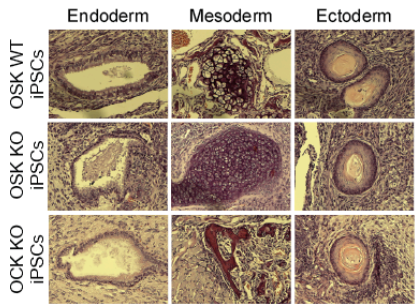






J

	MEF	OCK pre-iPSC	OCK iPSC KO 1	OCK iPSC KO 2	OSK iPSC WT	OSK iPSC KO 1	OSK iPSC KO 2	iPSC	ESC
MEF	1								
OCK pre-iPSC	0.892	1							
OCK iPSC KO 1	0.839	0.871	1						
OCK iPSC KO 2	0.854	0.883	0.988	1					
OSK iPSC WT	0.859	0.884	0.979	0.985	1				
OSK iPSC KO 1	0.846	0.879	0.983	0.987	0.992	1			
OSK iPSC KO 2	0.845	0.879	0.982	0.986	0.992	0.998	1		
iPSC	0.879	0.879	0.959	0.969	0.964	0.961	0.960	1	
ESC	0.886	0.874	0.951	0.963	0.971	0.967	0.966	0.977	1



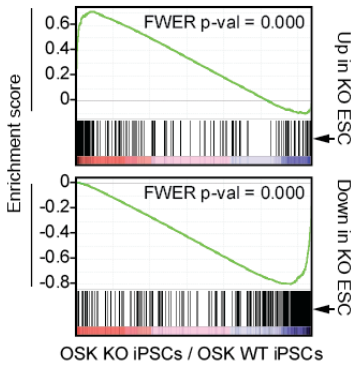


Figure S3.6. Characterization of *Tcf3*-Dependent Expression Changes in OCK Reprogramming, Related to Figure 3.6

A. qPCR showing the mRNA levels of *Tcf1* relative to siCtrl in OCK pre-iPSCs carrying the *Nanog*-GFP reporter, transfected with siRNAs against *Tcf1* for the experiment shown in Figures 3.6A and 3.6B. Analysis was performed 3 days after the initial siRNA transfection. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

B. As in (A), except for *Tcf3*.

C. As in (A), except for *Tcf4*.

D. qPCR showing the mRNA levels for the Wnt-regulated genes *Lef1*, *Tcf1*, and A*xin2* relative to siCtrl 3 days after the initial *Tcf3/Tcf4* siRNA transfection in OCK pre-iPSCs. The average of duplicate qPCR measurements of one experiment is shown, with error bars indicating standard deviation.

E. Representative bright-field (left panels) and GFP fluorescence (right panels) images of OCK pre-iPSCs carrying the *Nanog*-GFP reporter, 4 days after siRNA transfections. Note that cell number was similar under all conditions (data not shown). The appearance of GFP positive colonies depended on *Tcf3* and/or *Tcf4* depletion and was reduced upon *Tcf1* codepletion.

F. FACS analysis of the OCK pre-iPSC reprogramming experiment. The percentage of SSEA1 positive and SSEA1/*Nanog*-GFP double positive cells was determined 96 hr post-siRNA transfection of OCK pre-iPSCs. Duplicate wells were measured by FACS and the average of the two technical replicates wells are shown with error bars indicating standard deviation. Note that the overall number of SSEA1 positive cells was largely unaffected by the knockdowns but that the number of SSEA1/GFP positive cells varied strongly.

G. qPCR for transcript levels of indicated genes relative to levels in MEFs at various time points throughout the OCK reprogramming experiment described in Figures 3.6C–E, with (KO) or without (WT) tamoxifen treatment at day four to delete *Tcf3*. Relative expression levels in *Tcf3* WT and *Tcf3* KO OSK iPSCs are given to the right of each panel. Values represent the average of triplicate analysis and error bars indicate standard deviation where presented.

H. Significantly enriched GO terms for genes downregulated at least 1.5-fold in OCK pre-iPSCs 48 hr after siRNA-mediated *Tcf3* knockdown.

I. qPCR for transcript levels relative to levels in MEFs of various genes at the indicated days of the OCK reprogramming experiment described in Figures 3.6C–E for *Tcf3* WT and *Tcf3* KO OCK samples. Relative expression levels in WT OSK iPSCs are given. Values represent the average of triplicate analysis and error bars indicate standard deviation.

Figure S3.6

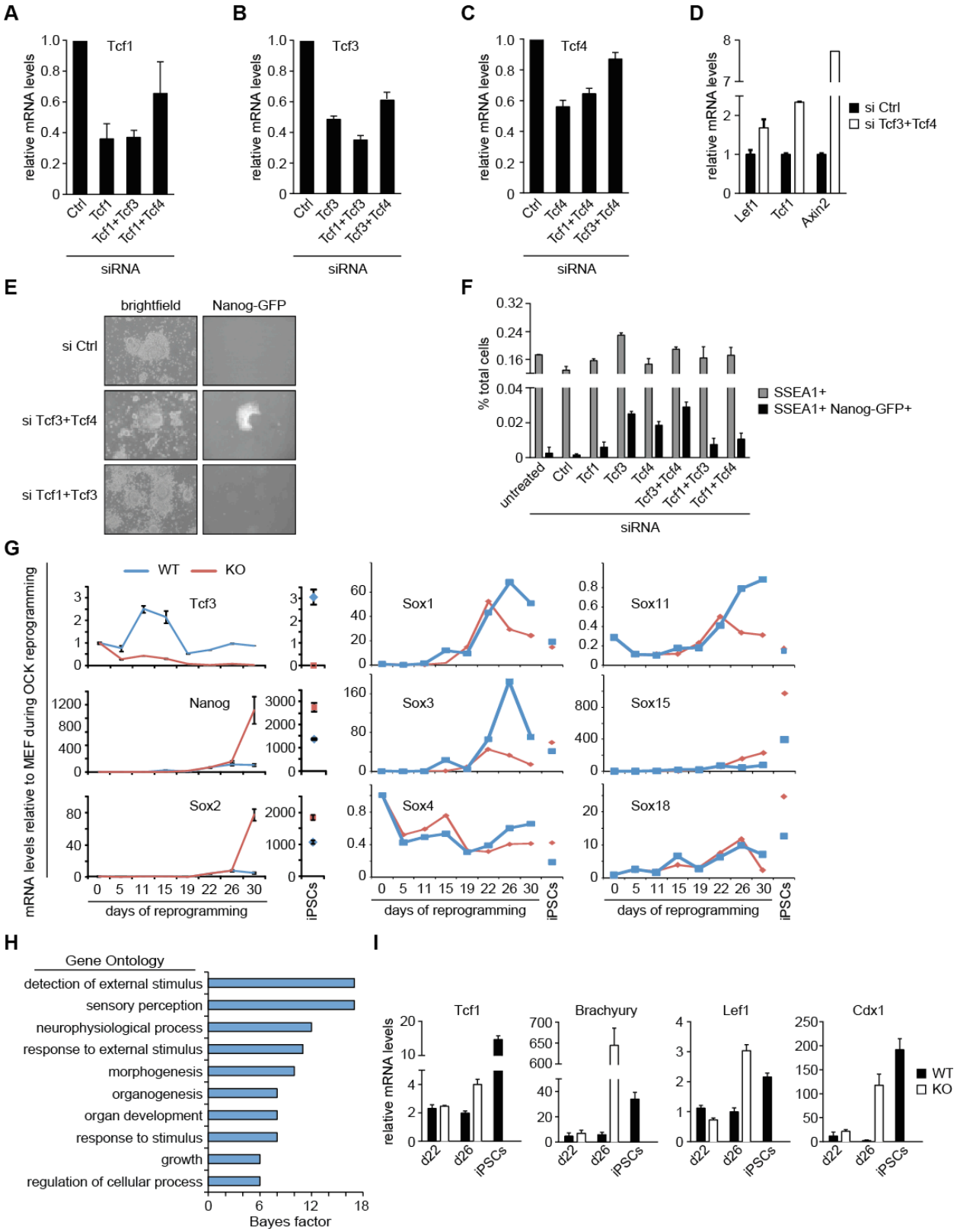

A
Tcf1
relative mRNA levels
Ctrl
Tcf1
Tcf1+Tcf3
Tcf1+Tcf4
siRNA
B
Tcf3
Ctrl
Tcf3
Tcf1+Tcf3
Tcf3+Tcf4
C
Tcf4
Ctrl
Tcf4
Tcf1+Tcf4
Tcf3+Tcf4
D
si Ctrl
si Tcf3+Tcf4
Lef1
Tcf1
Axin2
E
brightfield
Nanog-GFP
si Ctrl
si Tcf3+Tcf4
si Tcf1+Tcf3
F
% total cells
SSEA1+
SSEA1+ Nanog-GFP+
untreated
Ctrl
Tcf1
Tcf3
Tcf4
Tcf3+Tcf4
Tcf1+Tcf3
Tcf1+Tcf4
siRNA
G
WT
KO
mRNA levels relative to MEF during OCK reprogramming
Tcf3
Nanog
Sox2
Sox1
Sox3
Sox4
Sox11
Sox15
Sox18
days of reprogramming
iPSCs
H
Gene Ontology
detection of external stimulus
sensory perception
neurophysiological process
response to external stimulus
morphogenesis
organogenesis
organ development
response to stimulus
growth
regulation of cellular process
Bayes factor
I
relative mRNA levels
Tcf1
Brachyury
Lef1
Cdx1
d22
d26
iPSCs
WT
KO

Figure S3.7. Characterization of OCK iPSCs Obtained by Stepwise Modulation of *Tcf3* Levels, Related to Figure 3.7

A. *Tcf3* levels were titrated early in OCK reprogramming using a dox-inducible *Tcf3* retrovirus, as described in Figures 3.7A and 3.7B. Elevated *Tcf3* levels result in slightly more three-dimensional ESC-like colonies in OCK reprogramming by day eight, as shown by bright field images of representative colonies (top panels). Crystal violet staining of culture wells at day eight displays colony size and number (bottom panels). 0.02 μg/ml dox-induction of *Tcf3* levels enhances the size and number of colonies by day eight, but the lower dox concentration (0.002 μg/ml), which yielded ESC-like levels of *Tcf3*, does not affect colony size or number.

B. PCR on genomic DNA for integrations retroviral *Oct4* and *Sox2*. The three OCK iPSC lines (A, B, C) derived from the 0.002 μg/ml dox (early), si*Tcf3* (late) condition from the experiment shown in Figure 3.7C carry the retroviral *Oct4* vector, but not the *Sox2* vector, confirming the absence of ectopic *Sox2* in reprogramming. An OSCK pre-iPSC line served as a positive control for the PCRs.

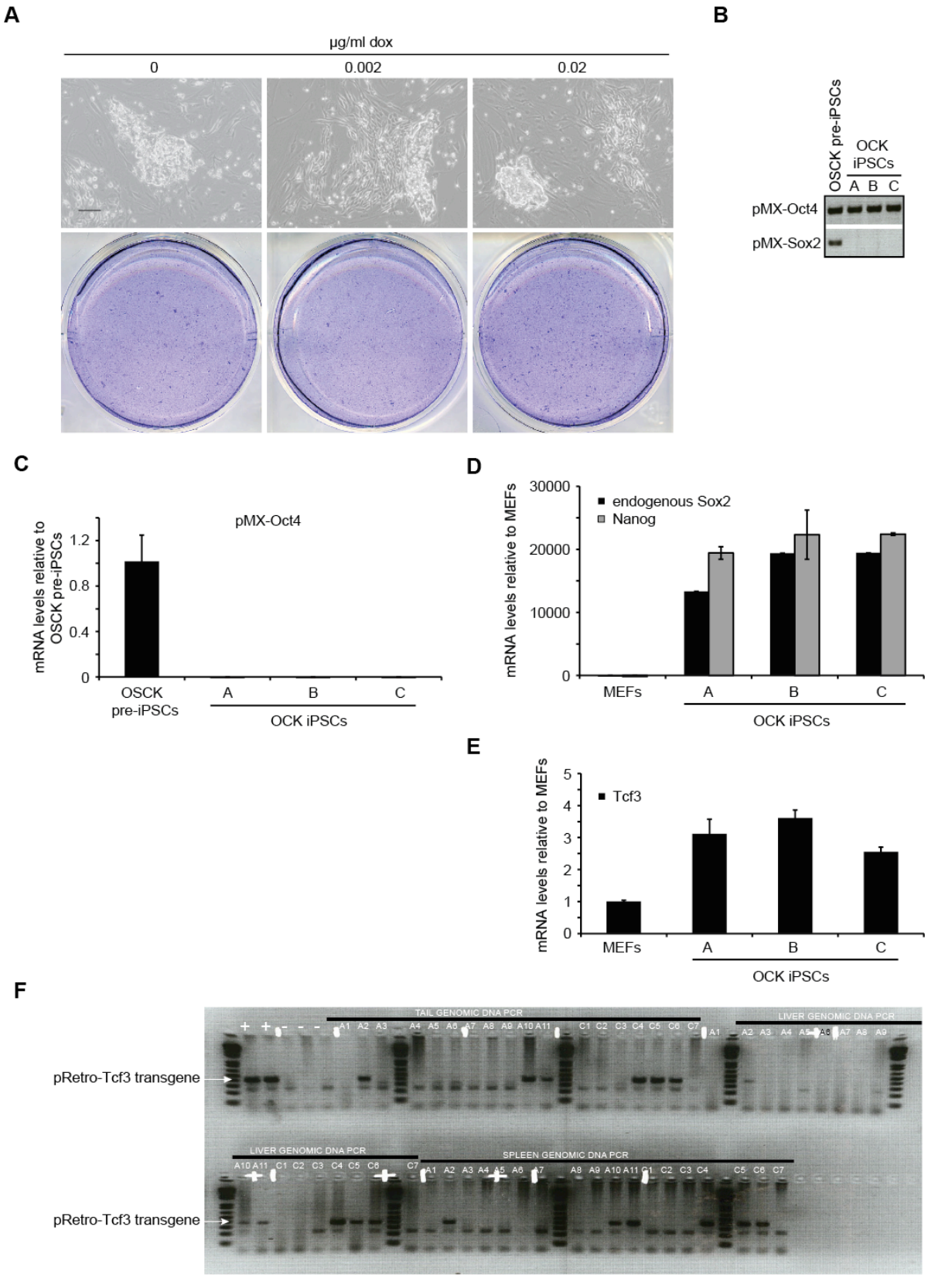
C. qPCR for transcript levels from the retroviral *Oct4* transgene (pMX-*Oct4*) in OCK iPSC lines A–C. OSCK pre-iPSCs serve as the reference sample, as they express high levels of retroviral reprogramming factors. Values represent the average of triplicate analysis and error bars indicate standard deviation.

D. qPCR shows transcript levels relative to levels in MEFs of the indicated genes in the three OCK iPSC lines (A, B, C). Note that endogenous *Sox2* levels are elevated in iPSCs, despite the lack of retrovirally integrated *Sox2* transgenes. Values represent the average of triplicate analysis and error bars indicate standard deviation.

E. As in (D), except for *Tcf3* transcript levels, which are 2-fold higher in iPSCs than MEFs, which is similar to the ratio between ESCs and MEFs shown in Figure 3.7B.

F. PCR for genomic integration of the inducible *Tcf3* retrovirus on DNA isolated from liver, spleen, and tails from chimeric mice obtained upon blastocyst injection of OCK iPSC lines A and C. The positive control (+) depicts genomic DNA isolated from OCK iPSCs A and C and the negative control (−) genomic DNA isolated from the WT mothers. Source data for Figure 3.7D.

Figure S3.7


A
μg/ml dox
0
0.002
0.02
B
OSCK pre-iPSCs
OCK iPSCs
A B C
pMX-Oct4
pMX-Sox2
C
pMX-Oct4
mRNA levels relative to OSCK pre-iPSCs
1.2
0.8
0.4
0
OSCK pre-iPSCs
A
B
C
OCK iPSCs
D
endogenous Sox2
Nanog
mRNA levels relative to MEFs
30000
20000
10000
0
MEFs
A
B
C
OCK iPSCs
E
Tcf3
mRNA levels relative to MEFs
5
4
3
2
1
0
MEFs
A
B
C
OCK iPSCs
F
TAIL GENOMIC DNA PCR
LIVER GENOMIC DNA PCR
pRetro-Tcf3 transgene
LIVER GENOMIC DNA PCR
SPLEEN GENOMIC DNA PCR
pRetro-Tcf3 transgene

REFERENCES

Aubert, J., Dunstan, H., Chambers, I., and Smith, A. (2002). Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation. Nature biotechnology *20*, 1240-1245.

Cole, M.F., Johnstone, S.E., Newman, J.J., Kagey, M.H., and Young, R.A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes & Development *22*, 746-755.

Ernst, J., and Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. BMC Bioinformatics *7*, 191.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S.*, et al.* (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell *6*, 71-79.

Galceran, J., Farinas, I., Depew, M.J., Clevers, H., and Grosschedl, R. (1999). Wnt3a-/--like phenotype and limb deficiency in Lef1(-/-)Tcf1(-/-) mice. Genes & Development *13*, 709-717.

Han, J., Yuan, P., Yang, H., Zhang, J., Soh, B.S., Li, P., Lim, S.L., Cao, S., Tay, J., Orlov, Y.L.*, et al.* (2010). Tbx3 improves the germ-line competency of induced pluripotent stem cells. Nature *463*, 1096-1100.

Lluis, F., Ombrato, L., Pedone, E., Pepe, S., Merrill, B.J., and Cosma, M.P. (2011). T-cell factor 3 (Tcf3) deletion increases somatic cell reprogramming by inducing epigenome modifications. Proceedings of the National Academy of Sciences of the United States of America *108*, 11912-11917.

Lluis, F., Pedone, E., Pepe, S., and Cosma, M.P. (2008). Periodic activation of Wnt/beta-catenin signaling enhances somatic cell reprogramming mediated by cell fusion. Cell Stem Cell *3*, 493-507.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55-70.

Mao, B., Wu, W., Li, Y., Hoppe, D., Stannek, P., Glinka, A., and Niehrs, C. (2001). LDL-receptor-related protein 6 is a receptor for Dickkopf proteins. Nature *411*, 321-325.

Marson, A., Foreman, R., Chevalier, B., Bilodeau, S., Kahn, M., Young, R.A., and Jaenisch, R. (2008a). Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell *3*, 132-135.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J.*, et al.* (2008b). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Gottgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491-504.

Merrill, B.J., Gat, U., DasGupta, R., and Fuchs, E. (2001). Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. Genes & Development *15*, 1688-1705.

Merrill, B.J., Pasolli, H.A., Polak, L., Rendl, M., Garcia-Garcia, M.J., Anderson, K.V., and Fuchs, E. (2004). Tcf3: a transcriptional regulator of axis induction in the early embryo. Development *131*, 263-274.

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.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nature Biotechnology *26*, 101-106.

Nguyen, H., Merrill, B.J., Polak, L., Nikolova, M., Rendl, M., Shaver, T.M., Pasolli, H.A., and Fuchs, E. (2009). Tcf3 and Tcf4 are essential for long-term homeostasis of skin epithelia. Nature Genetics *41*, 1068-1075.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313-317.

Papp, B., and Plath, K. (2013). Epigenetics of Reprogramming to Induced Pluripotency. Cell *152*, 1324-1343.

Pereira, L., Yi, F., and Merrill, B.J. (2006). Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. Molecular and Cellular Biology *26*, 7479-7491.

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.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells *27*, 543-549.

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.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes & Development *24*, 2239-2263.

Stadtfeld, M., Maherali, N., Borkent, M., and Hochedlinger, K. (2010). A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nature Methods *7*, 53-55.

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230-240.

Szabo, P.E., Hubner, K., Scholer, H., and Mann, J.R. (2002). Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. Mechanisms of Development *115*, 157-160.

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.

Takemaru, K.I., and Moon, R.T. (2000). The transcriptional coactivator CBP interacts with beta-catenin to activate gene expression. The Journal of Cell Biology *149*, 249-254.

Tam, W.L., Lim, C.Y., Han, J., Zhang, J., Ang, Y.S., Ng, H.H., Yang, H., and Lim, B. (2008). T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. Stem Cells *26*, 2019-2031.

ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nature Cell Biology *13*, 1070-1075.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.

Wu, C.I., Hoffman, J.A., Shy, B.R., Ford, E.M., Fuchs, E., Nguyen, H., and Merrill, B.J. (2012). Function of Wnt/beta-catenin in counteracting Tcf3 repression through the Tcf3-beta-catenin interaction. Development *139*, 2118-2129.

Yi, F., Pereira, L., Hoffman, J.A., Shy, B.R., Yuen, C.M., Liu, D.R., and Merrill, B.J. (2011). Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. Nature Cell Biology *13*, 762-770.

Yi, F., Pereira, L., and Merrill, B.J. (2008). Tcf3 functions as a steady-state limiter of transcriptional programs of mouse embryonic stem cell self-renewal. Stem Cells *26*, 1951-1960.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.

Yoshikawa, Y., Fujimori, T., McMahon, A.P., and Takada, S. (1997). Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse. Developmental Biology *183*, 234-242.

Zhang, X., Peterson, K.A., Liu, X.S., McMahon, A.P., and Ohba, S. (2013). Gene Regulatory Networks Mediating Canonical Wnt Signal Directed Control of Pluripotency and Differentiation in Embryo Stem Cells. Stem Cells.

CHAPTER 4

Regulation of Reprogramming to Induced Pluripotency by Polycomb Repressive Complexes

ABSTRACT

The Polycomb repressive complexes (PRC1 and PRC2) cooperatively regulate transcriptional silencing in embryonic stem cells (ESCs), particularly of genes involved in early embryonic development. Histone H3K27me3 and H2AK119Ub, mediated by individual Polycomb group proteins, are important to maintain the pluripotent state. As global changes in histone marks occur during the reprogramming of somatic cells to induced pluripotent stem cells (iPSCs), the role of the PRCs remain uncharacterized. Here, we describe the effects of inactivating individual components of the PRCs in the reprogramming process. We demonstrate that Ring1b, the H2AK119 ubiquitinylase of PRC1, is dispensable for reprogramming, despite its important role in regulating ESC pluripotency and early development. In contrast, the essential subunit of PRC2, Eed is required throughout the reprogramming process. Interestingly, the H3K27 methyltransferase of PRC2, Ezh2 is required for the early stage of reprogramming, but dispensable during later stages, suggesting a requirement for an alternative PRC2 incorporating the methyltransferase Ezh1 instead of Ezh2. Nevertheless, we are able to derive iPSCs generated in the absence of catalytically active Ezh2, and while we observe a global decrease in H3K27me3 genomic binding in these cells, this mark is retained at a minor group of loci. Together, our data uncouple the requirements for PRC proteins in early development from the acquisition of pluripotency during reprogramming.

INTRODUCTION

Somatic cells undergo genome-wide changes in chromatin state when they are induced to become pluripotent stem cells (iPSCs) by over expression of the transcription factors Oct4, Sox2, cMyc, and Klf4 (Chin et al., 2009; Maherali et al., 2007; Mikkelsen et al., 2007). Chromatin changes accompany the silencing of somatic genes and the activation of pluripotency genes. One of the most well-characterized histone modifications that changes its locations during the reprogramming process is histone H3 tri-methylated lysine 27 (H3K27me3),

a mark associated with transcriptional silencing. The transcriptional repressors responsible for depositing the H3K27me3 mark are the proteins comprising the Polycomb repressive complex 2 (PRC2). This complex is composed of the core subunits Embryonic ectoderm development (Eed), Suppressor of zeste 12 (Suz12), and one of the histone methyltransferases Enhancer of zeste (Ezh1) or Ezh2, which catalyze the transfer of methyl groups to the lysine tail of histone H3 through their SET domains (Cao et al., 2002; Shen et al., 2008).

At a subset of PRC2 targets in embryonic stem cells (ESCs), the established H3K27me3 mark subsequently recruits the Polycomb repressive complex 1 (PRC1), which mono-ubiquitinylates histone H2A on lysine 119 (H2AK119Ub1) through its catalytic subunit, Ring finger protein 1b (Ring1b), and to a lesser extent, Ring1a (de Napoles et al., 2004). This process is thought to mediate transcriptional silencing by inhibiting RNA Polymerase II elongation (Stock et al., 2007; Zhou et al., 2008). The location profile of H3K27me3 in ESCs is distinct from that in somatic cells: in pluripotent cells it is associated with genes regulating early embryonic development, which are repressed in these cells (Surface et al., 2010); whereas in somatic cells, genes specific to the cell type are not repressed, and the genes supporting the pluripotency network are typically H3K27me3-positive accompanying their repression.

As mentioned above, the profile of H3K27me3 changes dramatically during the reprogramming process, in line with expression changes of the corresponding genes (Maherali et al., 2007; Mikkelsen et al., 2007). It was recently shown that the demethylation of H3K27 by Ubiquitously transcribed tetratricopeptide repeat, X chromosome (Utx) at pluripotency genes is required for the generation of iPSCs (Mansour et al., 2012). However, it is unclear whether the reprogramming process requires a gain in H3K27me3 at developmental genes that are active in the somatic starting cells and repressed in the pluripotent state.

In addition to genes regulating development, genes regulating the cell cycle are also differentially expressed between ESCs and somatic cells. ESCs display a cell cycle structure distinct from somatic cells, characterized by a short G1 phase due to high levels of Cyclin E, which stimulates E2F activation of S-phase entry genes (Stead et al., 2002). As pluripotent cells progress to a more differentiated state, they reduce Cyclin E levels, become subject to Cyclin D regulation of the G1 checkpoint, and are susceptible to senescence through Arf and Ink4a upregulation, proteins which inhibit Cyclin E and Cyclin D, respectively (He et al., 2009; Savatier et al., 1996; Stead et al., 2002). Previous reprogramming studies reported that the expression of Arf and Ink4a inhibit reprogramming, and their downregulation improves the efficiency of reprogramming (Li et al., 2009; Utikal et al., 2009). Notably, the transcription of several of these senescence regulators is under the direct influence of PRC2 in several cells types, including somatic, stem, and cancer cells (Sauvageau and Sauvageau, 2010). However, it remains unclear whether the regulation of these genes by PRC2 and H3K27me3 plays a critical role during the reprogramming process.

The PRC1 and PRC2 subunits that catalyze the histone modifications involved with gene silencing in ESCs have been characterized in several studies (Bernstein et al., 2006; Boyer et al., 2006; Stock et al., 2007). Importantly, Ezh2 null ESCs can be established and maintain self-renewal but are affected in their differentiation capabilities (Shen et al., 2008). Notably, Ezh2 is part of the PRC2 Polycomb repressive complex that is responsible for establishing global H3K27me2/3 in ESCs. In its absence, the alternative PRC2 complex containing the catalytic protein Ezh1 is able to maintain H3K27me2/3 at developmentally regulated PRC2 target genes (Shen et al., 2008). The ability to methylate H3K27 by either Ezh2 or Ezh1 is dependent on another PRC2 subunit, Eed, and ablation of Eed in ESCs results in the complete erasure of all forms of H3K27 methylation (Montgomery et al., 2005). However, even Eed knockout ESCs can be generated, indicating the H3K27me3 is not required for the maintenance of the self-renewing

state but for subsequent differentiation processes. Notably, many of the developmental genes enriched for H3K27me3 in ESCs are not upregulated dramatically in the undifferentiated state, likely because the transcriptional activators for these genes are not yet expressed (Boyer et al., 2006). Regarding PRC1, Ring1b null ESCs are unstable and tend to easily differentiate (Leeb and Wutz, 2007), and Ring1b/Ring1a double knockout ESCs halt proliferation and upregulate PRC1 target genes (Endoh et al., 2008). These findings highlight a dependence on the H2AK119 ubiquitinylation activity of PRC1 to maintain ESC self-renewal.

Given the differential effects PRC protein depletion has in ESCs, it remains unclear if they are necessary for the reprogramming process. Here, we demonstrate that Ring1b surprisingly is dispensable for reprogramming, despite its critical role in embryonic development. In contrast, cells lacking a catalytically active Ezh2 gene, and thereby the ability to maintain and establish global H3K27me3, reprogram with dramatically reduced efficiency. Specifically, we observe that the methyltransferase activity of Ezh2 is required early on in reprogramming, at least in part, to prevent the upregulation of its previously characterized targets, the senescence effectors Arf and Ink4a (Sauvageau and Sauvageau, 2010). Nevertheless, we were able to establish a few iPSC lines generated in the absence of catalytically active Ezh2, and, observed global differences in H3K27me3 profiles compared to iPSC lines generated in the presence of wild-type Ezh2. Ezh2-deficient iPSCs retain H3K27me3, likely established by the redundant activity of Ezh1, at loci that are associated with regulators of early embryonic development, such as *Hox* genes. Interestingly, whereas the Eed subunit of PRC2 is necessary throughout the entire reprogramming process, the Ezh2 methyltransferase activity becomes dispensable as reprogramming proceeds into later stages, suggesting that the function of Ezh1 is important late in the reprogramming process. Taken together, this study begins to dissect the timing and necessity of PRC2 and PRC1 components during the reprogramming process and aims to reveal critical H3K27me3 targets during the reprogramming process.

RESULTS

Ring1b is not required for reprogramming

Given that Ring1a null mice have no lethal embryonic phenotype (del Mar Lorente et al., 2000), but Ring1b is required for differentiation of ESCs and the development of the early embryo (Leeb and Wutz, 2007; Voncken et al., 2003), we tested if Ring1b was required for reprogramming mouse embryonic fibroblasts (MEFs) to the pluripotent state. Microarray expression analysis of MEFs, a line of partially reprogrammed intermediate cells known as pre-iPSCs (Sridharan et al., 2009), completely reprogrammed iPSCs, and ESCs indicates that between the chromatin-modifying subunits Ring1a and Ring1b of the Polycomb repression complex 1 (PRC1), Ring1b is highly expressed in pluripotent cells relative to non-pluripotent cells (Figure S4.1A). Using primary, transgenic MEFs carrying an IRES Cre-ER constitutively expressed from the 3' UTR of the PolII gene and homozygous for conditional 2lox Rnf2 alleles (Cales et al., 2008), the gene for Ring1b, we addressed whether Ring1b is required for reprogramming to pluripotency. Oct4, Sox2, and Klf4 (OSK) were retrovirally overexpressed in these MEFs, and deletion of Rnf2 was induced at different time points of the reprogramming process by treating cultures with 1μM 4-hydroxy-tamoxifen (4OHT) for 24 hours (Figure 4.1A). We confirmed that 4OHT treatment induced efficient excision of both Rnf2 alleles in the absence or presence of the ectopic reprogramming factors (Figure S4.1B). The appearance of ESC-like colonies was observed twelve days post-transduction, at which cultures were fixed, immunostained, and quantified for the number of Nanog-positive colonies, a marker for faithfully reprogrammed iPSCs (Figures 4.1A and 4.1B). Co-staining for Ring1b confirmed the efficient deletion of the Ring1b gene (Figures 4.1B and 4.1C). The quantification indicated that 4OHT treatment at any time point of reprogramming resulted in no significant reduction of the number of Nanog-positive colonies compared to the wild-type control experiment (Figures 4.1C), despite the lack of Ring1b protein in Nanog-positive colonies formed in 4OHT-treated cultures (Figure

4.1C). The same result was obtained when we analyzed the reprogramming cultures for the expression of another pluripotency marker, Esrrb (Figure S4.1C). We conclude that Ring1b is not required for reprogramming of somatic cells to pluripotency and that its absence does not alter the efficiency of reprogramming. This dispensability could potentially be due to a compensatory function by its mammalian paralogue Ring1a, which shares many gene targets of Ring1b in ESCs (Endoh et al., 2008). In the future, Ring1a/1b double knockout reprogramming experiments would test this hypothesis.

Catalytically active Ezh2 is critical for the early phase of reprogramming but becomes less important during later stages

We next addressed if the catalytic subunit of the major Polycomb repression complex 2 (PRC2), Ezh2, is necessary for reprogramming. To test this, we used primary, MEFs homozygous for Ezh2 alleles in which exons 3-6, which encode the catalytic SET domain mediating histone H3K27 methyltransferase activity, are flanked by loxP sites (Figure S4.2A). Adenoviral transduction of *Cre* enabled efficient excision of the SET domain from both Ezh2 alleles in fibroblasts with and without OSK transduction (Figures S4.2A and S4.2B). After transduction with OSK in these MEFs, deletion of the Ezh2 SET domain was induced at different timepoints of reprogramming to determine when Ezh2 catalytic activity was critical for the process (Figure 4.2A). qPCR of *Cre*-infected reprogramming cultures demonstrated that mRNA coding for the Ezh2 SET domain is efficiently reduced compared to control cultures (Figure S4.2C). Cultures were allowed to reprogram over a 15 day period, fixed, and analyzed for iPSC formation. Quantification of the entire reprogramming culture for the presence of alkaline phosphatase (AP) -positive colonies, a marker for reprogramming intermediates and pluripotent cells, indicated that deletion of the SET domain at day zero of reprogramming almost completely abolished the appearance of AP-positive colonies. SET domain deletion at day six showed a milder inhibitory effect and deletion at day twelve showed no effect (Figures 4.2B and

4.2C). In line with this observation, adenovirally transduced *Cre* at day zero of reprogramming almost completely abolished the formation of colonies expressing Nanog, which is indicative for faithful reprogramming to iPSCs (Figure 4.2D). Notably, SET domain deletion at day three led to a small number of Nanog-positive colonies, indicating that the catalytic activity of Ezh2 is absolutely critical before day three of reprogramming. Deletions at later time points gradually reduced the inhibitory effect; day six deletion resulted in 50% of Nanog-positive colonies compared to control, and deletions at day nine and twelve were indistinguishable from the control (Figure 4.2D). Notably, if the reprogramming cultures were allowed to grow until day 25, a few Nanog-positive colonies can be observed in cultures in which Ezh2 was inactivated at day zero, and even at two days prior to OSK transduction (data not shown). This suggests that while catalytic Ezh2 is not absolutely required for reprogramming, the kinetics of the process is dramatically affected. Altogether, these results suggested that the methyltransferase activity of Ezh2 promotes early reprogramming events, but gradually becomes dispensable during the later reprogramming stage.

To verify that the iPSCs appearing in *Cre*-transduced cultures did not arise from cells that evaded SET domain excision, we co-immunostained colonies from the day three and day nine *Cre*-infected cultures for another faithful iPSC marker, Esrrb, and the biochemical product of the Ezh2 methyltransferase activity, H3K27me3 (Figure 4.2E). This analysis of the arising colonies indicated a similar trend in the number of Esrrb positive colonies as for Nanog, and that the majority of Esrrb-positive colonies lacked detectable H3K27me3 in cultures treated with Ad-*Cre* confirming the deletion of the Ezh2 SET domain (Figure 4.2E). Together, these data indicate that the catalytic activity of Ezh2 is required for reprogramming and particularly during the early stages of the reprogramming process.

To address the possibility that the inhibitory reprogramming effect due to deletion of the Ezh2 SET domain was not acting through a dominant-negative mechanism (potentially by generating a dominant negative protein product upon SET domain deletion), we also performed siRNA-mediated knockdown of Ezh2 in a time course fashion during reprogramming, knocking down Ezh2 at days five, eight, eleven, or continuously across all three time points (Figure S4.2D). The siRNA sequences used effectively reduced Ezh2 protein and global H3K27me3 levels in ESCs (Figure S4.2E), and pheno-copied the results of the SET domain deletion (Figure 4.2D), wherein conditions downregulating Ezh2 activity inhibited the formation of Nanog-positive colonies by day 14. These data demonstrate that loss of Ezh2 activity, rather than a dominant negative effect, is responsible for the inhibitory reprogramming phenotype. Considering the lack of effect upon deletion of Ezh2 late in reprogramming, it may be important to note that Ezh2 is not required for self-renewal of ESCs (Shen et al., 2008). It may also be possible that the SET domain deletion late in reprogramming is mostly due to the fact that the cells may have already acquired a stage close to pluripotency at the point of deletion rather than that cells experienced the effect of the deletion prior to entering the pluripotent state. Further analysis that carefully tracks the timing of when Nanog- and Esrrb-positive cells appear will resolve this uncertainty.

Ezh2 regulates proliferation during early reprogramming

As PRC2 has been shown to regulate the expression of cell cycle and apoptosis regulatory proteins such as Arf, Ink4a, Ink4b, p57, and Bim in multiple cell types (Sauvageau and Sauvageau, 2010), we tested whether the deletion of the Ezh2 SET domain during the early reprogramming stage affects this crucial aspect of iPSC generation: the ability of cells to proliferate. Previous studies have identified that Arf and Ink4a, inhibitors of cellular proliferation, are barriers to the reprogramming process and that the repression of the Arf/Ink4a locus is intrinsic to reprogramming (Li et al., 2009; Utikal et al., 2009). We found that the deletion of the Ezh2 SET domain in OSK-transduced MEFs upregulated *Arf* and *Ink4a* mRNA (Figure 4.2F)

and reduced the proliferative capacity in MEFs with or without OSK-transduction (Figure S4.2G and S4.2H). Taken together, these data suggest that Ezh2 may act critically during the early stage of reprogramming to suppress the expression of cell cycle inhibitors, thereby promoting early reprogramming events. A demonstration that Arf and or Ink4a downregulation can rescue this reprogramming inhibition by Ezh2 loss of function would confirm this hypothesis in the future.

Eed is Required Throughout the Reprogramming Process

Since we observed that catalytic Ezh2 is critical for the early reprogramming stage but not for later stages, we reasoned that the PRC2 complex containing Ezh1 as the catalytic subunit may retain H3K27me3 at sites critical for reprogramming. Therefore, we also tested how the loss of both PRC2 complexes and of all H3K27me3 affects the reprogramming process, particularly during the late stage. To directly address the necessity of PRC2 function during the late reprogramming stage, we knocked down the essential PRC2 subunit Eed (Montgomery et al., 2005) continuously and at early and late time windows during the reprogramming process using a doxycycline-inducible shRNA vector. This knockdown system achieved efficient reduction of *Eed* mRNA upon addition of doxycycline (Figure 4.2G) and resulted in a drastic suppression of Nanog-positive colony formation when Eed was continuously downregulated from days -2 to 14 (Figure 4.2H), phenocopying the Cre-mediated deletion of the Ezh2 SET domain at day 0 (Figure 4.2D). Furthermore, induced knockdown of Eed from days -2 to 7 effectively inhibited reprogramming, supporting the observation that a PRC2 complex incorporating Ezh2 is essential in the early reprogramming phase (Figure 4.2H). Importantly, Eed knockdown from days 7 to 14 also dramatically reduced the formation of Nanog-positive colonies (Figure 4.2H). This suggests that the function of both PRC2 complexes and H3K27me are also critical during the late phase of reprogramming, and that Ezh1-mediated H3K27me may be critical for the late stage of reprogramming.

iPSCs generated in the absence of Ezh2 methyltransferase activity are pluripotent

While deletion of the Ezh2 SET domain at day zero of reprogramming yielded no iPSCs by day 15, deletion at day three yielded a reasonable number of clonally expandable iPSC lines by day 15 that were Ezh2 mutants (Figure 4.2D). PCR genotyping of three iPSC lines that were independently derived from these reprogramming cultures confirmed the presence or absence of the SET domain in control- or Cre-infected cultures, respectively (Figure 4.3A). qPCR analysis of these 1lox iPSC lines (i.e iPSC lines lacking the SET domain) confirmed the lack of mRNA transcripts coding for the SET domain (Figure S4.3A), and immunostaining demonstrated a marked reduction of global H3K27me3 levels (Figure 4.3B). All iPSCs lines exhibited hallmarks for faithfully reprogrammed cells, including expression of Esrrb (Figure 4.3B) and Nanog (Figure S4.3A), and silencing of the exogenous reprogramming factors (Figure S4.3B). Importantly, all three Ezh2 mutant iPSCs produced teratomas, confirming their ability to give rise to cell types within the three germ layers (Figure 4.3C). These results demonstrate that the absence of Ezh2 methyltransferase activity from day three of the reprogramming process impairs the reprogramming process but allows the establishment of faithfully reprogrammed iPSCs. In subsequent reprogramming experiments, we found that deletion of the Ezh2 SET domain at day 0, and even 2 days prior to OSK transduction, can yield clonally expandable iPSC lines, but only after 25 days in culture. Pluripotency assays on these lines will confirm the faithfulness of their reprogramming and definitively test the dispensabilty of catalytic Ezh2.

Ezh2 mutant iPSCs reveal global loss of but retain H3K27me3 at specific loci

Because we could generate $Ezh2^{\Delta/\Delta}$ iPSCs when the Ezh2 SET domain was deleted at day 3 of reprogramming, we mapped the locations of H3K27me3 in three $Ezh2^{f/f}$ and three $Ezh2^{\Delta/\Delta}$ iPSC lines genome-wide by ChIP-Seq to examine where residual H3K27me3 is present, likely mediated by the Ezh1-containing PRC2. The ChIP-seq data confirmed that global

H3K27me3 marks are depleted genome-wide in the Ezh2$^{\Delta/\Delta}$ iPSCs (Figure S4.3C). Interestingly, H3K27me3 is also depleted from the PRC2 target loci of the cell cycle and apoptosis genes we assayed in the early phase of reprogramming (Figure 4.3D). This suggests that these genes may be derepressed in Ezh2$^{\Delta/\Delta}$ reprogramming and the resultant iPSC. However, the Ezh2$^{\Delta/\Delta}$ iPSCs would likely possess an unknown mechanism to counter the senescence effects expected by Arf/Ink4 upregulation. Gene expression analysis is needed to further investigate this paradoxical situation by determining the expression levels of these genes. While global ChIP-seq signals for H3K27me3 are reduced, focal spots of H3K27me3 are retained at a comparable level to Ezh2$^{f/f}$ iPSCs at distinct regions of the genome. Closer inspection of the loci retaining H3K27me3 reveals several developmental genes, including clusters of the classical Polycomb targets, the *Hox* genes (Figure 4.3E). This observation recapitulates the finding that focal H3K27me3 persists in pluripotent cells devoid of catalytically active Ezh2, due to the function of Ezh1 (Shen et al., 2008). Future work remains to investigate global gene expression differences between the Ezh2$^{f/f}$ and Ezh2$^{\Delta/\Delta}$ iPSC lines, and how these differences are associated with differential H3K27me3. It will be interesting to know if the genes whose expression are deregulated due the lack of Ezh2-mediated H3K27me deposition during the reprogramming process also undergo the same deregulation when Ezh2 activity is removed only in the pluripotent state. Deletion of the Ezh2 SET domain from Ezh2$^{f/f}$ iPSCs will address this question.

PRC2 overexpression enhances reprogramming

Expression array analysis of the core subunits of PRC2 demonstrated that the PRC2 subunits Ezh2, Eed, Suz12, but not Ezh1, are highly expressed in iPSCs and ESCs relative to MEFs and pre-iPSCs (Figure S4.4A). Immunostaining of wildtype reprogramming cultures at middle (day 6) and late (day 14) time points supports the expression array data, demonstrating

that Ezh2 protein levels gradually increase during the reprogramming process, first at an intermediate stage in E-cadherin positive cells, and then later again when the pluripotency marker Nanog is induced (Figure S4.4B). As loss-of-function of Ezh2 early in reprogramming effectively impaired the process, we also tested whether ectopically expressed Ezh2-PRC2 components would enhance the reprogramming process. To this end, we used primary, transgenic MEFs harboring the M2rtTA at the Rosa26 locus and various tet-inducible PRC2 subunits. These MEFs were obtained from mice crossed between the R26:M2rtTA mouse and several transgenic strains that carried a tet-inducible promoter driving the expression of the cDNA coding for either Ezh2, Eed, Suz12, or for all three PRC2 subunits (Figure 4.4A), allowing us to overexpress Ezh2, Suz12, Eed alone, or all three subunits of PRC2 together. The reprogramming factors OSK were transduced into these different MEF lines, and overexpression of the individual or combination of PRC2 components was induced by treating cultures with doxycycline continuously from day four post-transduction until cultures were fixed at day 15. The cultures were then immunostained and quantified for Nanog-positive colonies (Figure 4.4B). M2rtTA MEFs lacking any tet-inducible PRC2 transgene were used as a negative control. We confirmed that doxycycline treatment of reprogramming MEFs effectively induced expression of the respective PRC2 proteins (Figure 4.4A). Notably, while overexpression of each individual PRC2 component minimally enhanced the reprogramming process relative to their respective no dox control, overexpression of all PRC2 components together effectively enhanced reprogramming efficiency (Figure 4.4B). These data suggest that ectopic expression of the entire Ezh2-PRC2 complex promotes the generation of iPSCs. Taken together with the Ezh2 and Eed loss-of-function data, these gain-of-function data further support a positive role for PRC2 proteins in reprogramming.

DISCUSSION

Because PRC1 and PRC2 play important roles in regulating lineage commitment during development by establishing histone modifications, and these histone modifications are associated with changes in gene expression observed between somatic and pluripotent cells, we investigated the roles of PRC1 and PRC2 during the reprogramming process. We first observed that the PRC1 Ubiquitinylase, Ring1b is dispensable for reprogramming, despite its importance *in vivo* for early embryonic development (Voncken et al., 2003) and *in vitro* for maintaining silencing of PRC1 targets in ESCs (van der Stoop et al., 2008). While we found that Eed, an essential subunit of PRC2 required for all H3K27me3, is necessary throughout the entire reprogramming process, Ezh2 is not required throughout the entire reprogramming process. Specifically, Ezh2 is required particularly in the early stage of reprogramming in part to suppress the cell cycle inhibitors Arf and Ink4a, but gradually becomes dispensable in the later stages. This suggests that the alternative PRC2 containing Ezh1 fulfills essential H3K27 methylation late in the reprogramming process. Genome-wide mapping of H3K27me3 revealed that it is globally reduced in iPSCs derived in the absence of Ezh2 catalytic activity. However, various genomic loci retain H3K27me3, including the *Hox* genes, suggesting that the silencing of these genes is essential for reacquiring the pluripotent state. Further analysis of global expression and H3K27me3 accumulation will address what are the specific targets of Ezh2 methyltransferase activity during the reprogramming process.

It is surprising that Ring1b knockout embryos have a strong developmental phenotype (Voncken et al., 2003) yet reprogramming without Ring1b showed no effect. We noted that some Ring1b mutant ESCs can be maintained while other mutant ESCs rapidly lose ESC identity (Leeb et al., 2010; Leeb and Wutz, 2007; van der Stoop et al., 2008). The conditional allele that we used, when examined in ESCs, manifests the former phenotype and may thus explain our ability to obtain Ring1b mutant iPSCs from fibroblasts. It will be important to resolve the differences between the Ring1b alleles through their effects on reprogramming. It may be

that our Ring1b allele sensitizes reprogramming to the effects of Ring1a loss of function. On the other hand, we cannot assume that since Ring1a knockout embryos have no embryonic defect, it will not have a reprogramming effect if inactivated alone (del Mar Lorente et al., 2000). In fact, a recent study reported that the knockdown of RING1A reduced the reprogramming efficiency of human fibroblasts (Onder et al., 2012). Based on the high degree of homology between Ring1a and Ring1b (de Napoles et al., 2004), a potential redundancy between Ring1a and Ring1b during murine reprogramming will need to be addressed. Regardless of whether there is a requirement for both Ring1a and Ring1b during reprogramming, it will be interesting to see if the PRC1 targets overlap with those of PRC2 during the process. Up to this point, there has not been a comprehensive study characterizing the dynamics of H2AK119Ub1 during reprogramming, but these data will surely provide important insight into the process.

Our focus shifted to a striking effect on reprogramming by genetically inactivating Ezh2. In the early phase, the most apparent effect of Ezh2 loss of function is the derepression of the *Cdkn2a* transcripts coding for the cell cycle inhibitors Arf and Ink4a, resulting in reduced proliferation of the reprogramming fibroblasts. However, our data cannot exclude a potentially inhibitory effect due to the derepression of other Ezh2 targets during the early phase. The ability of Arf/Ink4a downregulation to rescue the reprogramming defect due to early Ezh2 inactivation would implicate Arf and Ink4a as the specific downstream Ezh2 targets mediating this effect. If Arf/Ink4a downregulation cannot rescue the early Ezh2 inactivation effect, gene expression profiling may reveal the downstream targets responsible for the early reprogramming effect.

The contrasting reprogramming effects due to the late inactivation of Ezh2 compared to Eed downregulation implicate an essential function for Ezh1 late in the reprogramming process. Current efforts are underway to identify the genes that are upregulated in the Ezh2$^{\Delta/\Delta}$ iPSCs and therefore are dependent on Ezh2-mediated deposition of H3K27me3. However, the

upregulation of these genes likely does not matter in the maintenance or reacquisition of the pluripotent state. Conversely, it will be interesting to find genes that lack H3K27me3 modifications in Ezh2$^{\Delta/\Delta}$ iPSCs but are not aberrantly expressed relative to Ezh2$^{f/f}$ iPSCs. This would suggest an alternative mechanism from H3K27me3 that is used to control gene expression that may play a critical role in reprogramming.

Because PRC2 modification of H3K27me3 is such a generally utilized mechanism of gene silencing in multiple cell types, isolating the targets that are required to be transcriptionally silenced during the reprogramming process poses a challenge. However, our data consistently demonstrate that PRC2 function is required for reprogramming. A recent report that iPSCs can be generated without Ezh2 confirmed our findings (Fragola et al., 2013). Furthermore, this study also found that global H3K27me3 levels were reduced in these Ezh2 mutant iPSC lines, but also retained focal H3K27me marks at embryonic developmental genes. Nevertheless, the question of which genomic targets between the catalytic subunits of PRC1 and PRC2 overlap remain unanswered and invite deeper investigation.

Overall, the findings from Testa and colleagues (Fragola et al., 2013), as well as our own, are in contrast to those of other reports, where deletions of many other repressive chromatin regulators typically enhance reprogramming (Luo et al., 2013; Pawlak and Jaenisch, 2011; Soufi et al., 2012; Sridharan et al., 2013). An integrated model of chromatin dynamics will provide practical implications for modeling cell state plasticity in the context of development and regenerative medicine.

EXPERIMENTAL PROCEDURES

Cell Lines, Expression Constructs and Reprogramming Experiments

For reprogramming with retroviral factors, Oct4, Sox2, Klf4, and cMyc were expressed from pMX retroviruses, as previously described (Maherali et al., 2007). MEFs cultured in mouse ESC media containing 15% Fetal bovine serum (FBS) were infected with respective retroviruses for 12-18 hr and split three days post-transduction onto coverslips that were pre-seeded with irradiated mouse feeders. Typically, mouse ESC media containing Leukemia inhibitory factor (Lif) and 15% Knockout serum replacement (KSR) was added on day five post-transduction and changed every three days thereafter until the indicated point of analysis. Reprogramming efficiency was determined by quantifying the number of colonies immunostaining for Nanog or Esrrb on fixed reprogramming cultures. In *Rnf2* (Ring1b) knockout reprogramming experiments, *Rnf2* conditional MEFs (Cales et al., 2008) carrying an IRES Cre-ER constitutively expressed from the 3' UTR of the PolII gene were used. Treatment with 1 μM 4-Hydroxytamoxifen (Sigma H7904) for 24 hr induced efficient ablation of *Rnf2*, and was performed at indicated time points. Ethanol was used as a vehicle control. In *Ezh2* SET domain knockout reprogramming experiments, transgenic MEFs homozygous for Ezh2 alleles in which exons 3-6, which code for the catalytic SET domain for histone methyltransferase activity, are flanked by loxP sites. Adenoviral transduction of *Cre* for one hour at different time points during reprogramming. Control Adenovirus was used in parallel. For inducible *Eed* knockdown experiments, the shRNA hairpin sequence 5TCTTGCTAGTAAGGGCACATA3 was cloned into the tet-pLKO-neo vector, allowing inducible expression in infected cells that were continuously selected with 400 μg/ml neomycin by adding 2 μg/ml doxycycline. For siRNA *Ezh2* knockdown experiments, 20nM siRNA with the sequence 5GGAGGGAGCUAAGGAGUUU3 was transfected with RNAi-MAX. siRNA targeting Luciferase was purchased from Dharmacon RNA Technologies and used for control knockdown.

To isolate and expand iPSC lines, colonies from primary reprogramming cultures were transferred with a micropipette to 96 well plates, trypsinized, plated into 24 well plates, and

subsequently passaged into 6 well plates with continuous culture in ESC media containing Lif and 15% of either FBS or KSR. Culture vessels were always gelatinized and seeded with irradiated feeders. For passaging of all iPSC lines derived from KSR reprogramming conditions, trypsin inhibitor (Invitrogen 17075-029) was used as instructed. For teratoma formation, 500,000 cells for each iPSC line were injected subcutaneously into the leg muscle of SCID mice. Teratomas were recovered three weeks post-injection, fixed overnight in 4% formaldehyde, paraffin embedded, and processed with hematoxylin and eosin.

Western Blots, Immunostaining, and AP-staining

The antibodies used for Western blotting were mouse anti-Ezh2 (BD Pharmingen 612667), mouse anti-Eed (Otte Lab), rabbit anti-Suz12 (Abcam ab12201-100), and anti-Gapdh (Fitzgerald RD1-TRK5G4-6C5). For immunostaining, cell lines or reprogramming cultures were grown on glass coverslips, fixed with 4% paraformaldehyde (PFA) in PBS for 10 min, permeabilized with 0.5% Triton 100-X in PBS for 5 min, and immunostained with mouse anti-Nanog (BD Pharmingen 560259) or rabbit anti-Nanog (Abcam ab80892), mouse anti-Esrrb antibody (R&D Systems H6705), rat anti-Cdh1 (Abcam ab11512), rabbit anti-Ezh2 (ab3748) or mouse anti-Ezh2 (BD Pharmingen 612667), mouse anti-Ring1b (MBL D139-3), and rabbit anti-H3K27me3 (Active Motif 39155). Clusters of cells with at least three nuclei staining positive for Nanog or Esrrb were scored as an iPSC colony. For AP detection, reprogramming cultures were fixed with 4% PFA in PBS for 20 min, washed with water, and incubated with a mixture of 1mg/ml Fast Red TR Salt (Sigma) and 4% Napthol AS-MX Solution (Sigma) for 10 min.

RT- and Genotyping PCRs

Total RNA was isolated from cells using the RNeasy kit (QIAGEN) and cDNA generated with Superscript III (Invitrogen). qPCR values were generated using the ddCT method normalized to U6. Primers used for detecting expression of pMX transgenes have been previously described

(Maherali et al., 2007). qPCR primers used for detecting mRNA expression are as follows: Ezh2 SET domain FW: 5TGTATGACAAATACATGTGCAGCT3, RV: 5GATTTACTGAATGATTAGCAAAACGA3; Eed: FW: 5GCACAGAGATGAAGTTCTGAGTGCTG3, RV: 5ATAAGACTCCTTAATTGCATTCATCATCCT3; Nanog: FW: 5AGGGTCTGCTACTGAGATGCTCTG3, RV: 5CAACCACTGGTTTTTCTGCCACCG3; Arf: FW: 5GCCGCACCGGAATCCT3, RV: 5TTGAGCAGAAGAGCTGCTACGT3; Ink4a: FW: 5GTGTGCATGACGTGCGGG3, RV: 5GCAGTTCGAATCTGCACCGTAG3; Ink4b: FW: 5GCAGGCCTTCCAAAACTTGA3, RV: 5AGCTGCAGAAAATGCGTAGGA3; p57: FW: 5CCGACTGAGAGCAAGCGAA3, RV: 5GGTCGAAGGCTGGCTGATTG3; Bim: FW: 5CCCTACAGACAGAACCGCAA3, RV: 5TGCAAACACCCTCCTTGTGT3; U6: FW: 5CGCTTCGGCAGCACATATAC3, RV: 5TCACGAATTTGCGTGTCATC3. For PCR genotyping Rnf2 alleles, primers with the following sequences were used: 2lox FW: 5TGCTCCTTTTTGATGGAACC3, RV: 5TCCTCAAACTGGTGTCCAAAC3; 1lox: FW: 5ATGGTCAAGCAAACATGAAGG3, RV: 5AAGCCAAAATTTAAAAGCACTGT3. For PCR genotyping Ezh2 SET domain alleles, primers with the following sequences were used: 2lox: FW: 5GTGACAGGTCTTAAAACTAAGAGTAA3, RV: 5GTTTTTCTCAACATTTCATTTTTGTAT3; 1lox: FW: 5GTGACAGGTCTTAAAACTAAGAGTAA3, RV: 5CTCTGAATGGCAACTCCACA3.

ChIP-seq for Transcription Factors

V6.5 embryonic stem cells and mouse embryonic fibroblasts (MEF) cells were grown to a final concentration of $5x10^7$ cells for each sequencing experiment. Cells were chemically cross-linked by the addition of formaldehyde to 1% final concentration for 10 minutes and quenched with 0.125 M final concentration glycine. Cells were then resuspended in buffer I (0.3M sucrose, 60mM KCl,15mM NaCl,5mM $MgCl_2$ 10mM EGTA,15mM Tris-HCl, 0.5mM DTT, 0.2% NP40 and protease inhibitor cocktail) and incubated on ice for 10 minutes. Nuclei were generated by

centrifugation via a sucrose cushion (1.2M sucrose, 60mM KCL,15mM NaCl, 5mM $MgCl_2$, 0.1mM Tris-HCl, 0.5mM DTT and protease inhibitor cocktail). Isolated nuclei were resuspended in sonication buffer (50mM Hepes, 140mM NaCl, 1mM EDTA, 1% TritonX-100, 0.1% Na-deoxycholate, 0.1% SDS) and sonicated using a Diagenode Bioruptor. Nuclear extracts were then incubated overnight at 4^{o}C with the respective antibodies: anti-Esrrb (RnD; H6705), anti-Klf4 (RnD; AF3158), anti-Myc (RnD; AF3696), anti-Nanog (cosmobio), anti-OCT4 (RnD; AF1759), anti-Sox2 (RnD AF2018), anti-p300 (SantaCruz;sc-585). Extracts were washed twice with RIPA, low salt buffer (20mM Tris pH 8.1, 150mM NaCl, 2mM EDTA, 1% Triton X-100, 0.1% SDS), high salt buffer (20mM Tris pH 8.1, 500mM NaCl, 2mM EDTA, 1% Triton X-100, 0.1% SDS), LiCl buffer (10mM Tris pH 8.1, 250mM LiCl, 1mM EDTA, 1% deoxycholate, 1% NP-40) and 1xTE. Reverse cross-linking occurred by overnight incubation at 65^{o}C with 1% SDS and proteinase K. All protocols for Illumina/Solexa sequence preparation, sequencing and quality control are provided by Illumina with minor modification by limiting the PCR amplification step to 10 cycles.

Native Chromatin Immunoprecipitation

All histone marks used for this study were determined using native ChIP (Wagschal et al., 2007). Briefly nuclei from non-crosslinked V6.5 and MEF cells were purified as descfibed above. Nuclei were then resuspended in Mnase-digestion buffer (0.32M sucrose, 50mM Tris-HCl, 4mM $MgCl_2$, 1mM $CaCl_2$, protease inhibitor cocktail) and digested with 3 units of MNase (Roche) for 10 minutes at 37^{o}C. Soluble chromatin fractions were then incubated with anti-H3K4me3 (Abcam ab8580), anti-H3K4me2(Abcam ab7766), anti-H3K4me1(Abcam ab8895), anti-H3K27me3 (Active Motif 39155), anti-H3K27ac (Abcam ab4729) and anti-H3K36me3 (Abcam ab9050). Extracts were washed twice with wash buffer A (50mM Tris-HCl, 10mM EDTA, 75mM NaCl), wash buffer B (50mM Tris-HCl, 10mM EDTA, 125mM NaCl), wash buffer C (50mM Tris-

HCl, 10mM EDTA, 250mM NaCl), 1x LiCl buffer and 1xTE. DNA extraction and library preparation as described above.

Figure 4.1. Ring1b Ablation Has No Effect on Reprogramming Efficiency

A. Schematic of the reprogramming experiment utilizing MEFs homozygous for a conditional Ring1b allele (fl/fl or 2lox/2lox), where Ring1b can be deleted at several timepoints by treatment with 1uM 4OHT for 24 hours. Reprogramming is assessed by counting the number of Nanog positive colonies at day 12 of reprogramming as determined by immunostaining.

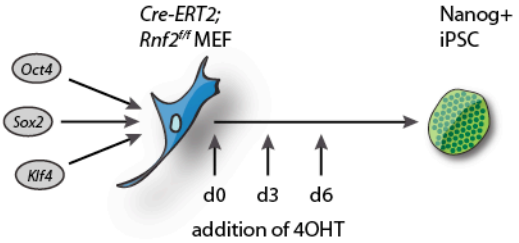
B. Representative immunofluorescence images of Nanog-positive colonies at day 12 of reprogramming. Ring1b is typically highly expressed in Nanog-positive cells present in reprogramming cultures treated with vehicle (Ctrl) at day zero, while absent from Nanog-positive cells present in cultures treated with 4OHT at day zero.

C. Quantification of colonies expressing Nanog and Ring1b in reprogramming cultures at day twelve of the reprogramming process. Deletion of Ring1b was induced by treatment with 4OHT at indicated timepoints. Upon addition of 4OHT almost all colonies lack Ring1b staining confirming efficient deletion of Ring1b.

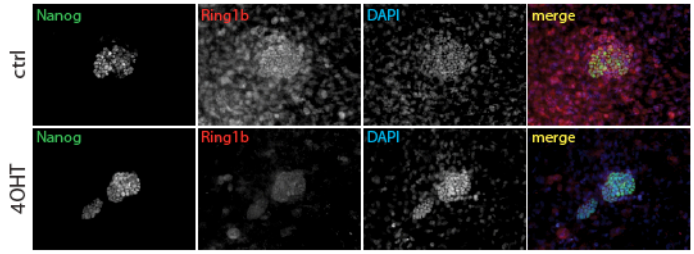
See Figure S4.1 for additional information.

Figure 4.1

A



B



C

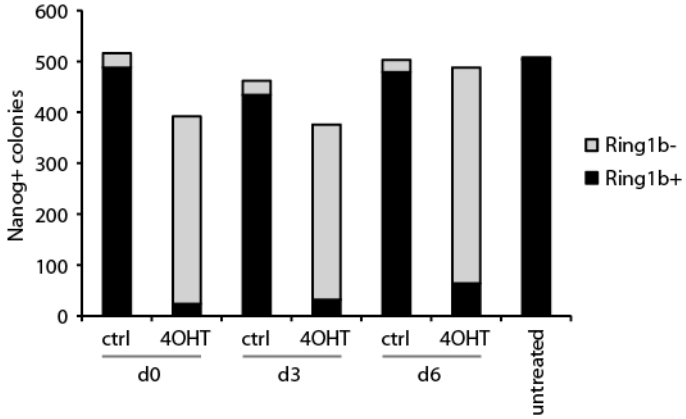


Figure 4.2. PRC2 is Required Throughout Reprogramming, But Ezh2 Activity is Dispensable in Late Stages

A. Schematic of the reprogramming experiments with MEFs homozygous for a conditional Ezh2 allele (where the catalytic SET domain is flanked by LoxP sites). Excision of the SET domain of Ezh2 can be induced at different timepoints of reprogramming by infection with adenoviral *Cre*. Reprogramming cultures were fixed at day 15 for analysis of pluripotency marker expression.

B. Alkaline phosphatase (AP) staining at day 15 of reprogramming cultures infected with adenoviral *Cre* or control (*ctrl*) virus at the indicated timepoints, or an uninfected control, at day 15 of reprogramming.

C. Quantification of AP-positive colonies from reprogramming cultures depicted in 2B.

D. Quantification of colonies expressing Nanog in reprogramming cultures infected with *Cre* or *ctrl* virus at indicated timepoints, at day 15 of reprogramming.

E. Quantification of colonies at day 15 expressing Esrrb in reprogramming cultures infected with *Cre* or *ctrl* virus at indicated timepoints. Co-staining for H3K27me3 was performed to asses the extent of Ezh2 deletion in these Esrrb positive colonies.

F. qPCR analysis for senescence and apoptosis regulator expression from reprogramming cultures infected at day zero with *Cre* or *ctrl* virus at day 4 of reprogramming. Values represent the average of triplicate sampling and error bars standard deviation.

G. qPCR analysis for Eed expression from reprogramming cultures described in (H) at day 4 of reprogramming. Values represent the average of triplicate sampling and error bars standard deviation.

H. Scheme for timed Eed depletion by shRNA knockdown in reprogramming cultures. MEFs were infected with a polycistronic, dox-inducible shRNA vector, which constitutively expresses a tTA repressor that silences expression of of the shRNA hairpin. Exposure to dox derepresses the shRNA, thereby enabling Eed knockdown. Yellow bars indicate the time windows in which cultures were exposed to dox, and the days are indicated with respect to when MEFs were transduced with OSK. Quantification of Nanog-positive colonies and their global H3K27me3 level at day 14 of reprogramming are shown to the right of the scheme. Values represent the average of technical replicates and error bars standard deviation.

See Figure S4.2 for additional information.

Figure 4.2

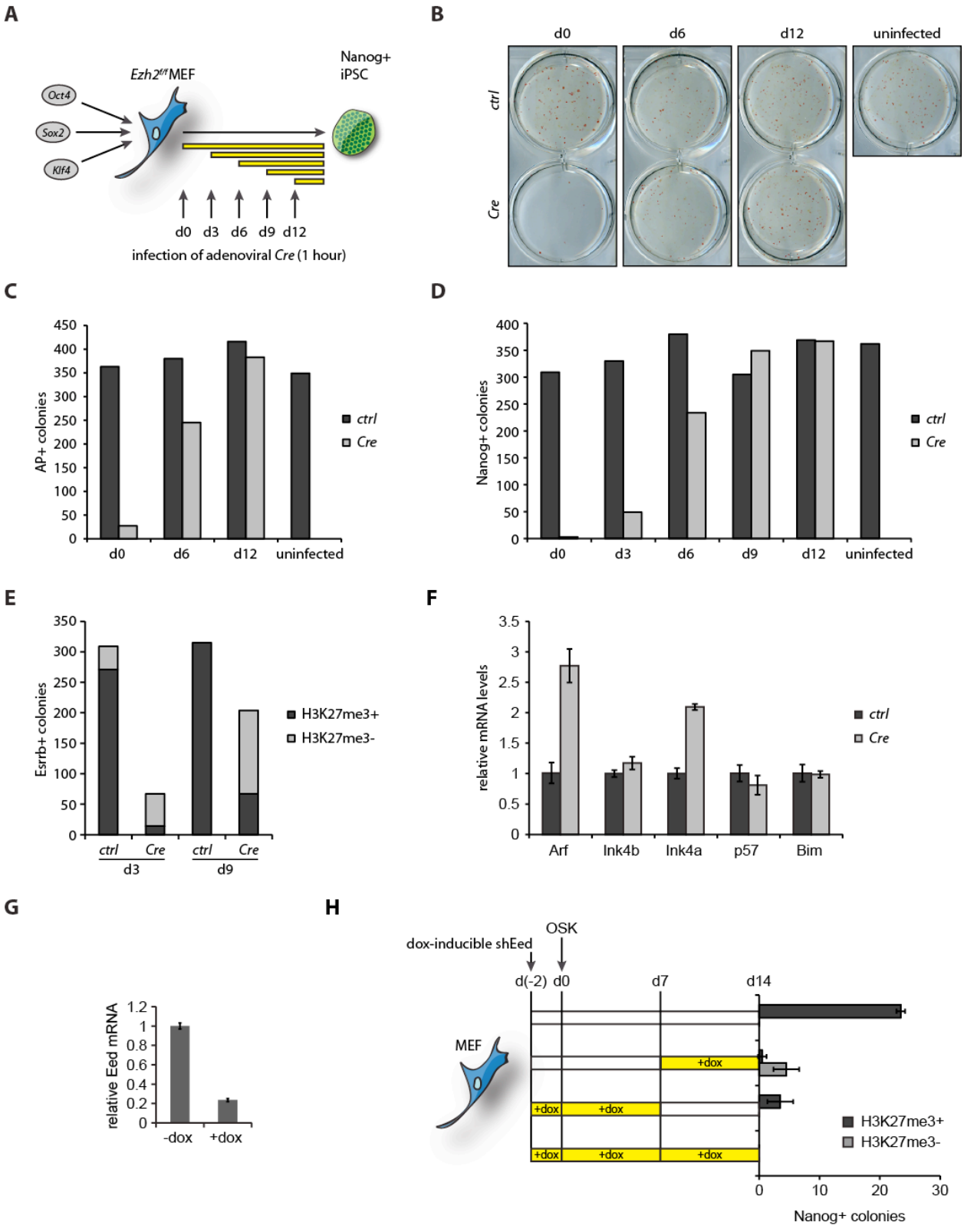

A
Ezh2fl/fl MEF
Nanog+ iPSC
Oct4
Sox2
Klf4
d0 d3 d6 d9 d12
infection of adenoviral Cre (1 hour)
B
d0
d6
d12
uninfected
ctrl
Cre
C
AP+ colonies
ctrl
Cre
d0
d6
d12
uninfected
D
Nanog+ colonies
ctrl
Cre
d0
d3
d6
d9
d12
uninfected
E
Esrrb+ colonies
H3K27me3+
H3K27me3-
ctrl
Cre
ctrl
Cre
d3
d9
F
relative mRNA levels
ctrl
Cre
Arf
Ink4b
Ink4a
p57
Bim
G
relative Eed mRNA
-dox
+dox
H
dox-inducible shEed
OSK
d(-2) d0
d7
d14
MEF
+dox
+dox
+dox
+dox
+dox
H3K27me3+
H3K27me3-
Nanog+ colonies

Figure 4.3. iPSCs Lacking Catalytic Ezh2 Can Be Established with Global Decrease in H3K27me3

A. PCR genotyping for Ezh2 SET domain 2lox and 1lox alleles in iPSC lines expanded from cultures infected with adenoviral *Cre* or *ctrl*, which was added at day 3 of reprogramming, confirming that iPSCs derived from the Cre-treated cultures lack the Ezh2 SET domain.

B. Representative immunofluorescence images of Esrrb and H3K27me3 staining on 2lox and 1lox Ezh2 iPSC lines A. H3K27me3 is typically highly detectable in Esrrb-positive cells derived from reprogramming cultures infected with control virus, while absent from Esrrb-positive cells derived from cultures infected with adeno-*Cre* virus.

C. Subcutaneous injections of ESCs, control, and mutant iPSC lines A into NOD/SCID mice produced teratomas exhibiting histological structures representative of all three germ layers after three weeks.

D. H3K27me3 ChIP-seq profiles at cell cycle regulator genes in Ezh2 wildtype (2lox) or mutant (1lox) iPSC lines A and B. Lower panel indicates locations bound by Ezh2 or Eed in wildtype ESCs. Note: H3K27me3 ChIP-seq values presented are raw values and not normalized to input, however the sequencing depth across all samples were comparable.

E. As in (D) but at the HoxC cluster.

See Figure S4.3 for additional information.

Figure 4.3

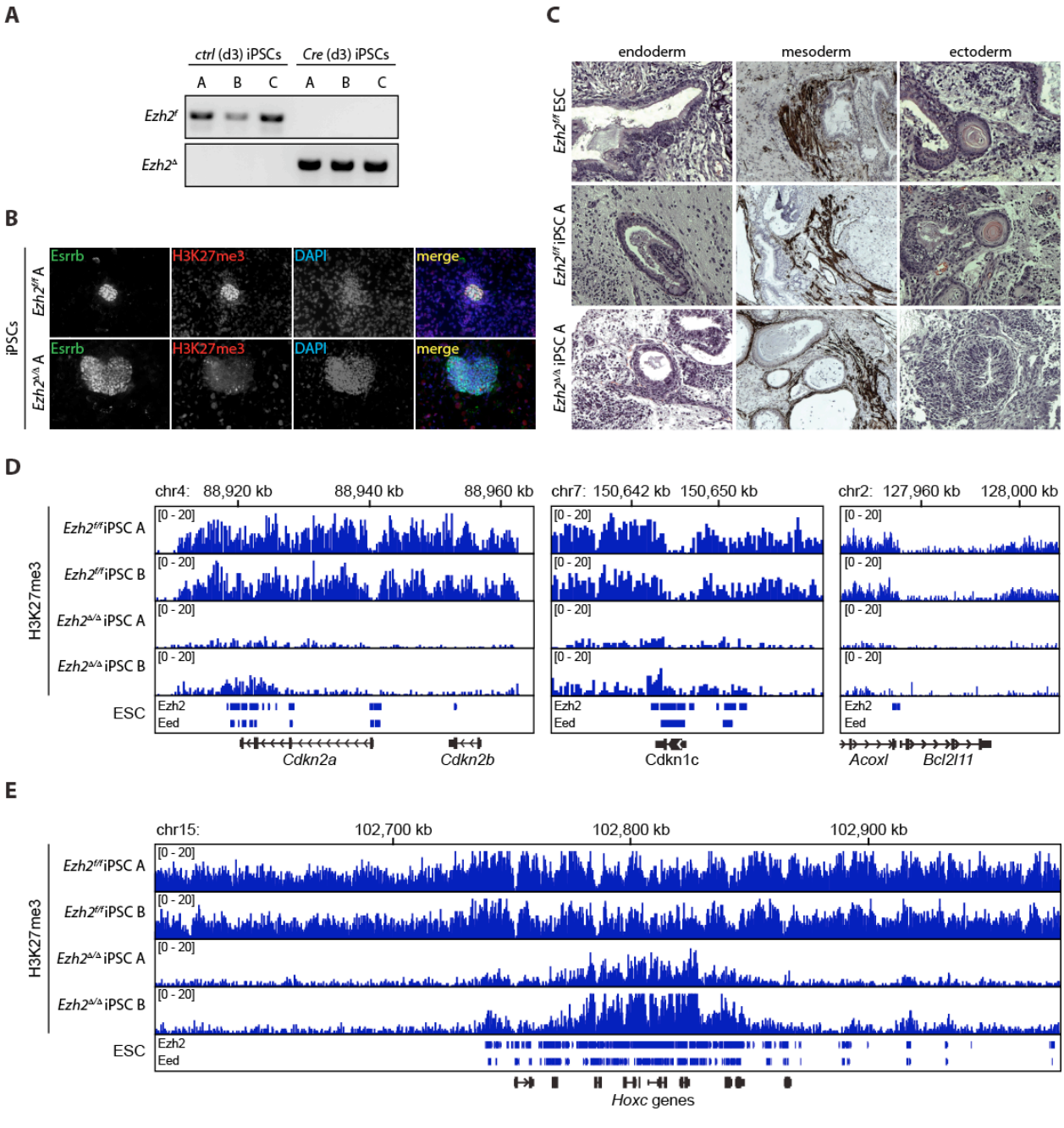

A
ctrl (d3) iPSCs
Cre (d3) iPSCs
A B C A B C
Ezh2f
Ezh2Δ
B
iPSCs
Ezh2f/f A
Ezh2Δ/Δ A
Esrrb
H3K27me3
DAPI
merge
C
endoderm
mesoderm
ectoderm
Ezh2f/f ESC
Ezh2f/f iPSC A
Ezh2Δ/Δ iPSC A
D
H3K27me3
chr4: 88,920 kb 88,940 kb 88,960 kb
chr7: 150,642 kb 150,650 kb
chr2: 127,960 kb 128,000 kb
Ezh2f/f iPSC A
Ezh2f/f iPSC B
Ezh2Δ/Δ iPSC A
Ezh2Δ/Δ iPSC B
[0 - 20]
ESC
Ezh2
Eed
Cdkn2a
Cdkn2b
Cdkn1c
Acoxl
Bcl2l11
E
chr15:
102,700 kb
102,800 kb
102,900 kb
Hoxc genes

Figure 4.4. Ectopic Expression of PRC2 Proteins Enhances Reprogramming

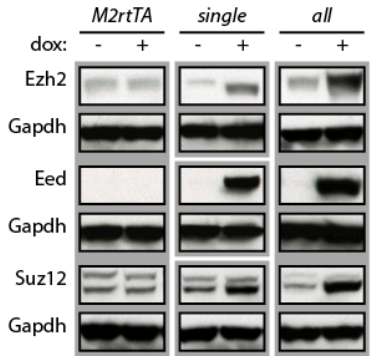
A. Transgenic MEFs were employed for overexpression of the PRC2 subunits Eed, Ezh2, and/or Suz12. These MEFs harbored the M2rtTA reverse tet-transactivator in the R26 locus. In addition, these MEFs carried the tetO Suz12 transgene in the second R26 allele, tetO-Ezh2 in the Col1A locus, and tetO Eed in the second Col1A allele (all), or only one of these tet-inducible transgenes (single tetO MEFs). Shown are western blots for Ezh2, Eed, and Suz12 expression in these transgenic MEFs with or without treatment with doxycycline. Left panel: M2rtTA MEF with no tet-inducible transgene; middle panel: M2rtTA MEFs that individually carry tet-inducible Ezh2 (top), Eed (center), and Suz12 (bottom); right panel: M2rtTA MEF that carry tet-inducible alleles for all three PRC2 subunits. Gapdh detection is shown directly beneath panels for which it serves as a loading control.

B. Fold increase of colonies expressing Nanog in reprogramming cultures induced at day 4 to overexpress each PRC2 component individually (Ezh2, Eed, Suz12), or in combination (all) over no dox induction, at day 15 of reprogramming. M2rtTA MEFs lacking a tet-inducible PRC2 transgene served as a control.

See Figure S4.4 for additional information.

Figure 4.4

A


M2rtTA
single
all
dox:
- + - + - +
Ezh2
Gapdh
Eed
Gapdh
Suz12
Gapdh

B

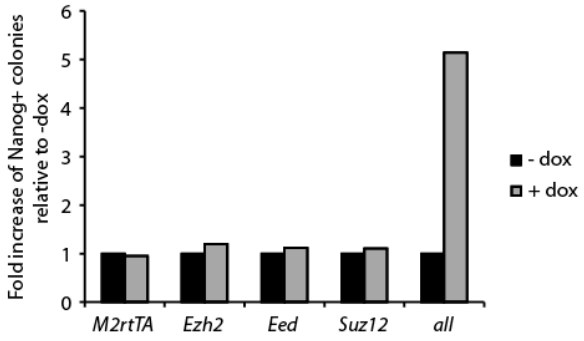

Fold increase of Nanog+ colonies relative to -dox
0
1
2
3
4
5
6
M2rtTA
Ezh2
Eed
Suz12
all
- dox
+ dox

Figure S4.1. Characterization of PRC1 Expression During Reprogramming, Related to Figure 4.1

A. Microarray expression values for the indicated PRC1 component genes in MEFs, a line of partially reprogrammed intermediate cells known as pre-iPSCs, iPSCs, and ESCs. Values represent the average of replicate microarray experiments (Sridharan et al., 2009).

B. PCR genotype for *Rnf2* 2lox and 1lox alleles (the gene for Ring1b) from MEFs treated with vehicle (ctrl) or 4OHT in the absence or presence of reprogramming factors OSK (Oct4, Sox2, and Klf4).

C. Quantification of colonies expressing Esrrb in reprogramming cultures treated with vehicle or 4OHT at indicated timepoints at day twelve of the reprogramming process.

Figure S4.1

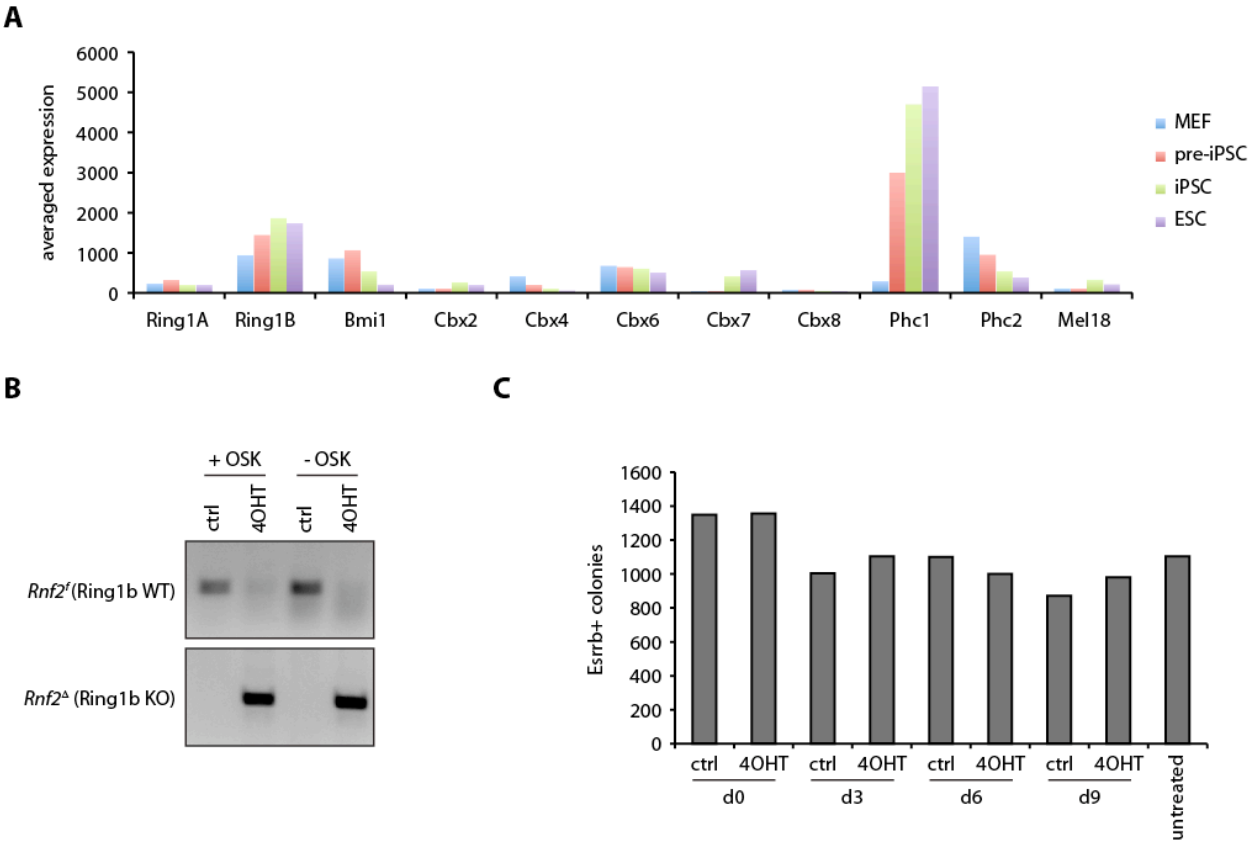

A
averaged expression
6000
5000
4000
3000
2000
1000
0
Ring1A
Ring1B
Bmi1
Cbx2
Cbx4
Cbx6
Cbx7
Cbx8
Phc1
Phc2
Mel18
MEF
pre-iPSC
iPSC
ESC
B
+ OSK
- OSK
ctrl
4OHT
ctrl
4OHT
Rnf2f (Ring1b WT)
Rnf2Δ (Ring1b KO)
C
Esrrb+ colonies
1600
1400
1200
1000
800
600
400
200
0
ctrl
4OHT
ctrl
4OHT
ctrl
4OHT
ctrl
4OHT
untreated
d0
d3
d6
d9

Figure S4.2. Characterization of Ezh2 Expression and Inactivation Effects During Reprogramming, Related to Figure 4.2

A. Schematic of the conditional Ezh2 allele (where the catalytic SET domain is flanked by LoxP sites). Excision of the SET domain of Ezh2 can be induced upon infection with adenoviral *Cre*.

B. PCR genotyping for the Ezh2 SET domain 2lox and 1lox alleles from MEFs infected with adenoviral *ctrl* or *Cre*.

C. qPCR analysis of day three reprogramming cultures infected with adenoviral *ctrl* or *Cre* on day zero or reprogramming. qPCR primers are internal to the SET domain which is deleted upon Cre addition. Values represent the average of triplicate sampling and error bars standard deviation.

D. Schematic of the reprogramming experiment in which Ezh2 was knocked down by siRNAs at several timepoints by transfection with siRNA. Nanog-positive colonies were quantified at day 14.

E. Immunofluorescence images of Ezh2 and H3K27me3 stainings in ESCs transfected with siRNAs targeting Ezh2 or Luciferase (ctrl).

F. Quantification of colonies expressing Nanog in reprogramming cultures transfected with siRNA targeting Ezh2, ctrl, or no siRNA (mock) at indicated timepoints, at day 15 of reprogramming. Values represent the average of technical replicates and error bars standard deviation.

G. Cell counts for Ezh2 SET domain 2lox reprogramming cells infected with *ctrl* or *Cre* virus on day zero. Time points refer to days after OSK infection.

H. Cell counts for Ezh2 SET domain 2lox MEFs infected with *ctrl* or *Cre* virus. MEFs were infected and counted up to five days after viral infection, passaged, and counted up to seven days thereafter.

Figure S4.2

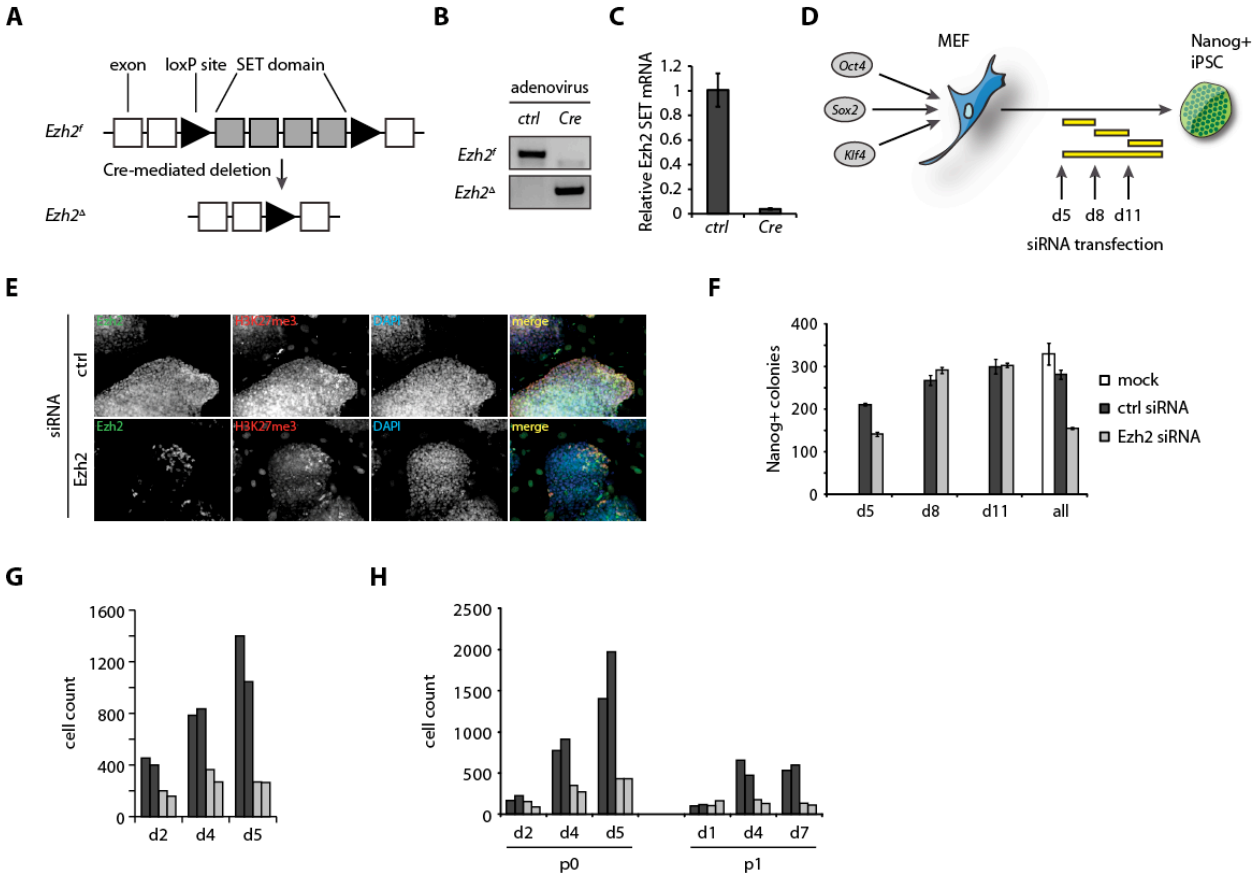

A
exon
loxP site
SET domain
Ezh2f
Cre-mediated deletion
Ezh2Δ
B
adenovirus
ctrl
Cre
Ezh2f
Ezh2Δ
C
Relative Ezh2 SET mRNA
1.2
1
0.8
0.6
0.4
0.2
0
ctrl
Cre
D
MEF
Oct4
Sox2
Klf4
Nanog+
iPSC
d5
d8
d11
siRNA transfection
E
siRNA
ctrl
Ezh2
Ezh2
H3K27me3
DAPI
merge
F
Nanog+ colonies
400
300
200
100
0
d5
d8
d11
all
mock
ctrl siRNA
Ezh2 siRNA
G
cell count
1600
1200
800
400
0
d2
d4
d5
H
cell count
2500
2000
1500
1000
500
0
d2
d4
d5
d1
d4
d7
p0
p1

Figure S4.3. Characterization of iPSCs Lacking Catalytic Ezh2, Related to Figure 4.3

A. qPCR analysis for the Ezh2 SET domain as well as Nanog expression in 2lox and 1lox iPSC lines. As described in Figure 3, these iPSC lines were expanded from reprogramming cultures infected with adenoviral *Cre* or *ctrl*, which was added at day 3 of reprogramming.

B. qPCR analysis for expression of retroviral Oct4, Sox2, or Klf4 in Ezh2 2lox and 1lox iPSC lines. RNA obtained from day seven reprogramming cultures served as a positive reference.

C. H3K27me3 ChIP-seq profiles at the whole genome level in Ezh2 wildtype or mutant iPSCs. Note: H3K27me3 ChIP-seq values presented are raw values and not normalized to input, however the sequencing depth across all samples were comparable.

Figure S4.3

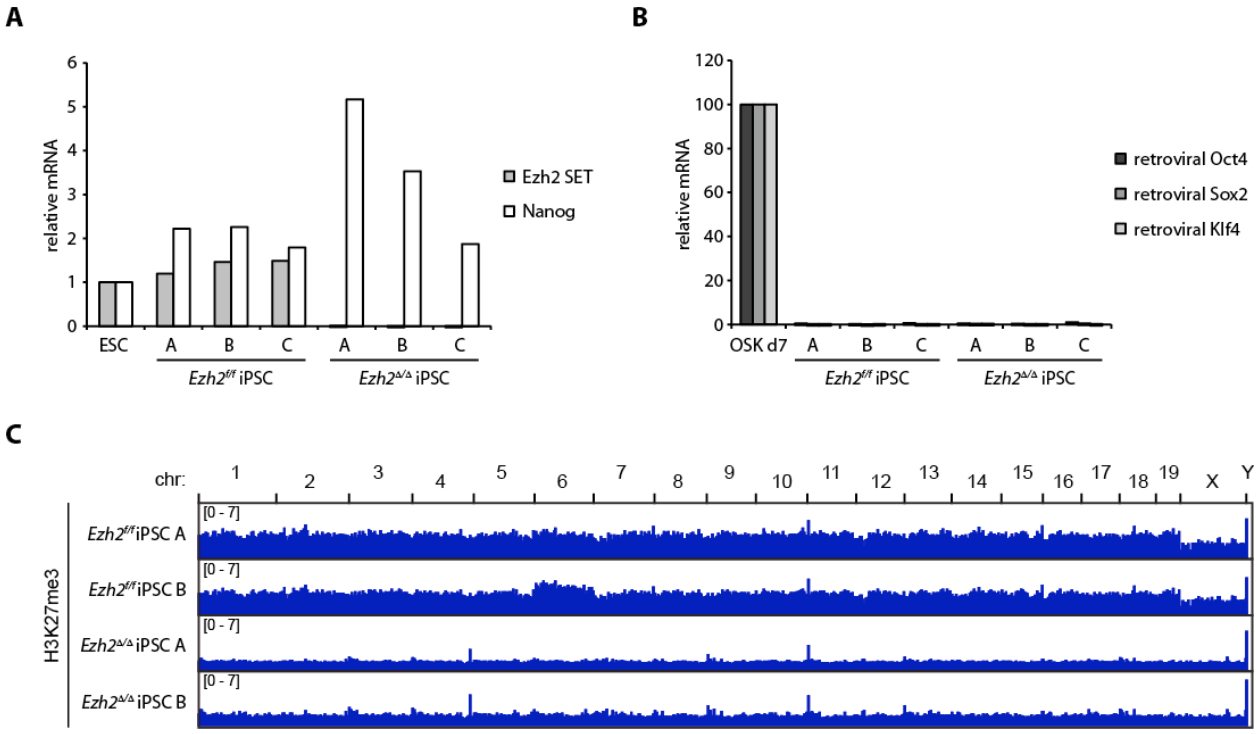
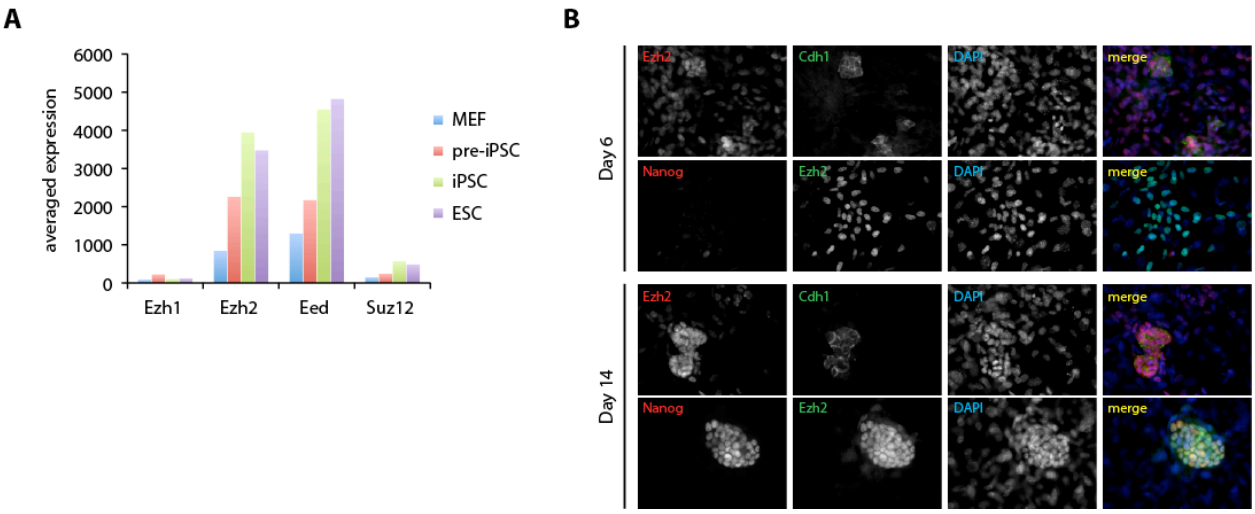

A
relative mRNA
Ezh2 SET
Nanog
ESC
A
B
C
A
B
C
Ezh2f/f iPSC
Ezh2Δ/Δ iPSC
B
relative mRNA
retroviral Oct4
retroviral Sox2
retroviral Klf4
OSK d7
A
B
C
A
B
C
Ezh2f/f iPSC
Ezh2Δ/Δ iPSC
C
chr:
H3K27me3
Ezh2f/f iPSC A
Ezh2f/f iPSC B
Ezh2Δ/Δ iPSC A
Ezh2Δ/Δ iPSC B
[0 - 7]

Figure S4.4. Characterization of PRC2 Expression During Reprogramming, Related to Figure 4.4

A. Microarray expression values for the indicated PRC2 component genes in MEFs, pre-iPSCs, iPSCs, and ESCs. Values represent average of replicate microarray experiments (Sridharan et al., 2009).

B. Ezh2 protein levels rise in reprogramming. Representative immunofluorescence images of reprogramming cultures at day 6 (top panels) and day 14 (bottom panels). At day 6, Ezh2 is moderately expressed in early reprogramming intermediate cells that are Cdh1-positive and Nanog-negative. At day 14, Ezh2 is highly expressed in more advanced reprogramming cells that are Cdh1-positive as well as in Nanog-positive.

Figure S4.4


A
averaged expression
6000
5000
4000
3000
2000
1000
0
Ezh1
Ezh2
Eed
Suz12
MEF
pre-iPSC
iPSC
ESC
B
Day 6
Day 14
Ezh2
Cdh1
DAPI
merge
Nanog
Ezh2
DAPI
merge
Ezh2
Cdh1
DAPI
merge
Nanog
Ezh2
DAPI
merge

REFERENCES

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Cales, C., Roman-Trufero, M., Pavon, L., Serrano, I., Melgar, T., Endoh, M., Perez, C., Koseki, H., and Vidal, M. (2008). Inactivation of the polycomb group protein Ring1B unveils an antiproliferative role in hematopoietic cell expansion and cooperation with tumorigenesis associated with Ink4a deletion. Mol Cell Biol *28*, 1018-1028.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J.*, et al.* (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell *5*, 111-123.

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M.*, et al.* (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell *7*, 663-676.

del Mar Lorente, M., Marcos-Gutierrez, C., Perez, C., Schoorlemmer, J., Ramirez, A., Magin, T., and Vidal, M. (2000). Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. Development *127*, 5093-5100.

Endoh, M., Endo, T.A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A.P., Okano, M., Brockdorff, N., Vidal, M.*, et al.* (2008). Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development *135*, 1513-1524.

Fragola, G., Germain, P.L., Laise, P., Cuomo, A., Blasimme, A., Gross, F., Signaroldi, E., Bucci, G., Sommer, C., Pruneri, G.*, et al.* (2013). Cell reprogramming requires silencing of a core subset of polycomb targets. PLoS Genetics *9*, e1003292.

He, S., Nakada, D., and Morrison, S.J. (2009). Mechanisms of stem cell self-renewal. Annu Rev Cell Dev Biol *25*, 377-406.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes & Development *24*, 265-276.

Leeb, M., and Wutz, A. (2007). Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. J Cell Biol *178*, 219-229.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136-1139.

Luo, M., Ling, T., Xie, W., Sun, H., Zhou, Y., Zhu, Q., Shen, M., Zong, L., Lyu, G., Zhao, Y.*, et al.* (2013). NuRD Blocks ReprogAramming of Mouse Somatic Cells into Pluripotent Stem Cells. Stem Cells *31*, 1278-1286.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55-70.

Mansour, A.A., Gafni, O., Weinberger, L., Zviran, A., Ayyash, M., Rais, Y., Krupalnik, V., Zerbib, M., Amann-Zalcenstein, D., Maza, I.*, et al.* (2012). The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. Nature *488*, 409-413.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P., and Magnuson, T. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr Biol *15*, 942-947.

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.

Pawlak, M., and Jaenisch, R. (2011). De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. Genes & Development *25*, 1035-1040.

Sauvageau, M., and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell *7*, 299-313.

Savatier, P., Lapillonne, H., van Grunsven, L.A., Rudkin, B.B., and Samarut, J. (1996). Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. Oncogene *12*, 309-322.

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.

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.

Sridharan, R., Gonzales-Cope, M., Chronis, C., Bonora, G., McKee, R., Huang, C., Patel, S., Lopez, D., Mishra, N., Pellegrini, M.*, 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.

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.

Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D., and Dalton, S. (2002). Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. Oncogene *21*, 8320-8333.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol *9*, 1428-1435.

Surface, L.E., Thornton, S.R., and Boyer, L.A. (2010). Polycomb group proteins set the stage for early lineage commitment. Cell Stem Cell *7*, 288-298.

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.

van der Stoop, P., Boutsma, E.A., Hulsman, D., Noback, S., Heimerikx, M., Kerkhoven, R.M., Voncken, J.W., Wessels, L.F., and van Lohuizen, M. (2008). Ubiquitin E3 ligase Ring1b/Rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. PLoS One *3*, e2235.

Voncken, J.W., Roelen, B.A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc Natl Acad Sci USA *100*, 2468-2473.

Wagschal, A., Delaval, K., Pannetier, M., Arnaud, P., and Feil, R. (2007). Chromatin Immunoprecipitation (ChIP) on Unfixed Chromatin from Cells and Tissues to Analyze Histone Modifications. CSH Protocols *2007*, pdb prot4767.

Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K.A., Lozach, J., Glass, C.K., and Rosenfeld, M.G. (2008). Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. Mol Cell *29*, 69-80.

CHAPTER 5

Multiplexed, Quantitative Analysis of Cytokine Secretion from Single Pluripotent Cells Reveals Heterogeneous, Secretory Subpopulations Throughout Colony Growth

ABSTRACT

Human pluripotent stem cells exist in culture as organized colonies that rapidly expand, yet they maintain the ability to differentiate into all cell types of the body. This ability is maintained not only by critical components of the medium in which they grow, but also by the crosstalk among individual stem cells via secretion of extracellular cytokines. Here, we utilize a high throughput, single-cell barcode chip to quantify the secretion of several proteins from individual pluripotent stem cells. We demonstrate that human pluripotent stem cell colonies are comprised of subpopulations of cells that possess distinct cytokine secretion patterns. One subpopulation secretes high levels of IGFBP2 and DKK3, while another subpopulation secretes high levels of BMP2 and FGF4 with moderate levels of DKK3. Additionally, the abundance of these subgroups depends on colony size, and they are distinctively localized to specific areas within colonies. Furthermore, the secretion profile of individual cells are not autonomous, and rather influenced by signaling interactions with other cells. Specifically, DKK3 regulates the secretion of BMP2 and FGF4. Together, our data establish a model in which subpopulations of pluripotent stem cells coordinate the maintenance of their pluripotency as they propagate in colonies.

INTRODUCTION

Human embryonic stem cells (hESCs) are derived from the inner cell mass of a developing blastocyst, propagated in *in vitro* culture systems as colonies, and defined by the ability to self renew their pluripotent state (Thomson et al., 1998). The ability to grow hESCs in culture relies heavily on components of the medium in which they are cultured, particularly in the absence of feeder cells, where specific cytokines and growth factors are necessary to maintain the self-renewal of hESCs and prevent spontaneous, terminal differentiation into somatic cells. Furthermore, human induced pluripotent stem cells (hiPSCs), cells that are functionally equivalent to hESCs in terms of self-renewal and pluripotency, are reprogrammed and

maintained in the same culture environment as hESCs (Lowry et al., 2008; Takahashi et al., 2007; Yu et al., 2007). This illustrates that the extracellular environment plays a prominent role in controlling the stability of the pluripotent state and the transitions between the somatic and pluripotency. While the contents of the basal media and supplemental growth factors are reasonably well defined and controllable, the profile of molecules secreted by pluripotent cells are not well understood. Thus the microenvironment, defined as the interface in which hESCs within a colony exchange signals with each other, remains to be further characterized.

Immunofluorescent labeling of hESC surface receptors suggests that the cells expressing those receptors respond to particular signaling pathways (Bendall et al., 2007). However, these methods provide limited insight into whether these pathways are autocrine or paracrine in nature. Bulk analysis of media conditioned by hESCs has been performed with proteomic studies to identify the most abundantly secreted proteins from hESCs (Bendall et al., 2009), but the distinct molecules within a heterogenous mixture of cells cannot be traced back to the secretory cell of origin. Open questions remain as to which cytokines are secreted by hESCs, whether there is coordinated secretion among several cytokines, whether there is a heterogeneity of cells with distinct secretory profiles in the same colony or culture, and what mechanisms control all of the above.

We have previously developed the single cell barcode chip (SCBC), a platform that uniquely enables the high-throughput, quantitative profiling of candidate signaling proteins secreted by a single cell (Wang et al., 2012). To provide a brief description, a fabricated silicone chip has 10,000 microchambers, into which we load a suspension of dissociated, single cells that are isolated into chambers and incubated at 37 °C for 12 hours. At the bottom of each microwell is an antibody array that can detect up to 8 proteins secreted from each cell. Here, we apply single cell analysis through the SCBC chip to pluripotent stem cells. We simultaneously

assay a multiplex panel of proteins secreted from these cells and consequently define subpopulations of cells with distinct secretion profiles. Specifically, one notable subpopulation exhibits high levels of Insulin-like growth factor binding protein 2 (IGFBP2) and Dickkopf-related protein 3 (DKK3) secretion, and another subpopulation secretes Bone-morphogenetic protein 2 (BMP2), Fibroblast growth factor 4 (FGF4), and DKK3. By profiling fibroblasts, hESCs, hiPSCs and differentiated hESCs, we find a pluripotency-specific, single cell secretion pattern for these molecules. Furthermore, we describe the specific localization of these secretory cells within hESC colonies and the amount of protein they secrete, both properties which are dependent on colony size. Finally, we observe that interactions between single, dissociated hESCs can control the secretion profiles of one another, potentially mediated through the function of DKK3. Overall, our study unveils a heterogeneous and dynamic hESC microenvironment and may reflect varying differentiation capabilities of hESCs as colonies grow. This concept will impact future studies dissecting coordinated signaling networks among pluripotent cells and thereby improve our understanding and practice of manipulating these cells for the purposes of regenerative medicine.

RESULTS

Heterogeneous hESCs Secrete Specific Profiles of Cytokines

To observe whether single pluripotent cells have distinct secretion patterns, we profiled the secretion of cytokines that were highly detectable in conditioned medium of hESC populations (Bendall et al., 2009), or whose mRNA expression changes dynamically among human fibroblasts, early and late passage iPSCs, and ESCs (Chin et al., 2009) (Figure S5.1A). In addition, we were limited to commercially available ELISA sets that included a capture antibody, a detection antibody, and a recombinant protein standard. The panel of proteins we assayed consisted of IGFBP2, DKK3, BMP2, FGF4, Transforming growth factor beta 1 (TGFβ1), Insulin-like growth factor 1 (IGF1), and Epidermal growth factor (EGF). Using feeder-

free hESC culture medium that displayed low background signals to our selected cytokines (Figure S5.1B), we cultured H1 hESCs, H9 hESCs, and a hiPSC cell line generated in our lab to near confluency before dissociating the colonies with Accutase into single cells, loading the suspension into the SCBC platform, and incubating them for 12 hours at 37 °C. Primary adult human fibroblasts as well as an H9-hESC derived fibroblast culture were treated similarly.

After the incubation, the protein secretion values from each single cell were determined, normalized, and all data obtained were clustered together to define patterns of protein secretion (Figure 5.1A). Firstly, the most observable feature is that the secretion profiles of the somatic cells are distinct from the pluripotent cells as well as each other (Figure 5.1A, somatic cells are represented by light and dark orange cell type labels). H9-hESC derived fibroblasts (H9dFibroblast) are characterized by relatively high secretion of IGFBP2, DKK3, and EGF, while the primary fibroblasts are characterized by a relatively high secretion of DKK3. The different secretion profiles observed between these two fibroblast cultures probably reflect different cellular states, as our method of *in vitro* differentiating fibroblasts from hESCs did not result in a cell type identical to primary fibroblasts. Secondly, the secretion profiles of the three pluripotent cell lines are distinct from the somatic cell profiles and indistinguishable from each other and therefore cluster together (Figure 5.1A, light and dark purple and black cell type labels). A subpopulation of these pluripotent cells is characterized by relatively high secretion of BMP2 and FGF4. Third, among the pluripotent cell lines, as well as H9-hESC derived fibroblasts, there is a notable co-secretion between IGFBP2 and DKK3. However, their values, when presented as normalized to IGFBP2 and DKK3 levels in fibroblasts, appear low (Figure 5.1A). These observations together suggest that hESCs display a secretion profile that is distinct from somatic cells that can also be re-established once somatic cells are reprogrammed back to the pluripotent state.

To further investigate differences in secretion patterns among the cytokines in our panel, specifically in the pluripotent cells, we focused on organizing the secretion data from H1 and H9 hESCs, and iPSCs into groups of cells with similar secretion patterns (Figure 5.1B). The most evident pattern was revealed by k-means clustering into three groups characterized by 1) high IGFBP2 and DKK3 secretion with low secretion of the other proteins; 2) high level BMP2 and FGF4, middle level DKK3, variable TGFβ1, and low level secretion of the other proteins; and 3) variable secretion of IGF1, EGF, and TGFβ1. Despite each of the three pluripotent cell lines clustering separately from each other, the co-secretion pattern between IGFBP2 and DKK3 and between BMP2 and FGF4 across all three cell types remains very similar (Figure 5.1B). A more direct comparison for co-secretion between these proteins within each cell type was analyzed using Spearmann rank correlation (Figure 5.1C), which supported strong, consistent correlations between IGFBP2 and DKK3 and between BMP2 and FGF4 secretion for single hESCs.

To address whether these heterogeneous hESC populations are arising simply as a result of single cell dissociation or isolated incubation in the micro-chamber environment, we looked for the presence of these subpopulations *in situ* by co-immunostaining hESC colonies for IGFBP2 and FGF4, proteins that share a low correlation (Figure 5.1C). Strikingly, fluorescent signals for each protein appeared in disparate locations within each hESC colony (Figure 5.1D), and not all cells within a colony (marked by the nuclear stain DAPI) are associated with each of these proteins. Moreover, IGFBP2 and FGF4 signals occur independently of one another (Figure 5.1D), lending further support to the existence of heterogeneous subpopulations of IGFBP2 and FGF4-secreting hESCs suggested by the micro-chamber experiments (Figures 5.1A-C). We confirmed these findings by staining cultures with or without prior treatment of a protein transport inhibitor cocktail (PTI or CTRL, respectively), which disrupts the ability of cells to secrete proteins (Figure 5.1E). Untreated cells display FGF4 staining patterns resembling the

Golgi apparatus and moderate levels of IGFBP2 staining in a less specific focal pattern (Figure 5.1E, CTRL). A four hour treatment with the protein transport inhibitory cocktail, which includes Brefeldin A, a small molecule known to disrupt the Golgi apparatus (Bershadsky and Futerman, 1994; Rosa et al., 1992), dramatically eliminated FGF4-stained structures while enriching for IGFBP2 stains (Figure 5.1E, PTI). This effect by PTI treatment suggests that the immunostaining signals we observe mark cells secreting IGFBP2 and FGF4.

Another potential explanation for the heterogeneous secretion observed from hESC and hiPSC cultures is that the pluripotent cells could be spontaneously differentiating in our cultures. Thus the different secretory profiles we observe are possibly from the hESC- and hiPSC-derived differentiated cells. However, fluorescence-activated flow cytometry of hESCs grown in our culture conditions indicate that the overwhelming majority of single cells analyzed are positive for cell surface markers associated with pluripotency, including the tetraspannin protein TG30, the Insulin-like growth factor 1 receptor (IGF1R), and Stage-specific embryonic antigen 4 (SSEA4) (Adewumi et al., 2007; Bendall et al., 2007) (Figure S5.1C). Furthermore, all hESCs grown in our culture conditions immunostain for the master transcription factors regulating pluripotency, OCT4 and NANOG (Figure 5.2A and data not shown). Taken altogether, these results demonstrate that a heterogeneity in secretory cells exists within the pluripotent cell pool.

Localization of hESC Subtypes Changes as Colony Size Increases

We considered that the observed heterogeneity in hESC cultures may change with colony size. While we observed that all cells within hESC colonies are OCT4-positive, we also noted that the staining intensity of IGFBP2 differed between large colonies, defined as consisting of greater than 100 cells, and small colonies, defined by less than 100 cells. Closer inspection of various image fields revealed that smaller colonies display a higher intensity of IGFBP2 staining (Figure 5.2A, bottom panel) compared to large colonies (Figure 5.2A, middle

panel). We also noted that small colonies expressed a higher level of IGF1R, a marker enriched on hESCs (Bendall et al., 2007) and a receptor for the IGF proteins, which are regulated by IGF-binding proteins (Clemmons et al., 1995), suggesting that the IGF pathway is highly regulated in small colonies. To quantify this trend of IGFBP2 enrichment in smaller colonies, imaging analysis of fluorescent signal intensities along linear segments across large and small colonies highlighted a stronger IGFBP2 signal across cells in small colonies compared to cells in large colonies (Figure 5.2B, lower panels). Conversely, FGF4 signal intensity is higher in large colonies compared to small colonies (Figure 5.2B, middle panels). Together, these data suggest that the secretion patterns of these proteins change as hESCs progress from small to large colonies, wtih IGFBP2 secretion becoming reduced and FGF4 secretion increased.

In addition to the intensity of these protein signals, we further investigated if the heterogeneous localization of IGFBP2 and FGF4 within a colony are correlated to colony size. For large colonies, we observed that strong IGFBP2 signals preferentially occurred at the edge of the colonies relative to the central colony (Figure 5.1D). Small colonies displayed strong IGFBP2 signals throughout. In contrast, FGF4 signals were evenly distributed throughout both large and small colonies (Figure 5.1D). This suggests IGFBP2-secretion initially occurs across all cells in small colonies and, as small colonies grow large, the central cells of the colony adopt a secretion profile with reduced IGFBP2. However, the ability of cells to secrete FGF4 increases with colony size. To illustrate this trend, multiple immunofluorescent images of colonies ranging from about 10 to 1000 cells per colony were quantified for the number of objects staining positive for IGFBP2 or FGF4 and correlated to the number of cells present in each hESC colony (Figure 5.2C). These data strongly support the conclusion that as the number of cells in a colony increased, the percentage of IGFBP2-secreting cells within the colony decreases, consistent with a decreasing ratio of edge cells to central cells as colonies increase in size. In contrast, the ratio of FGF4-secreting cells to all cells in a colony remained constant. The

preferential localization of IGFBP2 secretion at the colony edge is also observable in Figure 5.2B (lower right panel), where the normalized, linear segments (bins) representing the edge of all colonies analyzed show a higher peak relative to the bins representing the central area of all colonies analyzed. Consistent with its staining pattern, FGF4 does not show a higher intensity at the colony edges (Figure 5.2B, middle panels).

While the averaging of staining intensity across colony cross sections reveals a gross distribution pattern of IGFBP2 and FGF4, we looked in finer detail as to whether IGFBP2 and FGF4 demonstrate nonrandom patterns with respect to the center and edge of a single, representative large colony (Figure 5.2D). We also examined the pattern of DKK3 staining, whose secretion appeared to correlate with that of IGFBP2 (Figure 5.1B/C). To this end, we plotted fluorescence intensities in concentric, radial segments, normalized to the surface area of the segment and the number of cells within that segment (Figure 5.2D/E). The moving average line for IGFBP2 and DKK3 track each other remarkably well, lending further support to their correlative secretion patterns discussed in Figure 5.1B and 5.1C, and showed an dramatic increase towards the periphery of the colony. On the other hand, FGF4, did not display such an increase. Interestingly, we noted periodic patterns of increased fluorescent signals for all three proteins moving from the center to the edge of the colony. These peaks suggest that IGFBP2, DKK3 and FGF4 are secreted from a population of nonrandom, regularly spaced cells and act locally within the colony. Altogether, these observations illustrate that the heterogeneity of protein secretion in hESC cultures differs with colony size, and particularly that large colonies are very heterogeneous.

Intercellular Influence on the Secretion Profile occurs in Large Colonies, but not in Small Colonies

The nonrandom, periodic positioning of single hESCs secreting IGFBP2, DKK3 or FGF4 suggested that these cytokines act in a local manner. We therefore asked what mechanisms could dictate that one cell in a colony takes on a specific secretion profile. We hypothesized that the ability of cells to sense the presence of one another is an instructive mechanism regulating the identity and output of their secretory function. Noting the differences in secretion patterns between large and small colonies as determined by immunostaining, we compared the secretion profile of single hESCs isolated from either large or small colonies, utilizing the SCBC (Figure 5.3A). Since the loading of the single cell suspension into the microchambers is random, a proportion of the microchambers will end up with two cells that landed there by chance, thereby providing a protein secretion data set for when two cells were co-incubated. Based on the protein secretion data set from microchambers with one cell, which are obtained from the same SCBC experiment, a simulation can be performed to produce a secretion profile modeling the random pairing of single cells. If we assume that two secretory cells bear no influence on the secretory capabilities of one another when incubated in the same microchamber, we would expect that the protein secretion readout would simple be a sum of the secretory profiles of the two cells. Alternatively, if one cell affects the secretory capabilities of another cell, then the protein secretion values obtained from the experimental two cell chambers would show a data distribution that differs from the null distribution produced by the simulation.

We first assayed the effects of single cell pairing on the magnitude of protein secretion. Comparing the distribution of IGFBP2 molecules that were actually detected in two cell chambers for small hESC colonies (exp) versus simulated two cell pairings (sim), there is no notable difference (Figure 5.3B, top panel). However, comparing the experimental two cell data to the simulated data for hESCs from large colonies, there is a notable increase in IGFBP2 secretion (Figure 5.3B bottom panel). This result indicates that the number of IGFBP2 molecules secreted by a cell in a large colony can be modulated in the presence of another cell,

but this property is not apparent in small colonies. Similar relationships were not found for any of the other tested proteins (Figure 5.3B).

We next assayed the ability of two cell pairs to affect, in each other, the coordinated secretion of the proteins in our panel. To do so, we determined the secretion pattern of our 7 candidate proteins when two cells were together in a chamber, both for small and large colonies. Specifically, we performed unsupervised hierarchical clustering of the SCBC data, either from chambers that incubated one cell, two cells, or for the simulated data for two cells based on the secretion data for one cell, for either small or large colonies (Figure 5.3C). For the one cell data from small colonies (Figure 5.3C top, left), we observed two prominent groups of cells. One group of single cells secretes high levels of IGFBP2 and DKK3 and is typically silent for BMP2 and FGF4, which we classified as the red group. The criteria for this grouping can also include chambers with detectable EGF, IGF1, and TGFβ1. We also noted another group of single cells characterized by BMP4 and FGF4 secretion that also have moderate levels of DKK3 secretion, classified as the blue group. This group can also include detectable EGF, IGF1, TGFβ1, and IGFBP2.

The red group comprises 36.56% of cell of small colony cultures and correlate IGFBP2 and DKK3 secretion highly, as indicated by the tight clustering of the red labels, while the blue group comprises 13.89% of the cells from small colony cultures and also correlate with each other. The clustered secretion data for the one cell chambers from large colonies, when compared to those from small colonies, retains the presence of the red and blue groups of cells, however, the percentage of the red cluster is drastically reduced while the blue cluster is expanded (Figure 5.3C, bottom, left). This is consistent with the results described in Figure 5.2C, where the number of FGF4 secreting cells increase with colony size while the number of IGFBP2 secreting cells decreases.

Clustering of the simulated two cell chamber data set from small colonies illustrated the expected organization pattern of the proteins and chambers (Figure 5.3C, top right), as well as the expected percentage of the red and blue groups of cells if the single cell data from this experiment are randomly combined. The simulation result follows the assumption that pairing of two single cells will not change their individual secretion profiles, thus the resulting secretion profile generated from this simulation would be the summed values for each protein from both cells. Clustering the experimental two cell chamber data set and comparing it to the simulated two cell data, we found that the experimental data look very similar to what was expected by the simulation (Figure 5.3C, top middle), including the percentage of red and blue cells. These findings suggest that hESCs in small colonies are not competent to change their secretion profile in response to nearby cells.

The simulation of two cell data from large colonies shows that the red and blue groups do not display a coordinated secretion pattern, as they do not cluster tightly together (Figure 5.3C, bottom, right). In contrast to the results from the small colonies, the experimental two cell chamber clusters (Figure 5.3C, bottom, center) demonstrate a dramatically expanded percentage of both red and blue groups compared to the one cell experimental and two cell simulation data. This suggests that hESCs in large colonies are competent to change their secretion profiles in response to nearby cells. Furthermore, the tight clustering of red and blue groups in the experimental two cell data indicates that two cell pairings result in very similar secretion profiles, compared to their behavior in the simulation. Taken together, these data argue for a competency of hESCs in large colonies to induce and respond to each other to affect their secretion profiles. This trait is lacking in small colonies, at least for the secretion of proteins in our panel.

DKK3 Activates BMP2 and FGF4 Secretion

Next, we tested the possibility that any one of the proteins we assayed for could be the effector responsible for the profile conversion observed from the one cell to the two cell data sets for large colonies. We were particularly interested in how the addition or depletion of IGFBP2, DKK3, BMP2, or FGF4 to hESC cultures could affect the secretion of IGFBP2, DKK3, BMP2, and FGF4. To do this, we treated H1 or H9 hESC cultures with recombinant proteins or neutralizing antibodies over the course of four to six days. We assayed the secretion of IGFBP2, DKK3, BMP2, and FGF4 proteins in the media conditioned by these cultures by taking samples every two days (Figure 5.4A). As an initial screen, assaying protein secretion in conditioned medium from bulk hESC cultures was sufficient to address this question.

The conditioned medium obtained from the control hESC cultures (CTRL) indicate that DKK3 and IGFBP2 secretion increase over the course of four days while BMP2 and FGF4 secretion remain relatively static (Figure 5.4B). Out of the large set of secretion values (data not shown), the most consistent trend of effects between H1 and H9 hESCs were associated with DKK3 stimulation and neutralization. In conditions where recombinant DKK3 protein (rDKK3) was added to the culture, IGFBP2 secretion was not affected, whereas BMP2 and FGF4 secretion increased (Figure 5.4B). This indicates that rDKK3 is sufficient to induce secretion of BMP2 and FGF4 in hESC cultures.

Conversely, BMP2 and FGF4 secretion markedly decreased when a neutralizing antibody to DKK3 (αDKK3) was added to the hESC culture, suggesting that DKK3 is necessary for BMP2 and FGF4 secretion (Figure 5.4C). Consistent with Figure 5.4B, IGFBP2 secretion continuously increases and is not affected by neutralizing DKK3. The immunoglobulin isotype controls (IgG) did not diverge from the control conditions, indicating that the effect of neutralizing DKK3 was specific. There were no signs of overt differentiation or cell death, as judged by

morphology, in any of these conditions (data not shown). In contrast, we noted that treatment with recombinant BMP2 resulted in rapid differentiation of hESCs within the first two days (data not shown), consistent with previous reports (Pera et al., 2004; Valera et al., 2010). Because ectopically applied rDKK3 upregulated BMP2 secretion but did not induce rapid differentiation, this suggests that either the BMP2 levels reached here were not high enough to cause differentiation, or there are other, unobservable mechanisms at play to buffer the effects of upregulated BMP2 signaling. Given that i) the secretion of BMP2 and FGF4 changes when DKK3 levels are altered (Figures 5.4A-C) and ii) the two cell interactions are able to expand the percentage of BMP2 and FGF4 secreting cells in a population (Figure 5.3C), these data suggest that DKK3 is a potential mediator of this conversion of secretion profiles in large hESC colonies.

DISCUSSION

Using a single cell based protein secretion assay, we identified subpopulations of hESCs that are defined by distinct secretion profiles. hESCs have a secretion profile distinguishable from differentiated cells that is re-established upon reprogramming to iPSCs. We defined two hESC subpopulations: one secreting high levels of IGFBP2 and DKK3, and another secreting BMP2, FGF4, and moderate levels of DKK3. Using fluorescence microscopy, we confirmed that these subpopulations exist within hESC colonies and further dissected the localization of these subpopulations with respect to colony structure. Further analysis revealed that the amount of protein that each cell secretes, as well as the localization of distinct secretory cells are affected by colony size. The SCBC data obtained from small and large colonies reveal that the population of IGFBP2 and DKK3 secreting cells constitutes a majority of small colonies, but decreases in large colonies. Conversely, the population of BMP2, FGF4, and DKK3 secreting cells increases with colony size. Utilizing the capacity of the single cell barcode chip to capture multiple cells per chamber, we were able to obtain data for proteins secreted from two cells incubated in the same chamber, and infer the effects of two cell interactions based on

comparisons to modeled data from one cell chambers. We found disparate effects between cells isolated from small colonies and cells isolated from large colonies, where cells from the latter changed the secretion profile and magnitude in response to direct cell-to-cell influence. Finally, we reveal DKK3 to be a regulator of BMP2 and FGF4 secretion, suggesting a role in coordinating protein secretion in large colonies.

Integrating of all the data presented here, a model for the transition of hESCs from a small colony to a large colony can be constructed (Figure 5.4D). Small colonies are characterized by a dominant subpopulation of cells secreting high levels of IGFBP2 and DKK3 (red cells). We also noted that these colonies expressed high levels of IGF1R, a receptor involved in the PI3-AKT survival pathway in hESCs (Bendall et al., 2007), suggesting that cells at this stage in culture are highly regulated by the IGF pathway. A minor subpopulation of cells secreting BMP2, FGF4, and DKK3 (blue cells) are also present, as well as other cells that do not fall into this category (grey cells). As small colonies grow and become larger, the localization of secretory subpopulations takes on a distinct orientation. Cells secreting high levels of IGFBP2 and DKK3 remain predominantly associated with the colony periphery while DKK3+/BMP2+/FGF+ cells are regularly interpersed throughout the colony.

Considering the previously reported functions of BMP2 and FGF4 in regulating hESC differentiation and self-renewal (Johannesson et al., 2009; Mayshar et al., 2008; Pera et al., 2004; Valera et al., 2010), BMP2 and FGF4 secretion at this stage must be critically regulated for the large colony to sustain the pluripotent state until the point of passaging, where it is dissociated into a small colony, and the colony growth cycle is renewed. The appearance of the BMP2 and FGF4 subpopulation may arise from cells that originally secreted DKK3 and IGFBP2 in the small colony, and this transition is perhaps mediated through autocrine or paracrine DKK3

signaling. One may speculate that this phenomenon represents an inherent mechanism of the inner cell mass to poise itself for developmental cues as it grows *in vivo* (Peerani et al., 2007).

The DKK family of proteins have classically been described as inhibitors of the Wnt signaling pathway (Glinka et al., 1998). While the Wnt signaling pathway has been implicated in directing differentiation of hESCs (Davidson et al., 2012; Singh et al., 2012), no function for DKK3 in regulating the self renewal of hESCs has previously been described. Furthermore, of the four mammalian homologs, DKK3 is distinct from its other family members in that it does not act to inhibit Wnt signaling through the classical Wnt receptor LRP6 (Davidson et al., 2002; Mao et al., 2002), but likely through some other pathway (Nakamura and Hackam, 2010). Thus, it remains unknown why among the four family members, DKK3 is the most highly secreted protein in hESCs (Bendall et al., 2009).

Based on our findings, several open questions remain. More functional and mechanistic experiments are needed in order to address the following: How are IGFBP2 and DKK3 secretion regulated? Does IGFBP2 affect the activity of the IGF signaling pathway? How does DKK3 affect BMP2 and FGF4 secretion? Are these mechanisms controlling secretion plasticity found at the transcriptional or post-transcriptional level? Furthermore, while we observed that two-cell interactions can affect secretion dynamics, the identification and characterization of the effector cell and the affected cell, as well as the signaling pathway through which it acts, remains to be investigated. mRNA expression profiling of small and large colonies will provide more data driven approaches to identify genes and pathways that can fit the model driven by our secretion data. Single-cell mRNA expression profiling technology can be considered as an alternative approach if the mRNA expression profiles from bulk hESC colonies are not informative. An important additional question is whether the subpopulations of large hESC colonies differ in their differentiation potential. Finding surface markers through the RNA-seq approach that

characterize these different cell populations should be instrumental towards purifying these populations and testing their differentiation potential.

We note that the hESCs secreting EGF, IGF1, and TGFβ1 did not reveal themselves as cells belonging to a readily identifiable subpopulation. This by no means indicates that EGF, IGF1, and TGFβ1 secretion patterns are not coordinated with other cytokines that we did not assay in our system. Future experiments with the SCBC platform may yield insightful associations for these proteins. In summary, our study begins to functionally dissect heterogeneous signaling networks occurring between hESCs within the same colony and how these networks may impact colony growth coupled with the maintenance of self renewal and pluripotency.

EXPERIMENTAL PROCEDURES

Single Cell Barcode Chip (SCBC) Design and Fabrication

The SCBC is composed of a single elastomer microfluidics layer bonded on top of a barcode-patterned glass slide. Details of microchip design and fabrication can be found in our previous publication (Wang et al., 2012). In brief, standard photolithographic techniques were used to generate hard molds with 3D features, which were translated into microchambers and microchannels in the elastomer microfluidics layer after casting. Each chip has 10,825 microchambers, each of which allows an independent secretion profile analysis. For on-chip experiments, cells are randomly loaded to microchambers and counted under a microscope through transparent microchip. With a cell concentration at 3×10^5 / ml for loading, we can achieve ~2000 single cell measurements, ~500 two cell measurements, ~200 three cell measurements and so on.

Antibody Microarray

The microarrays are initially patterned as DNA arrays, since DNA has the physical and chemical stability to withstand the various processing steps of microfluidics fabrication. The DNA itself is patterned onto a poly-lysine coated glass slide by two sequential microchannel-based flow patterning steps. The two microchips for those two steps are based on polydimethylsiloxane (PDMS) elastomer sealed with a glass slide with 20 microchannels winding from one end to the other end of the glass slide. The first flow patterning step generates 20 μm wide, 50 μm pitch lines of 3 unique DNA oligomers, while the second DNA patterning step is carried out at right angles to the first. We design the ssDNA sequences for the first and second patterning in such a way that the intersection of the two sets of lines will remain a unique ssDNA for assembly location for complementary ssDNA - antibody conjugates. In the current study, we chose 3 by 3 array, which has the capacity of multiplexed measurement of up to 8 different proteins. This addressable 9 element array has been repeated for ~ 19,000 times with various orientations across the whole glass slide, which will be assembled with a SCBC elastomer layer.

We validated the patterned DNA array using 100 nM cy5 conjugated complementary DNA oligomers in 1% bovine serum albumin (BSA) in PBS. After incubation for 1 h at 37 °C, the slide was rinsed with 3% BSA in PBS twice and DI water, and dried under flowing N2 before signal readout by a Genepix scanner (Molecular Devices, LLC.). The validation procedure provides a check on the cross-reactivity between the anchor, bridging, and terminal ssDNA oligomers. In addition, the fluorescence intensity per unit area can be compared against standard DNA spotting approaches as a means of gauging surface coverage. Finally, it provides an assessment of the fidelity of the microfluidics flow-patterning steps.

The DNA array was converted into antibody array using complementary ssDNA-antibody conjugates. The procedure of conjugation has been described before (DNA-Encoded Antibody Libraries: A Unified Platform for Multiplexed Cell Sorting and Detection of Genes and Proteins.

2007, JACS). Calibration and cross-reactivity characterizations of the barcode assay have been performed before application to single cells. We used recombinant EGF (RnD 236-EG), TGFβ1 (RnD 240-B), IGFBP2 (RnD 674-B2), IGF1 (RnD 291-G1), BMP2 (RnD 355-BM), FGF4 (RnD 235-F4) and DKK3 (RnD 1118-DK) proteins (R&D Systems, Inc.) at various concentrations to calibrate the fluorescence intensity of a barcode assay with molecule number in a cell chamber with a volume of 0.87 nl. Antibody cross-reactivity assays were carried out using spotted arrays and identical biomolecular reagents.

Cell Sample Preparation

Single cells were suspended at 3×10^5 / ml in ReproFF2 (ReproCell) medium supplemented with 20 ng/ml bFGF after treating cells in a flask with Accutase (Stem Cell Technologies). For the SCBC profile experiments, the following cells were used: H1 and H9 hESC cell lines, one iPSC line generated by the UCLA Stem Cell Core facility, a primary human fibroblast line from the mother of a Lesch-Nyhan patient, and *in vitro* differentiated fibroblasts from the H9 ESCs. Media sampled include ReproFF2 supplemented with 20 ng/ml bFGF, mTeSR (Stem Cell Technologies), and mouse embryonic fibroblast (feeder) conditioned standard hESC media, which contains 20% Knockout Serum Replacement, and supplemented with PenStrep, glutamine, non-essential amino acids, 2-Mercaptoethanol, and 20 ng/ml bFGF. Pluripotent cells grown in Matrigel–coated wells and ReproFF2 were cultured according to manufacturer's instructions, and were passaged weekly using Dissociation Solution (ReproCell). To differentiate H9 hESCs into Fibroblasts, ReproFF was replaced with DMEM containing 20% fetal bovine serum (FBS) for one week, passaged with Trypsin, and cultured in FBS supplemented media for an additional week. To inhibit protein transport, cells were grown in media supplemented with PTI (eBioscience 00-4980) according to manufacturer's instructions for four hours prior to fixation.

For secretion analysis of bulk conditioned hESC media, hESCs were passaged regularly, plated for one day, and cultured for two days in ReproFF2 supplemented with 20 ng/ml bFGF, which served as the baseline media. Thereafter, treatment was started on cells, and conditioned media was collected every two days, 0.22 μm filtered, and frozen until all timepoints were completed. Samples were thawed in parallel and assayed on spotted arrays. For ectopic stimulation with recombinant DKK3 or neutralization with anti-DKK3 antibody, the following were used at the indicated concentrations: 2 μg/ml DKK3 (RnD 1118-DK) and 2.14 μg/ml goat anti-DKK3 (Abcam ab2459), with PBS or 2.14 μg/ml normal goat IgG (AB-108-C) used as a respective controls. Conditioned media were normalized to their respective treatment media that had no contact with any cells, but placed in tissue culture vessels and incubated at 37°C in parallel with experimental cell cultures.

Immunostaining, Imaging Analysis, Flow Cytometry, and Gene Expression Data.

For immunostaining, cell lines were grown on glass coverslips coated with Matrigel, fixed with 4% paraformaldehyde (PFA) in PBS for 10 min, permeabilized with 0.5% Triton 100-X in PBS for 5 min, and immunostained with mouse anti-FGF (R&D Systems MAB635), rabbit anti-IGFBP2 (Abcam ab124930), chicken anti-IGF1R (Abcam ab32823), mouse anti-OCT4 (Santa Cruz Biotechnology sc-5279), or mouse anti-DKK3 (R&D Systems MAB11181). Images were captured on Zeiss Axio Imager M1 Microscope using Axiovision Imaging Software. Quantitative fluorescence analysis was performed on Image J using line region of interest tool set to a width of 50 pixels. Intensity data normalized using R software by setting histogram to 100 bins and averaging the intensity of all pixels within a bin. To quantify fluorescent objects in Image J, colonies selected in grey scale channel images were binarized to black and white objects using intensity thresholds, and the Analyze Particles method was used to count objects within the region of interest greater that 5 pixels^2. Object counts in the FGF4 and IGFBP2 channels were normalized to the number of DAPI objects in the corresponding region of interest. For flow

cytometry, cells were accutase treated, washed with PBS solution containing 0.5% bovine serum albumin and 2mM EDTA, pelleted, and resuspended in wash solution containing the following antibodies for 15 minutes at 4°C: mouse anti-TG30 (Millipore MAB4427), mouse anti-IGF1R (BD Pharmingen 555998), mouse anti-SSEA4 (kindly provided by the laboratory of Amander Clark). After a wash, a 15 minute incubation in wash solution containing secondary goat anti-mouse antibody conjugated with FITC, and a subsequent wash, cells were sorted on a FACSDiva (BD Biosciences) and FACS plots were analyzed using Flowjo software. Microarray expression data obtained from (Chin et al., 2009), and values represent average gene expression from three cell lines representative of each cell type.

On-Chip Assay and Quantification

The PDMS layer of SCBC chip is mated onto a barcode glass slide. The microchambers are aligned with the microarray so that each microchamber is ensured of overlapping a complete barcode. Then the microchambers and microchannels were blocked with 3% BSA in PBS for 1h before filling a cocktail of 10 μg/ml antibody-DNA conjugates. After incubation for 1 h and removal of unlinked conjugates, the SCBC was filled with a cell sample, and then was mounted onto a clamp. With appropriate pressure exerted by the clamp, all microchambers were closed where random number of cells resided in each microchamber. The whole structure was incubated for 12 h at 37 °C in a 5% CO2 incubator. The clamp was then released to form the open state, and biotinylated detection antibody cocktails in 1:18 dilution prepared according to product instruction were injected into the microchannels. After washing off unbound antibodies, 1 μM cy5-Streptavidin and 100 nM H-cy3 in PBS with 3% BSA were used to complete sandwich-ELISA procedure and label barcode arrays. Finally, the glass slide containing the developed arrays was detached and cleaned for scanning by a Genepix Scanner. Fluorescence intensity of each barcode across the whole slide was digitalized and exported to a custom

algorithm written in Matlab (MathWorks, Inc.). The intensity can be further converted into protein copy number using a calibration curve.

Data Processing

Cluster 3.0 and Treeview (Michael Eisen) were used to cluster subpopulations and plot heatmap. The quantification of fluorescence spatial distribution was achieved by a Matlab algorithm. Statistical analysis was performed using laboratory developed Matlab programs and Prism (GraphPad Software). To compare protein levels under various conditions, we obtained the mean protein number, and used *P* value to evaluate the whether the difference was significant through two-tailed Student's *t* tests assuming unequal variance. A *P* value less than 0.05% is considered statistically significant and denoted with *, while ** and *** represent $P<0.01$ and $P<0.001$, respectively. The error bars for mean protein levels represent the standard error of the mean, which is the standard deviation of the mean protein level divided by square root of cell number. This type of analysis was carried out to establish the statistical uniqueness of the 0, 1, 2, and 3 cell data sets, as well as the uniqueness of the different 2 cell data sets, as a function of cell-cell separation.

Two-Cell Distribution Simulation

We simulated the protein levels of two cells from one cell data to interrogate whether two cells together in one microchamber were influenced by each other. Two single cell profiles randomly selected from one cell secretion data were added to compose a simulated two cell secretion profile, assuming no influence on protein copy numbers due to cell interactions between those two single cells. The simulated two cell secretion data were then compared with experimental two cell data, either through dot plots or heatmap after clustering.

Figure 5.1. Heterogeneous hESCs Bear Distinct Secretion Profiles

A. Unsupervised hierarchical clustering of $\log_2$ secretion ratios from H1 and H9 hESCs, hiPSC, primary human fibroblasts, and H9-hESC derived fibroblasts (H9dFibroblast) relative to the median intensity of each chamber (column centered). Each row represents a chamber in which one cell was assayed. Right-most column indicates the corresponding cell type of that chamber.

B. K-means clustering of $\log_2$ secretion ratios of H1 and H9 hESCs, and hiPSC relative to the median z-score of each chamber and each protein (column and row centered). Each row represents a chamber in which one cell was assayed. Right-most column indicates corresponding cell type of that chamber.

C. Spearman correlation coefficients for protein pairs whose respective chambers were ranked based on secretion values in either H1 (top) or H9 (bottom) hESCs. The correlation values indicate the similarity between the corresponding proteins listed above and to the right of each value.

D. Co-immunostaining of FGF4 and IGFBP2 in H9 hESC colonies. DAPI staining marks nuclei. Immunoglobulin isotype was used as a negative control for primary antibodies, depicted on the left. Images captured at 50x magnification.

E. As in (D), except that the hESC culture was treated with protein transport inhibitor (PTI) 4 hours prior to fixation and staining. CTRL indicates control, untreated cultures. Images were captured at 200x magnification, and further enlarged digitally to reveal the subcellular staining pattern. The image suggests that PTI treatment disrupts the trans-Golgi network structure outlined by FGF4 staining in CTRL while increasing intracellular IGFBP2.

See Figure S5.1 for additional information.

Figure 5.1

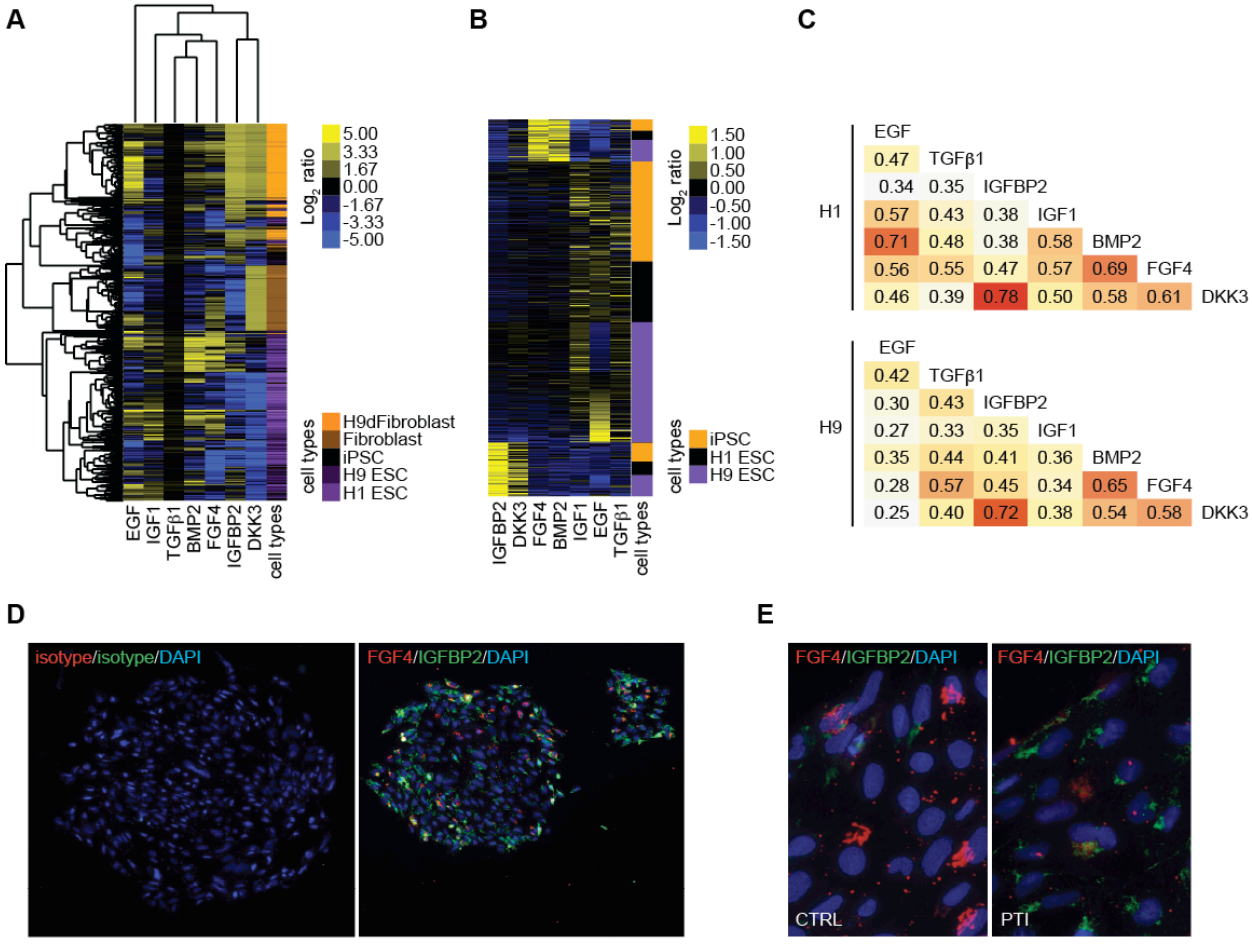

A
Log2 ratio
5.00
3.33
1.67
0.00
-1.67
-3.33
-5.00
cell types
H9dFibroblast
Fibroblast
iPSC
H9 ESC
H1 ESC
EGF
IGF1
TGFβ1
BMP2
FGF4
IGFBP2
DKK3
cell types
B
Log2 ratio
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
cell types
iPSC
H1 ESC
H9 ESC
IGFBP2
DKK3
FGF4
BMP2
IGF1
EGF
TGFβ1
cell types
C
H1
EGF
0.47 TGFβ1
0.34 0.35 IGFBP2
0.57 0.43 0.38 IGF1
0.71 0.48 0.38 0.58 BMP2
0.56 0.55 0.47 0.57 0.69 FGF4
0.46 0.39 0.78 0.50 0.58 0.61 DKK3
H9
EGF
0.42 TGFβ1
0.30 0.43 IGFBP2
0.27 0.33 0.35 IGF1
0.35 0.44 0.41 0.36 BMP2
0.28 0.57 0.45 0.34 0.65 FGF4
0.25 0.40 0.72 0.38 0.54 0.58 DKK3
D
isotype/isotype/DAPI
FGF4/IGFBP2/DAPI
E
FGF4/IGFBP2/DAPI
CTRL
FGF4/IGFBP2/DAPI
PTI

Figure 5.2. Secretion Intensity and Localization As a Function of Colony Size

A. Co-immunostaining for IGF1R, IGFBP2, and OCT4 in H9 hESC colonies. DAPI staining marks nuclei. Immunoglobulin isotype was used as a negative control for primary antibodies, depicted on the top panels. Images captured at 200x magnification.

B. Average fluorescence intensity profiles along multiple lines bifurcating hESC colonies co-immunostained with FGF4 and IGFBP2. “n” indicates number of lines calculated, “c” indicates number of colonies assayed. All lines approximately represent colony diameters and encompass 2 colony edges. The length of all lines were set to 100 bins, values within each bin were averaged, and the summed averages of all corresponding bins in each line are plotted along the y-axis. Bins 0-20 and 80-100 represent average staining intensities at the colony edge. Small and large colonies were distinguished by nuclei number, labels on the right side of each panel indicate the stain that was measured.

C. Scatterplot of the ratio of nuclei per colony versus FGF4 or IGFBP2 stained objects. 32 hESC colonies, each composed of a variable number of DAPI stained nuclei, were quantified in Image J for the number of objects demarcated by either FGF4 or IGFBP2 staining. The number of these objects was normalized to the number of nuclei within the region of interest, and the resulting ratios are plotted against the corresponding number of nuclei. See Experimental Procedures for more details.

D. Co-immunostaining for IGFBP2, DKK3, and FGF4 in H9 hESC colonies. DAPI staining marks nuclei. Immunoglobulin isotype was used as a negative control for primary antibodies, depicted in the left panel. Images captured at 50x magnification.

E. Average fluorescence intensity per surface area per nuclei. Concentric rings with increasing radii demarcated exclusive areas over a representative ESC colony in which we quantified the total fluorescence intensity for each immunostain. Total fluorescence intensity for each radial section was normalized by the total area and the number of DAPI stained nuclei present within the radial section. The resulting value is plotted against the radius. Best-fit lines represent the the moving average of 4 data points.

Figure 5.2

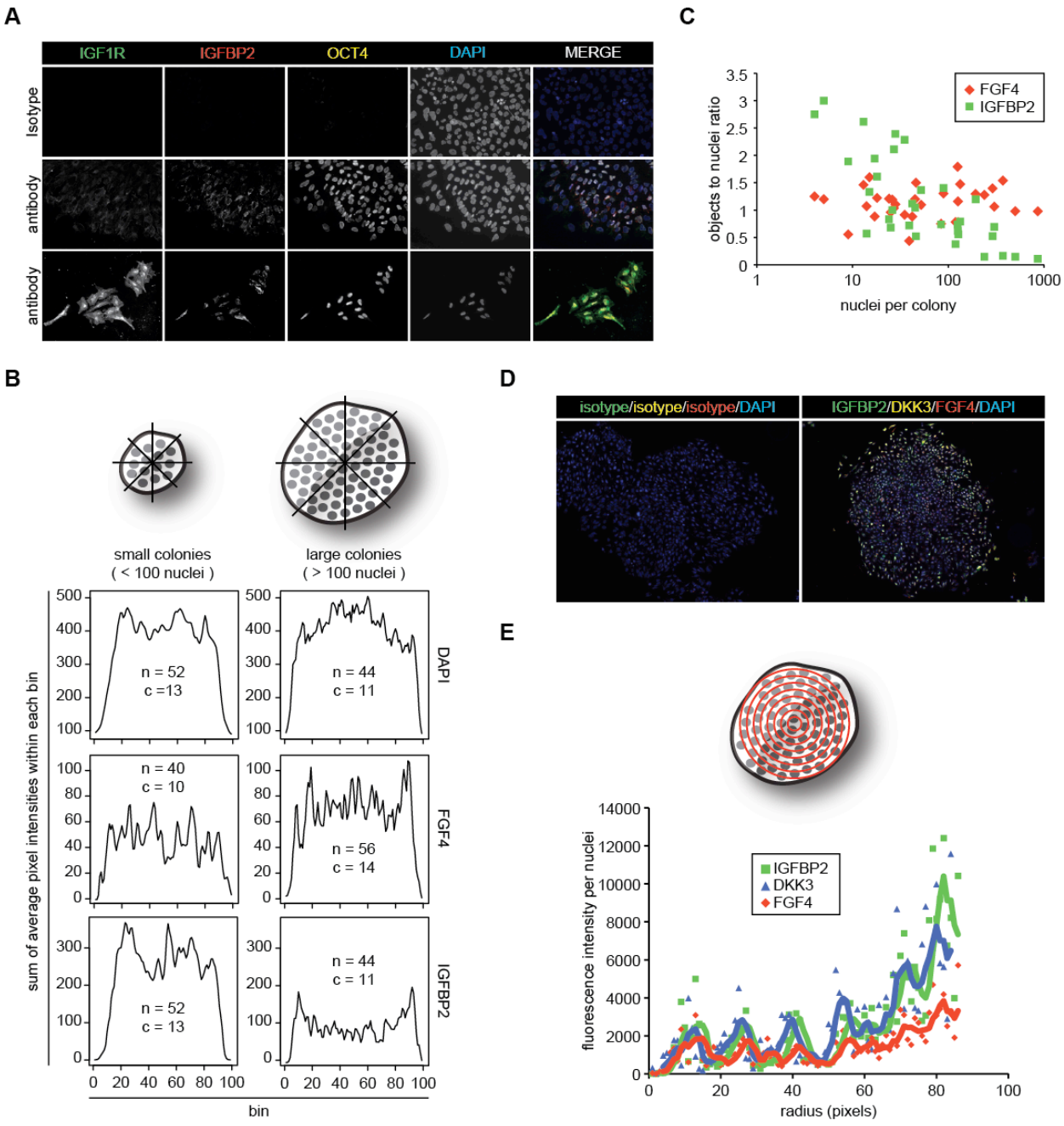

A
IGF1R
IGFBP2
OCT4
DAPI
MERGE
Isotype
antibody
antibody
C
objects to nuclei ratio
nuclei per colony
FGF4
IGFBP2
B
small colonies
(< 100 nuclei)
large colonies
(> 100 nuclei)
sum of average pixel intensities within each bin
DAPI
FGF4
IGFBP2
n = 52
c =13
n = 44
c = 11
n = 40
c = 10
n = 56
c = 14
n = 52
c = 13
n = 44
c = 11
bin
D
isotype/isotype/isotype/DAPI
IGFBP2/DKK3/FGF4/DAPI
E
fluorescence intensity per nuclei
radius (pixels)
IGFBP2
DKK3
FGF4

Figure 5.3. Direct Intercellular Interaction Influences Secretion Profiles and Intensity

A. Representative bright field images of H9 hESCs in small or large colonies prior to loading into micro-well chambers.

B. Plots comparing the distribution of secreted proteins detected in chambers in which 2 dissociated cells were assayed (exp) versus the distribution of the same proteins based on simulated data from experimental 1 cell data (sim), where secreted protein values from a 1 cell per well data set are randomly paired together. Values produced by the simulated experiment are under the assumption that the secretion intensity or profile of a cell is not changed due to pairing with another cell. Blue horizontal bars represent the median value of the distribution. The y-axis is in logrithmic scale.

C. Unsupervised hierarchical clustering of $\log_2$ secretion ratios from H9 hESCs relative to chambers with no detectable protein secretion. Each row represents a chamber in which either one cell was assayed (left panels), two cells were assayed (middle panels), or two cells were simulated from the experimental one cell data set (right panels). Experiments from small hESC colonies are on the top panel, from large colonies bottom panels. Right-most column of each heat map indicates the secretion profile of that chamber, and numbers indicate percentage of that subgroup within the data set. Red: chambers in which both IGFBP2 and DKK3 are detected and BMP2 and FGF4 are not detected. This group can also include detectable EGF, IGF1, and TGFβ1. Blue: chambers in which DKK3, BMP2, and FGF4 are detected. This group can also include detectable EGF, IGF1, TGFβ1, and IGFBP2.

Figure 5.3

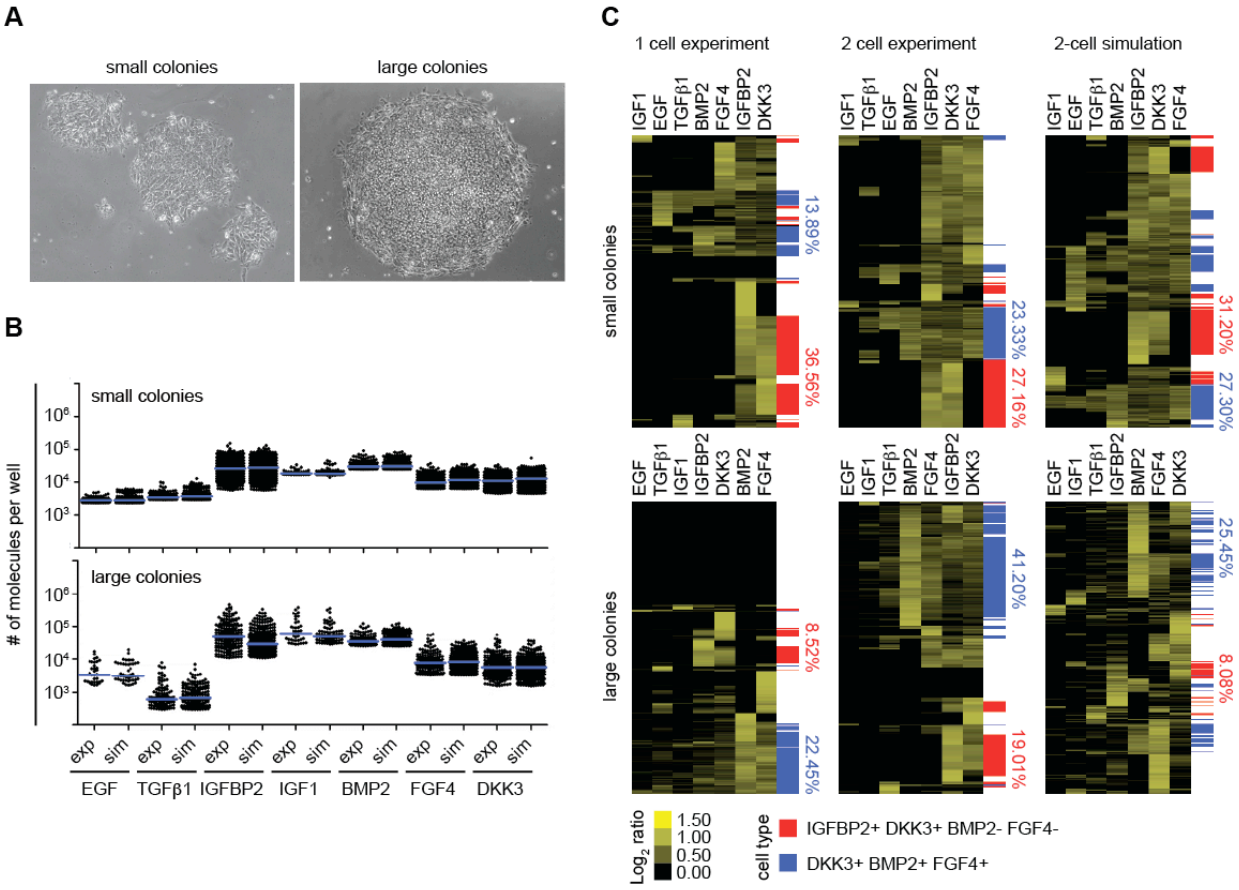

A
small colonies
large colonies
B
small colonies
large colonies
# of molecules per well
exp
sim
EGF
TGFβ1
IGFBP2
IGF1
BMP2
FGF4
DKK3
C
1 cell experiment
2 cell experiment
2-cell simulation
small colonies
large colonies
13.89%
36.56%
23.33%
27.16%
31.20%
27.30%
8.52%
22.45%
41.20%
19.01%
25.45%
8.08%
Log2 ratio
1.50
1.00
0.50
0.00
cell type
IGFBP2+ DKK3+ BMP2- FGF4-
DKK3+ BMP2+ FGF4+

Figure 5.4. DKK3 Affects BMP2 and FGF4 Secretion

A. Schematic depicting the experimental setup for the data shown in (B) and (C).

B. Effect of recombinant DKK3 (rDKK3) treatment on the secretion of IGFBP2, BMP2, and FGF4 over the course of 4 days, measurements taken every 2 days. Secretion values for conditioned media were first normalized to unconditioned treatment media, and the resulting values are then depicted as fold-change relative to the amount of protein secreted by the same culture for 2 days prior to starting rDKK3 treatment.

C. As in (B), except antibody recognizing DKK3 (αDKK3) was used as treatment instead over the course of 6 days.

D. Model summarizing the changing heterogeneity of hESC colonies as they grow larger in culture. Cells secreting high levels of IGFBP2 and IGF1R, represented by IGFBP2+/DKK3+ cells are apparent in small colonies, suggestive of an activated survival pathway. As colonies grow larger, the majority of IGFBP2 secretion localizes to the colony periphery, while BMP2 and FGF4 secreting cells are activated and localize to interspersed positions throughout the colony.

Figure 5.4

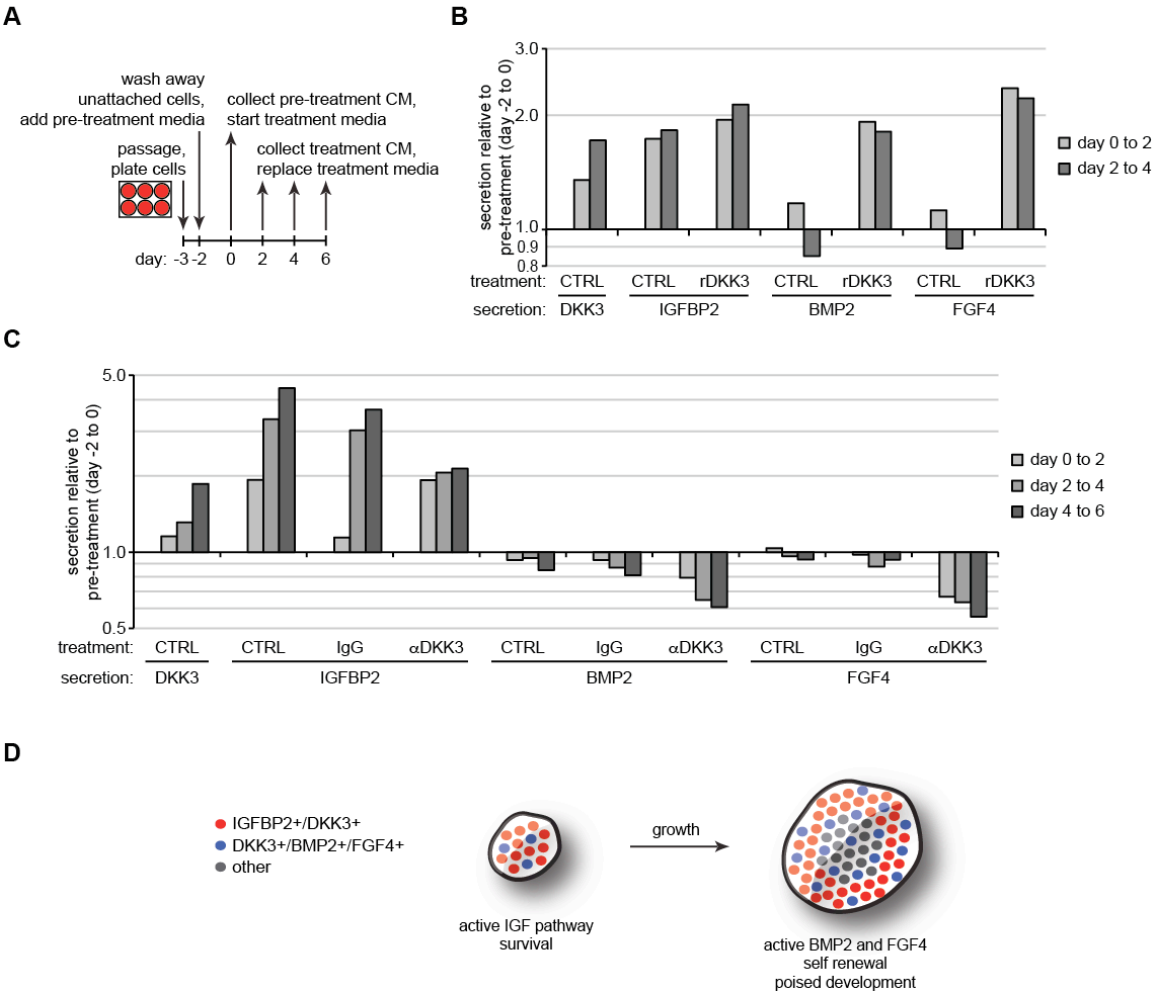

A
wash away
unattached cells,
add pre-treatment media
collect pre-treatment CM,
start treatment media
passage,
plate cells
collect treatment CM,
replace treatment media
day: -3 -2 0 2 4 6
B
secretion relative to
pre-treatment (day -2 to 0)
3.0
2.0
1.0
0.9
0.8
day 0 to 2
day 2 to 4
treatment: CTRL CTRL rDKK3 CTRL rDKK3 CTRL rDKK3
secretion: DKK3 IGFBP2 BMP2 FGF4
C
secretion relative to
pre-treatment (day -2 to 0)
5.0
1.0
0.5
day 0 to 2
day 2 to 4
day 4 to 6
treatment: CTRL CTRL IgG αDKK3 CTRL IgG αDKK3 CTRL IgG αDKK3
secretion: DKK3 IGFBP2 BMP2 FGF4
D
IGFBP2+/DKK3+
DKK3+/BMP2+/FGF4+
other
growth
active IGF pathway
survival
active BMP2 and FGF4
self renewal
poised development

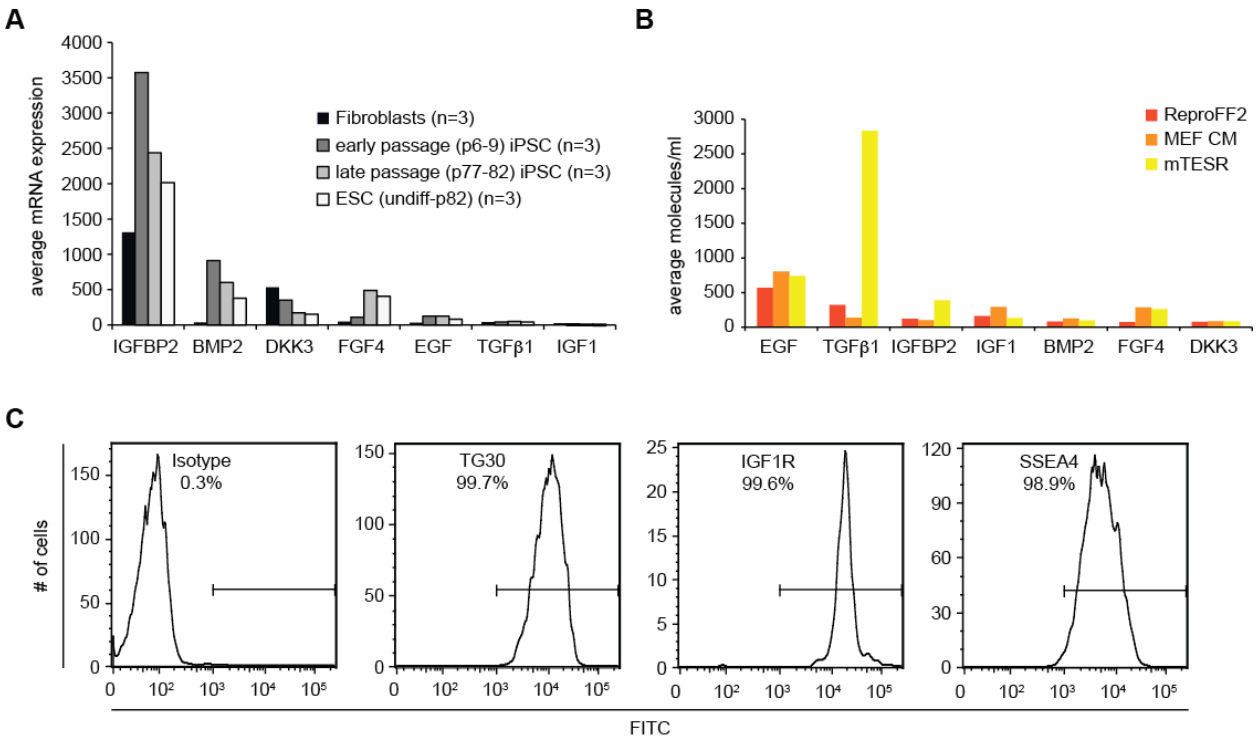
Figure S5.1. Selection of Reagents and Characterization of Culture Conditions Related to Figure 5.1

A. Microarray expression profile data of cytokines that display variable expression across somatic and pluripotent cells (Chin et al., 2009). Values represent average mRNA expression for each cytokine obtained from three independent cell lines representing each cell type. The range of passage numbers are indicated for iPSCs and ESCs.

B. Comparison among hESC-qualified culture media for detection of proteins being assayed.

C. Flow cytometry profiles for hESCs immunostained for markers associated with pluripotency. Immunoglobulin isotype is used as a control for primary antibody staining.

Figure S5.1


A
average mRNA expression
4000
3500
3000
2500
2000
1500
1000
500
0
IGFBP2
BMP2
DKK3
FGF4
EGF
TGFβ1
IGF1
Fibroblasts (n=3)
early passage (p6-9) iPSC (n=3)
late passage (p77-82) iPSC (n=3)
ESC (undiff-p82) (n=3)
B
average molecules/ml
3000
2500
2000
1500
1000
500
0
EGF
TGFβ1
IGFBP2
IGF1
BMP2
FGF4
DKK3
ReproFF2
MEF CM
mTESR
C
# of cells
Isotype
0.3%
TG30
99.7%
IGF1R
99.6%
SSEA4
98.9%
FITC

REFERENCES

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., Bevan, S.*, et al.* (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nature Biotechnology *25*, 803-816.

Bendall, S.C., Hughes, C., Campbell, J.L., Stewart, M.H., Pittock, P., Liu, S., Bonneil, E., Thibault, P., Bhatia, M., and Lajoie, G.A. (2009). An enhanced mass spectrometry approach reveals human embryonic stem cell growth factors in culture. Molecular & Cellular Proteomics : MCP *8*, 421-432.

Bendall, S.C., Stewart, M.H., Menendez, P., George, D., Vijayaragavan, K., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Rouleau, A., Yang, J., Bosse, M.*, et al.* (2007). IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. Nature *448*, 1015-1021.

Bershadsky, A.D., and Futerman, A.H. (1994). Disruption of the Golgi apparatus by brefeldin A blocks cell polarization and inhibits directed cell migration. Proceedings of the National Academy of Sciences of the United States of America *91*, 5686-5689.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J.*, et al.* (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell *5*, 111-123.

Clemmons, D.R., Busby, W.H., Arai, T., Nam, T.J., Clarke, J.B., Jones, J.I., and Ankrapp, D.K. (1995). Role of insulin-like growth factor binding proteins in the control of IGF actions. Progress in Growth Factor Research *6*, 357-366.

Davidson, G., Mao, B., del Barco Barrantes, I., and Niehrs, C. (2002). Kremen proteins interact with Dickkopf1 to regulate anteroposterior CNS patterning. Development *129*, 5587-5596.

Davidson, K.C., Adams, A.M., Goodson, J.M., McDonald, C.E., Potter, J.C., Berndt, J.D., Biechele, T.L., Taylor, R.J., and Moon, R.T. (2012). Wnt/beta-catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. Proceedings of the National Academy of Sciences of the United States of America *109*, 4485-4490.

Glinka, A., Wu, W., Delius, H., Monaghan, A.P., Blumenstock, C., and Niehrs, C. (1998). Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. Nature *391*, 357-362.

Johannesson, M., Stahlberg, A., Ameri, J., Sand, F.W., Norrman, K., and Semb, H. (2009). FGF4 and retinoic acid direct differentiation of hESCs into PDX1-expressing foregut endoderm in a time- and concentration-dependent manner. PloS One *4*, e4794.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proceedings of the National Academy of Sciences of the United States of America *105*, 2883-2888.

Mao, B., Wu, W., Davidson, G., Marhold, J., Li, M., Mechler, B.M., Delius, H., Hoppe, D., Stannek, P., Walter, C.*, et al.* (2002). Kremen proteins are Dickkopf receptors that regulate Wnt/beta-catenin signalling. Nature *417*, 664-667.

Mayshar, Y., Rom, E., Chumakov, I., Kronman, A., Yayon, A., and Benvenisty, N. (2008). Fibroblast growth factor 4 and its novel splice isoform have opposing effects on the maintenance of human embryonic stem cell self-renewal. Stem Cells *26*, 767-774.

Nakamura, R.E., and Hackam, A.S. (2010). Analysis of Dickkopf3 interactions with Wnt signaling receptors. Growth Factors *28*, 232-242.

Peerani, R., Rao, B.M., Bauwens, C., Yin, T., Wood, G.A., Nagy, A., Kumacheva, E., and Zandstra, P.W. (2007). Niche-mediated control of human embryonic stem cell self-renewal and differentiation. The EMBO Journal *26*, 4744-4755.

Pera, M.F., Andrade, J., Houssami, S., Reubinoff, B., Trounson, A., Stanley, E.G., Ward-van Oostwaard, D., and Mummery, C. (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. Journal of Cell Science *117*, 1269-1280.

Rosa, P., Barr, F.A., Stinchcombe, J.C., Binacchi, C., and Huttner, W.B. (1992). Brefeldin A inhibits the formation of constitutive secretory vesicles and immature secretory granules from the trans-Golgi network. European Journal of Cell Biology *59*, 265-274.

Singh, A.M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A.L., Sun, Y., Menendez, L., Kulik, M., and Dalton, S. (2012). Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. Cell Stem Cell *10*, 312-326.

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.

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.

Valera, E., Isaacs, M.J., Kawakami, Y., Izpisua Belmonte, J.C., and Choe, S. (2010). BMP-2/6 heterodimer is more effective than BMP-2 or BMP-6 homodimers as inductor of differentiation of human embryonic stem cells. PloS One *5*, e11167.

Wang, J., Tham, D., Wei, W., Shin, Y.S., Ma, C., Ahmad, H., Shi, Q., Yu, J., Levine, R.D., and Heath, J.R. (2012). Quantitating cell-cell interaction functions with applications to glioblastoma multiforme cancer cells. Nano Letters *12*, 6101-6106.

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.

CHAPTER 6

Conclusions and Outlook

This body of work represents studies involved in understanding the signaling and transcriptional mechanisms regulating the maintenance of pluripotent cells and reprogramming to the pluripotent state. In Chapter 2, we establish a framework for understanding the challenges involved in reprogramming a somatic cell to the pluripotent state by overexpression of a defined set of transcription factors. These vary with cell type of origin, include aspects of cell cycle, chromatin state, differentiation status of the starting cell, and reprogramming methods used, which encompass the overexpression or downregulation of biological molecules such as transcription factors, chromatin modifiers, or microRNAs, as well as culture conditions. We learn that the regulation of the reprogramming process is complicated not only by the number of barriers involved, but the timing of their effect. Chapter 2 also addresses the qualitative analysis of iPSCs and calls for considerations in standardizing methods to assess reprogramming data across laboratories, which will be beneficial for the clinical usage of iPSCs as the technology becomes more refined.

In Chapter 3, we examined Wnt signaling, and its transcriptional effector proteins, the T-cell factors (Tcfs), for their role in regulating the reprogramming process. Given that the Wnt signaling pathway is ubiquitously involved in embryonic development, in the maintenance of pluripotency in ESCs, as well as in fusion mediated reprogramming, it was surprising that the ectopic activation of Wnt signaling had minimal effects on transcription factor-mediated reprogramming to iPSCs. By dissecting the reprogramming process into early and late stages, we understand that the minimal effects of continuous Wnt pathway activation during iPSC reprogramming reported previously (Marson et al., 2008; Takahashi and Yamanaka, 2006) are due to the differential effects of the Wnt pathway in the early and late stages of reprogramming. Wnt pathway stimulation has an inhibitory effect during the early stage of reprogramming and an enhancing effect during the late stage. A systematic analysis of the involvement of all four Tcfs, the family of transcriptional effectors within the canonical Wnt pathway, in this stage-specific

response to Wnt activation provided a consistent model of the early and late reprogramming processes with regards to the Wnt pathway.

We subsequently centered our focus on Tcf3, primarily a transcriptional repressor of Wnt target genes that is well characterized for its function in the regulation of the pluripotency transcriptional network in ESCs. Modulation of Tcf3 levels throughout the reprogramming process confirmed the biphasic nature of the reprogramming process. During the early phase, the presence of Tcf3 promotes the activation of cell cycling genes that resemble an ESC expression pattern, while during the late phase, the absence of Tcf3 prevents the aberrant upregulation of developmental genes that would stymie the establishment of the pluripotent state. Importantly, this study illustrates that the timing and levels of Tcf3 can effectively modulate the requirement for one of the reprogramming factors, Sox2. This study highlights that signaling and transcriptional networks are constantly changing throughout the reprogramming process, and demonstrates that one pathway, such as that of Wnt signaling, can have multiple and even opposite effects on reprogramming if activated at a particular phase. This temporal aspect adds another dimension of complexity to the reprogramming process. Therefore the potentially changing effects of other signaling and transcriptional networks besides the Wnt signaling pathway should be carefully considered in attempts to improve reprogramming efficiency.

In Chapter 4, we investigated the role of Polycomb proteins in the reprogramming process based on prior characterization of the changing H3K27me3 profiles between somatic and pluripotent cells (Maherali et al., 2007; Mikkelsen et al., 2008). We find that despite its important role in development (Voncken et al., 2003), Ring1b is not required for the reprogramming process. This highlights potential differences between differentiation and reprogramming in terms of PRC1 requirement. Much work remains to be done to determine the

exact role PRC1 may have in the reprogramming process, in comparison to its role during development. In contrast to Ring1b, Ezh2 function is required for reprogramming, particularly at the early stages, which, in the absence of Ezh2, appears to be stalled in part by upregulated cell cycle inhibitors Arf and Ink4a. However, Ezh2 becomes dispensable in the later reprogramming stages, and subsequently dispensable in the pluripotent state, despite the requirement for the essential PRC2 component Eed at all reprogramming stages. The derivation of iPSCs lacking catalytic Ezh2 enabled us to observe a global reduction of H3K27me3 in these cells, but focused regions of H3K27me3 were maintained at distinct genomic loci, likely due to the function of another PRC2 complex containing Ezh1. These methylation events include early embryonic development genes typically kept silent in ESCs, suggesting that the expression of these genes is not compatible with the pluripotent state. Future expression analysis from Ezh2 mutant iPSC lines will enable us to distinguish those targets that depend on Ezh2-mediated H3K27 methylation in actively reprogramming cells from those targets that depend on Ezh2 methylation in established pluripotent cells.

In Chapter 5, we discovered that hESC colonies are heterogeneous, with single cells secreting distinct profiles of cytokines throughout colony growth. Combining microfluidic, immunofluorescence imaging, computational modeling, and cell biology methods, we uncover a unique relationship among the cytokines IGFBP2, DKK3, BMP2, and FGF4, and also a relationship between the cells that secrete them. We used human, as opposed to mouse, ESCs and iPSCs to investigate signaling at the single cell level based on two characteristics: 1) human pluripotent stem cell colonies grow in two-dimensional structures, thereby simplifying the analysis of effects due to a cell's localization and orientation within a colony, and 2) the fact that ELISA antibody sets are more commercially available for human proteins due to their clinical use in cancer diagnostics. Future experiments with our single cell barcode chip will incorporate a wider panel of antibodies and may be applicable for characterizing iPSCs to be used clinically.

It will be interesting to gain more mechanistic insights into the causes and consequences of the heterogeneity that we discovered in this study.

The overall findings described in this body of work highlight novel considerations for future investigations into mechanisms affecting the reprogramming and applications of iPSCs. One consideration, is that the reprogramming process is a multiphasic process, where a transcription factor such as Tcf3 can regulate a particular set of genes that affect the early stages of reprogramming, but regulate another set of genes that affect the late stages of reprogramming.

This theme of distinct, multiphasic reprogramming effects also extends to PRC2 functions. Inactivation of Ezh2-containing PRC2 during early reprogramming inhibits the process in part by derepressing cell cycle inhibitors. While this Ezh2-containing PRC2 is dispensable during the late reprogramming phase, Eed, and thus PRC2-mediated H3K27 methylation, is not. In line with this idea of stage-dependent contexts, other studies have also characterized the multiphasic nature of reprogramming. These commonly report a widespread initiation of gene expression changes that are characterized by a mesenchymal to epithelial transition (MET), a stabilization phase in which a subset of pluripotency genes begins to be expressed, and a maintenance phase where the complete pluripotency transcriptional network is established after silencing of the reprogramming factors (Golipour et al., 2012; Polo et al., 2012; Samavarchi-Tehrani et al., 2010). Since iPSC reprogramming undergoes MET, it was generally assumed that applying conditions that stimulate MET throughout the duration of the reprogramming process would increase efficiency (Li et al., 2010). However, a more recent finding demonstrated that by inducing an epithelial to mesenchymal transition (EMT) prior to MET, reprogramming efficiency was surprisingly enhanced (Liu et al., 2013). These results resemble the opposite effects of Wnt signaling we observed, where conditions that promote one segment

of the reprogramming process may in fact be detrimental to another segment. Future reprogramming studies operating under this paradigm will likely result in improved reprogramming techniques.

Another consideration for future studies in reprogramming is the heterogeneous nature of reprogramming cultures. While molecular analysis across whole cell cultures has revealed much insight into the gene expression changes and epigenetic states during the reprogramming process (Golipour et al., 2012; Polo et al., 2012; Samavarchi-Tehrani et al., 2010), sorting out the deterministic and stochastic events that are intrinsic to successful reprogramming will require single cell analysis. Despite the situation that single cell analysis on the genome-wide level is currently in its infancy, several studies have attempted to characterize reprogramming cells to identify definitive events that mark a faithful path to pluripotency (Buganim et al., 2012; Polo et al., 2012). Heterogeneous populations arise stochastically and persist even into late reprogramming intermediates, which are marked by SSEA1-positivity (Polo et al., 2012). However, much work remains to identify true reprogramming events within these populations. Our single cell studies take an alternative approach to understand how heterogeneity is intrinsic to pluripotent cells in culture.

In terms of the panel of cytokines that we assayed, the human fibroblasts and H9 hESC-derived fibroblasts that we secretion profiled in Chapter 5 bear a relatively homogeneous population of secretory cells. Understanding how fibroblasts undergoing reprogramming recapitulate the heterogeneous composition of pluripotent stem cell colonies will likely improve not only the efficiency of the reprogramming process, but also methods to directly differentiate pluripotent cells into pure populations of somatic cells for usage in disease modeling and regenerative medicine.

REFERENCES

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.

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.

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.

Liu, X., Sun, H., Qi, J., Wang, L., He, S., Liu, J., Feng, C., Chen, C., Li, W., Guo, Y.*, et al.* (2013). Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT-MET mechanism for optimal reprogramming. Nature Cell Biology *15*, 829-838.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R.*, et al.* (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55-70.

Marson, A., Foreman, R., Chevalier, B., Bilodeau, S., Kahn, M., Young, R.A., and Jaenisch, R. (2008). Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell *3*, 132-135.

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.

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.

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.

Voncken, J.W., Roelen, B.A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proceedings of the National Academy of Sciences of the United States of America *100*, 2468-2473.