

Global Indiscriminate Methylation in Cell-Specific Gene Promoters following Reprogramming into Human Induced Pluripotent Stem Cells

Jonathan Nissenbaum,^{1,2,5} Ori Bar-Nur,^{1,2,3,4,5} Eyal Ben-David,² and Nissim Benvenisty^{1,2,*}

¹Stem Cell Unit

²Department of Genetics

Institute of Life Sciences, The Hebrew University of Jerusalem, Israel

³Massachusetts General Hospital Cancer Center and Center for Regenerative Medicine, 185 Cambridge Street, Boston, MA 02114, USA

⁴Present address: Harvard Stem Cell Institute, 1350 Massachusetts Avenue, Cambridge, MA 02138, USA

⁵These authors contributed equally to this work

*Correspondence: nissimb@cc.huji.ac.il

<http://dx.doi.org/10.1016/j.stemcr.2013.11.007>

This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-No Derivative Works License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

SUMMARY

Molecular reprogramming of somatic cells into human induced pluripotent stem cells (iPSCs) is accompanied by extensive changes in gene expression patterns and epigenetic marks. To better understand the link between gene expression and DNA methylation, we have profiled human somatic cells from different embryonic cell types (endoderm, mesoderm, and parthenogenetic germ cells) and the iPSCs generated from them. We show that reprogramming is accompanied by extensive DNA methylation in CpG-poor promoters, sparing CpG-rich promoters. Intriguingly, methylation in CpG-poor promoters occurred not only in downregulated genes, but also in genes that are not expressed in the parental somatic cells or their respective iPSCs. These genes are predominantly tissue-specific genes of other cell types from different lineages. Our results suggest a role of DNA methylation in the silencing of the somatic cell identity by global nonspecific methylation of tissue-specific genes from all lineages, regardless of their expression in the parental somatic cells.

INTRODUCTION

Forced expression of transcription factors in human somatic cells allows the generation of induced pluripotent stem cells (iPSCs) (Takahashi et al., 2007). These cells are equivalent to inner-cell mass-derived embryonic stem cells (ESCs) and hold a great promise for regenerative medicine and cell replacement therapy. The way somatic cells transition to the pluripotent state is not yet fully elucidated. The activation of the pluripotent state depends on the ability to upregulate a set of pluripotency genes. Unraveling the way by which pluripotent factors interact with the genome is key to understanding cellular reprogramming (Plath and Lowry, 2011). Although the thorough studies concerning the action of the pluripotent factors illuminate some aspect of the silencing of the somatic cell identity, the induction of pluripotency gene targets by itself is insufficient to solely explain the conversion of somatic cells into pluripotent cells, and other cellular processes need to occur for the erasure of the somatic cell identity. Methylation of cytosine in the context of CpG dinucleotides in gene promoters has been acknowledged for many years as a mechanism for regulation of gene expression in mammalian cells (Cedar and Bergman, 2009). Differential gene expression between somatic cells and ESC cells has been shown to be governed by methylation of gene promoters (Meissner et al., 2008). The genomic landscape affects the location and level of

DNA methylation by the content of the CpG dinucleotides in a given genomic region. DNA methylation density varies in a CpG rich versus CpG poor regions (Hawkins et al., 2010; Lister et al., 2011). Overall, gene promoters are generally characterized by a high content of CpG dinucleotide (HCpG) known as well as CpG Islands, or by a low content of CpG dinucleotide (LCpG). Given the complex interplay between DNA methylation and gene expression, comprehensive correlation analysis can illuminate our understanding of the reprogramming process. Recent studies that have focused on DNA methylation profiling of different CpG regions during reprogramming, included limited expression analysis, mainly in the form of preselected genes sets with an a priori knowledge regarding their mode of action (Nishino et al., 2011; Weber et al., 2007). Other studies have focused on CpG regions from an opposite direction, i.e., the methylation processes that occur when pluripotent cells differentiate in culture (Brunner et al., 2009; Xie et al., 2013). Here, we set out to investigate the methylation and expression dynamics of somatic cells representative of three different embryonic cell types (mesoderm, endoderm, and teratoma cells derived from parthenogenetic germ cells) and their respective iPSCs. We thus aimed at deciphering the involvement of DNA methylation in silencing the somatic cell identity in the context of different somatic cells with distinct genetic and epigenetic backgrounds.

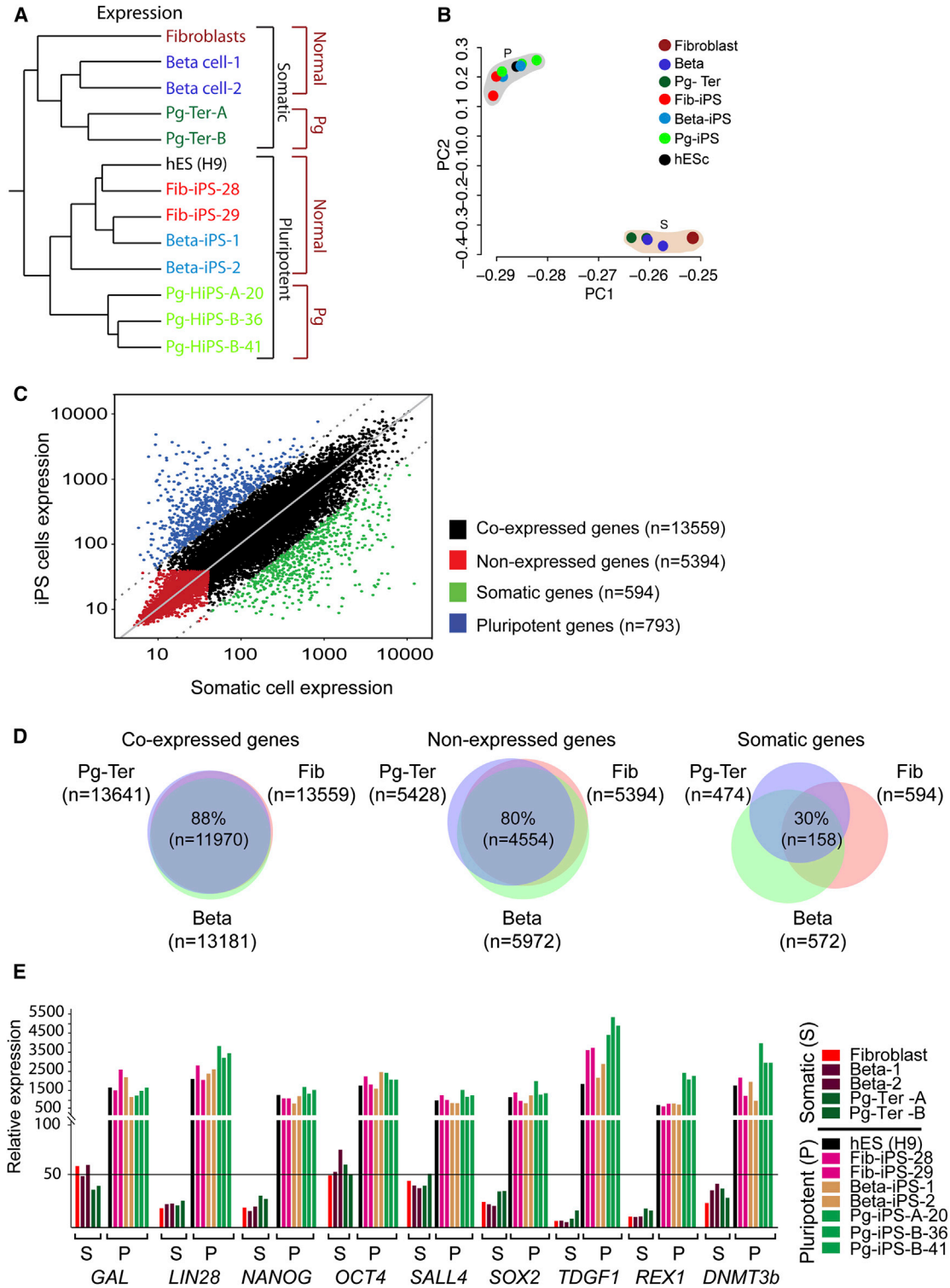


Figure 1. Expression Analysis of Somatic and Pluripotent Cell Lines

(A) Hierarchical clustering analysis based on mRNA expression. Somatic cells and their derived iPSCs are designated in dark/light matching colors.

(B) Principal-component analysis (PCA) based on RNA expression levels across somatic and pluripotent cell samples. Coloring indicates classification of samples into subgroups. Light brown and gray shades represent somatic and pluripotent cells, respectively.

(legend continued on next page)



RESULTS

To study the status of DNA methylation during cellular reprogramming, we have analyzed the gene expression and methylation profiles of somatic cells from three different lineages, representative of different embryonic germ layers, and the iPSCs derived from them, as well as control human ESCs. For mesoderm, we have chosen human fibroblasts and the iPSCs (Fib-iPSCs) generated from them (Pick et al., 2009; Urbach et al., 2010). For endoderm, we have used human pancreatic beta cells and beta-iPSCs (Bar-Nur et al., 2011), and for the germline we have used human parthenogenetic ovarian teratoma-derived cells and parthenogenetic iPSCs (Pg-iPSCs) generated from them (Stelzer et al., 2011). For each lineage, we have used between two and three iPSC clones in all analyses. We initially compared the somatic and pluripotent cells by gene expression microarrays. As expected, an unsupervised hierarchical clustering separated the somatic and pluripotent cells into two distinct groups (Figure 1A). Within the somatic group, further separation was observed based on the origin of the somatic cells; however, for the pluripotent cells, this distinction was only seen in the Pg-iPSCs versus other iPSCs, probably due to the lack of expression of paternal imprinted genes in the parthenogenetic cells (Figure 1A). Further separation of somatic versus pluripotent cells was observed following principal-component analysis (PCA; Figure 1B) that placed all pluripotent cells in one group while all somatic cells were clustered in a separate group. As expected, pluripotent cells, although generated from various origins, were clustered tightly together (Figure 1B). We then correlated each iPSC clone to the somatic cell from which it was derived (Figure 1C). This has allowed us to generate four different groups of genes: (1) genes that are upregulated following reprogramming (“pluripotency genes”), (2) genes that are downregulated following reprogramming (“somatic genes”), (3) genes that are not expressed in both cell types (“nonexpressed genes”), and (4) genes that are expressed in both cell types (“coexpressed genes”) (Figure 1C). This analysis shows that only a relatively small subset of genes are differentially expressed between iPSCs and the somatic cells from which they are derived (Figure 1C). Next, we looked at the overlap among the three different somatic lineages within each of the four

gene expression categories. Coexpressed and nonexpressed genes show a high overlap (88% and 80%, respectively) among the three groups (Figure 1D). Somatic-specific genes show low overlap among the three groups (30%), which is expected since each cell type represents a different embryonic lineage. Lastly, the expression levels of well-known pluripotency related genes were significantly higher in all the pluripotent cells compared to the somatic cells (Figure 1E). We then profiled the methylation status of the cells. We have used Illumina's *Infinium* Methylation 27 BeadChip Platform arrays that sample 27,578 CpG sites in promoter regions covering about 15,000 genes in the genome. Here, again, unsupervised hierarchical clustering separated the somatic and pluripotent cells into two distinct groups, with continued subdivision into smaller groups based on the cell type (Figure 2A). Similar to the above expression analysis, the PCA method was applied and a clear distinction between pluripotent and somatic cells was observed (Figure 2B). We next looked at the global methylation levels in each cell type. The Illumina array samples CpG across the genome and gives a β value for methylation in each CpG, ranging between 0 and 1 (where 0 is not methylated and 1 is fully methylated). CpG that is fully methylated (i.e., score of “1”) is representative to the sampled region. We divided the β values into three categories: (1) hypomethylated (β between 0 and 0.3), (2) hemimethylated (β between 0.3 and 0.6), and (3) hypermethylated (β between 0.6 and 1). The Illumina platform allows dividing the methylation of the genes into two distinct groups: CpG island promoters (characterized by HCpG content) and nonisland CpG promoters (characterized by LCpG content). The methylation patterns in the somatic cells show that they are more methylated in the LCpG and that CpG islands are overall not methylated (Figure 2C). All the iPSCs showed higher β values than their corresponding somatic cells and were similar to the levels observed in ESCs (Figure 2C), thus corroborating previous observations (Hawkins et al., 2010; Lister et al., 2011). The higher methylation levels observed in iPSCs were predominantly located in LCpG, while the HCpG promoters showed little change in methylation levels following reprogramming (p value < 0.0001 and p value = 0.61, respectively; Figure 2C). Our analysis shows that most of the genes having HCpG content show low levels of

(C) Scatterplot of somatic and iPSC gene expression. x axis = fibroblast cell expression; y axis = Fib-iPSC expression. Transcripts were divided into four gene expression categories: (1) coexpressed genes, (2) nonexpressed genes, (3) somatic-specific genes, and (4) pluripotency-specific genes (black, red, green, and blue, respectively).

(D) Gene overlap for the coexpressed genes, nonexpressed genes, and somatic genes are presented in the Venn diagrams. The beta cell (Beta), fibroblast (Fib), and parthenogenetic teratoma (Pg-Ter) lineages are represented by light green, light red, and purple, respectively.

(E) Gene expression of typical core pluripotency genes among the somatic and pluripotent cells. x axis = pluripotency genes for the somatic (s) and pluripotent (p) cells; y axis = gene expression level. Parental origin cell line and their respective iPSCs are designated by a different color (dark/light hues). Horizontal black line represents an expression threshold.

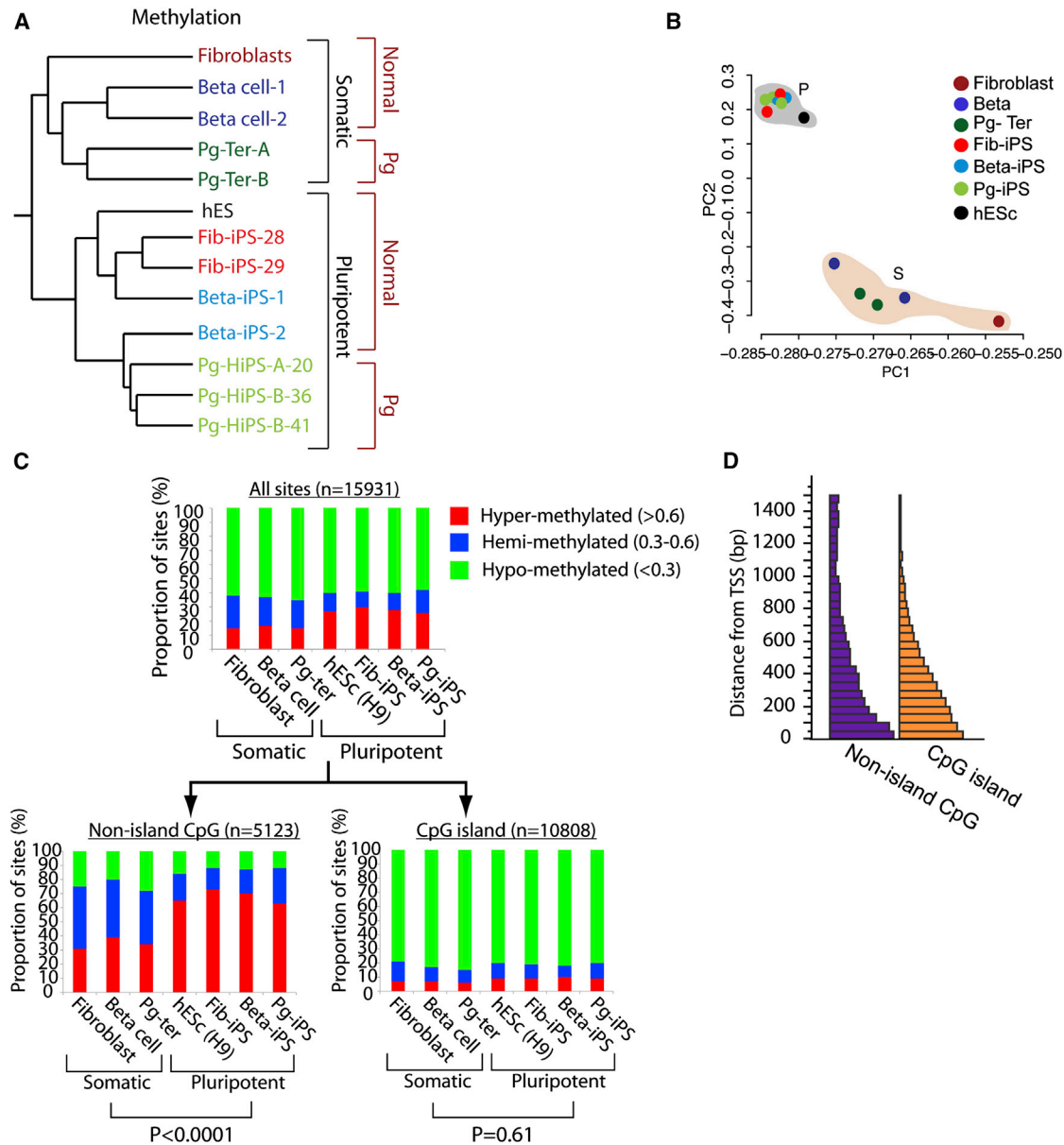


Figure 2. DNA Methylation Analysis of Somatic and Pluripotent Cell Lines

(A) Hierarchical clustering analysis based on DNA methylation. Somatic cells and their derived iPSCs are designated in dark/light matching colors.

(B) PCA based on CpG methylation levels across somatic and pluripotent cell samples. Coloring indicates classification of samples into subgroups. Light brown and gray shades represent somatic and pluripotent cells, respectively.

(C) Distribution of methylation scores across somatic and pluripotent cells. The three histograms represent the distribution of hyper-methylated (>0.6), hemimethylated (0.3–0.6), and hypomethylated (<0.3) genes (red, blue, and green, respectively) for all island CpG and nonisland CpG sites. p values = Fisher's exact test.

(D) CpG site distance from transcription start site (TSS). The histogram shows the distribution of distance from the TSS for the CpG island and nonisland CpG sites (orange and purple, respectively). y axis = distance from TSS in base pairs.

methylation, while only a small subset maintain a high degree of methylation. Overall, the methylation levels in these gene promoters does not dramatically change following reprogramming and they remain either hypo-

or hypermethylated (Figure 2C). Our analysis also shows that methylation is much more prominent following reprogramming than demethylation (Figure 2C). The distribution of the distance of the CpG dinucleotides from the



transcription start site (TSS) shows that most LCpG and HCpG sites are located in 1,000 bp range from the TSS (89% and 97%, respectively; Figure 2D), in accordance with other studies that compared tissue-specific genes and highly expressed housekeeping genes (characterized by nonisland CpG and CpG islands, respectively) (Brenet et al., 2011; Morita et al., 2012). In summary, our data show that reprogramming is accompanied by massive methylation of gene promoters that have a low number of CpG dinucleotides, resembling the state in ESCs. This de novo methylation occurred in all the iPSC clones, regardless of the starting cell type or lineage.

We next sought to look at the correlation between the methylation and expression levels of each somatic cell and the iPSCs that were generated from it. We looked at the methylation status for CpG island or nonisland CpG promoters in the four expression groups that we generated previously (Figure 1C). Although the overall methylation proportions of HCpG genes provide a rather static picture (Figure 2C), a certain fraction of the genes change their methylation status following reprogramming (Figure 3A). Plotting methylation versus expression revealed that many upregulated pluripotency genes underwent promoter demethylation following reprogramming, especially those that are located in HCpG promoters, as the vast majority of pluripotent genes are characterized by CpG island promoters (Figure 3A). Concomitantly, the somatic-specific gene group that is downregulated in iPSCs underwent extensive de novo methylation in LCpG rather than in HCpG promoters (Figures 3A and 3B; Figures S1 and S2 available online). Intriguingly, genes that are not expressed in either somatic cells or iPSCs generated from them showed a great increase in methylation in LCpG promoters, regardless of the parental somatic cell type (Figure 3B; Figure S2). The nonexpressed gene group was significantly different from the coexpressed genes in the LCpG category and also greatly differs from the nonexpressed gene group in the HCpG category (Figures 3B; Figure S2). This group of genes behaves in a similar fashion to the somatic-specific gene group, which shows a similar de novo methylation trend. This raised an interesting issue: why would genes that are not expressed by the somatic cells undergo extensive de novo methylation? To address this issue, we used the Amazonia database (<http://amazonia.transcriptome.eu/>; Le Carrour et al., 2010), which shows the expression levels of most human genes in hundreds of human tissue samples representing the cell repertoire of the human body. Analysis of the genes that are not expressed in any of the three somatic cell types or their respective iPSCs by the Amazonia database demonstrated that they are tissue-specific genes, each representing a different cell type (Figure 3C; Table S1) such as skin, intestine, blood, bone marrow, testis, or brain. The differences in methyl-

ation between the somatic and the pluripotent cells were highly significant for all cell-specific genes. A gene ontology annotation analysis of these genes showed enrichment for processes such as epidermal, keratinocyte, and epithelial cell differentiation and developmental processes.

DISCUSSION

DNA methylation is an epigenetic modification that plays a pivotal role in the silencing of gene expression, such as the silencing of imprinted genes or genes that reside in the inactive X chromosome (Cedar and Bergman, 2009; Weber et al., 2007). How the role of DNA methylation in resetting the epigenome of somatic cells following reprogramming to pluripotency is not yet fully understood. Since reprogramming to pluripotency dictates the activation and repression of a subset of genes, it is expected that methylation will play major role in this process. Repression of lineage-specific genes that are not expressed in a given somatic cell can be maintained due to the fact that specific transcription factors are absent. This regulation can then be augmented by DNA methylation executed by de novo methylases. Our results, using three different somatic cells and their iPSC progenies, corroborates previous observations that reprogramming is accompanied by a wave of de novo methylation in iPSCs (Deng et al., 2009; Doi et al., 2009; Lister et al., 2011; Nagae et al., 2011; Nishino et al., 2011). Although the de novo methylase *DNMT3b* is expressed only in the pluripotent cells (Figure 1E), the trajectory in which hypermethylation is acquired can occur alongside gene repression, to precede the regulatory repression, or to occur after the gene is silenced. In addition, our work refines these observations by analyzing the expression levels of the cells, showing that this methylation occurs not only for the somatic-specific expressed genes, but also for many other nonexpressed genes.

Our results can be explained by the following model. During reprogramming, the de novo DNA methyltransferases (*DNMTs*) methylates not only the promoters of the somatic-specific genes that are undergoing reprogramming, but also the promoters of most cell-specific genes, even if they are not expressed in the somatic cell undergoing reprogramming (Figure 4). This suggests that the *DNMTs* cannot separate between the different tissue-specific promoters of the cell undergoing reprogramming and that gene silencing occurs in an indiscriminate mode that does not distinguish between parental cells from different lineages. For example, the gene *CMTMC5* is expressed predominantly in ectodermal cells of the adult brain; however, this gene undergoes de novo methylation in iPSCs that are generated from mesoderm, endoderm,

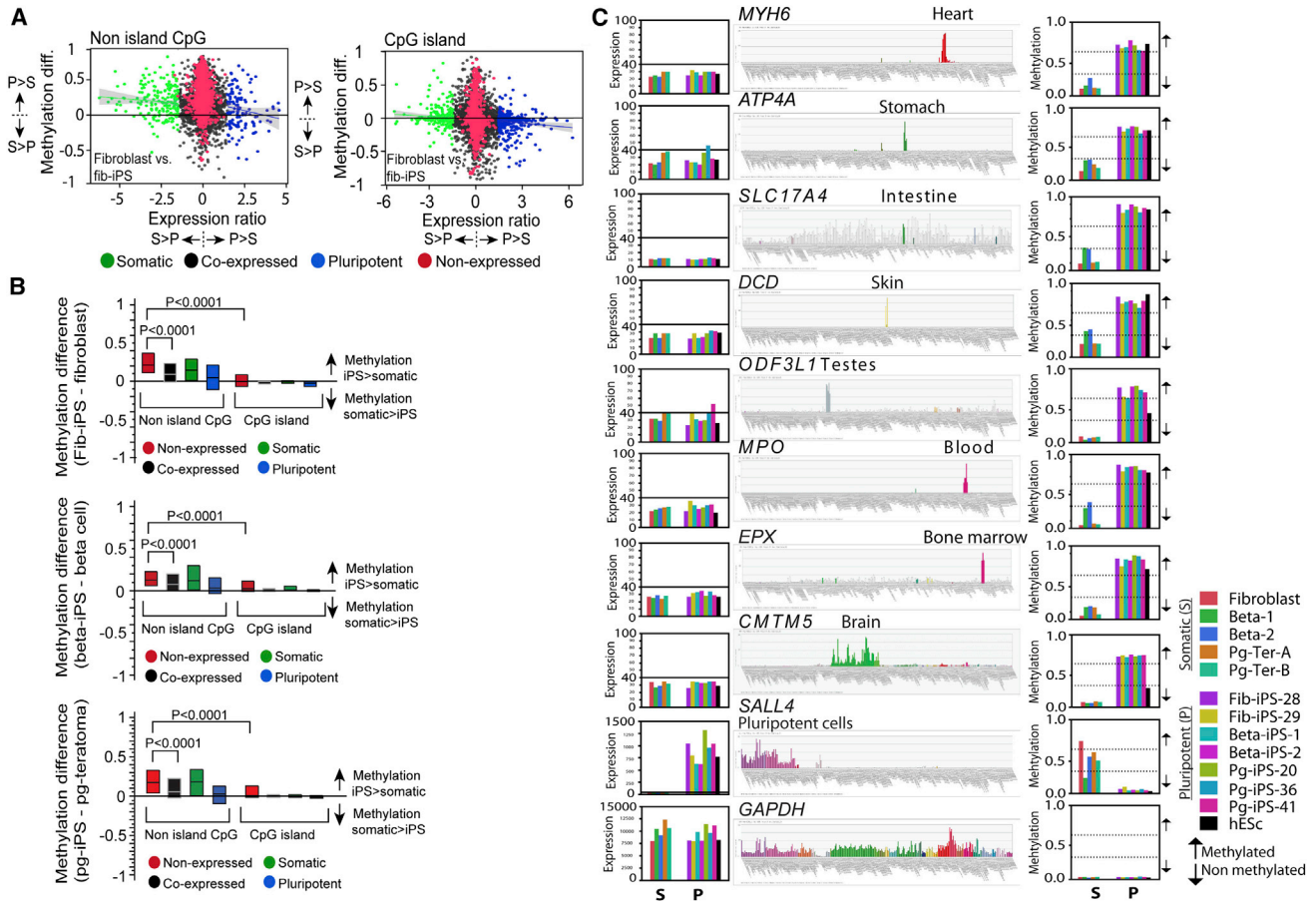


Figure 3. Analysis of Combined Expression and Methylation Data

(A) Expression versus methylation. For fibroblast versus fib-iPSCs, island CpG and nonisland CpG are presented separately. x axis = expression ratio (log iPSC cell expression – log somatic cell expression); y axis = methylation difference (iPSC methylation – somatic cell methylation). A positive value on the y and x axes reflects higher methylation or expression in iPSCs than in the parental somatic cells and vice versa. Colors represent the four expression categories as in Figure 1B. Regression lines are presented for the somatic and pluripotent gene categories (green and blue, respectively), and shaded gray boundaries represent confidence interval (CI). See also Figure S1.

(B) Comparison of DNA methylation of fibroblast and their respective iPSC cell line. y axis = methylation difference based on fib-iPSCs values minus fibroblast values. A positive value reflects higher methylation in iPSCs than in the parental somatic cells and vice versa. The box plots illustrate the methylation difference for each expression category (Figure 1C) and for the CpG and nonisland CpG groups. Each panel represents different cell origins (fibroblasts, beta cells, and parthenogenetic teratomas). p value = Student's t test. See also Figure S2.

(C) Tissue-specific gene expression profile using the Amazonia database. The displayed genes, which are not expressed in the somatic (S) and pluripotent (P) cells, become hypermethylated during reprogramming. mRNA expression and methylation values are presented in the histograms. Left side = expression; right side = methylation. Center histogram: expression levels in various tissues and cells. *GAPDH* and *SALL4* serve as control for housekeeping and pluripotency genes, respectively. Black line = expression threshold; dotted black line = threshold for hypermethylation and hypomethylation. See also Table S1.

and teratoma cells derived from parthenogenetic germ cells (Figure 3C). A different interpretation is that iPSCs undergo extensive aberrant methylation in vitro (Lister et al., 2011). However, the high concordance between methylation of the tissue-specific genes among our iPSCs that are generated from distinct lineages suggests otherwise. If indeed the methylation of tissue-specific genes represents an aberrant phenomenon, then we would not expect it to be

consistent across three diverse cell types and many independent and different reprogramming experiments. In addition, a recent study shows that keratinocytes reprogram much faster than fibroblasts because they are more methylated than fibroblasts (Barrero et al., 2012). In this case, the *DNMTs* may not need to methylate all the tissue-specific gene promoters, thus enhancing reprogramming efficiency.

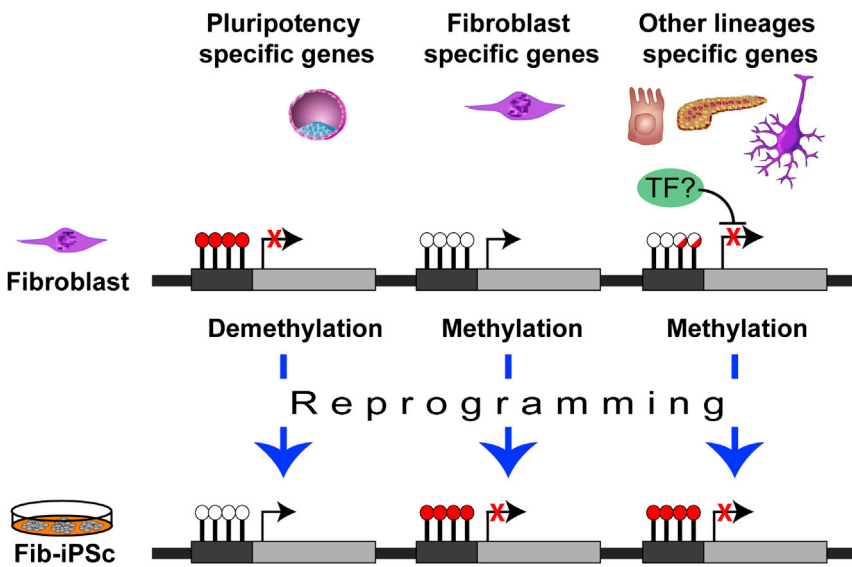


Figure 4. A Model that Illustrates the Possible Silencing of the Somatic Cell Identity following Reprogramming by DNA Methylation

Following reprogramming, pluripotency-related gene promoters undergo extensive DNA demethylation regardless of the starting somatic cell type. Somatic cell-specific gene promoters undergo extensive methylation (illustrated by a fibroblast cells). Other cell-specific gene promoters also undergo extensive DNA methylation, as the methylating enzymes cannot distinguish between various somatic cell-specific promoters, thus acting in a “blind fashion.”

In mouse ESCs, it was recently shown that the cells are hypermethylated when grown with serum; however, a switch to serum-free medium supplemented with mitogen-activated protein/extracellular signal-regulated kinase kinase (MEK) inhibitor or GSK3 inhibitor (2i medium) results in genome-wide hypomethylated mouse ESCs that resemble an earlier developmental “naïve state” (Leitch et al., 2013). Human ESCs may represent a later developmental stage than mouse ESCs, more similar to the mouse epiblast stem cell stage (Tesar et al., 2007). Human iPSCs resemble human ESCs, and thus the high degree of methylation we observe in both cell types may represent a depiction of their in vivo postimplantation state, just before the cells start to acquire their somatic cell identity. It will be interesting to examine the methylation status of naïve human PSCs and see if they are relatively hypomethylated (Hanna et al., 2010). Which enzymes mediate the extensive de novo methylation during reprogramming is yet to be determined. A recent work shows that non-CpG methylation in PSCs is mediated by *DNMT3a* and *DNMT3b*, as knocking down these genes eliminated most of the non-CpG methylation (Ziller et al., 2011). It is likely that these enzymes are also responsible for the de novo methylation we detect in iPSCs in LCpG promoters as the expression of *DNMT3b* was upregulated in all our pluripotent cells (Figure 1E). How HCpG gene promoters are protected from the de novo methylation is yet not fully understood. Binding of proteins to CpG islands may interfere with the *DNMTs* attempt to methylate these HCpG promoters, but more work is needed to show if this is indeed the case. Finally, several recent works suggest that methylation plays a key role in mediating the ability of cells to differentiate (Bar-Nur

et al., 2011; Kim et al., 2010; Lister et al., 2011; Nagae et al., 2011; Polo et al., 2010), sometime due to an epigenetic memory (Bar-Nur et al., 2011; Kim et al., 2010; Lister et al., 2011; Polo et al., 2010). We propose that a thorough dissection of the methylation status of each iPSCs will greatly benefit the use of iPSCs for direct differentiation protocols and in using the cells for potential cell replacement therapy.

EXPERIMENTAL PROCEDURES

Generation and Culture of iPSCs

Generation of fibroblast iPSCs, beta iPSCs, and parthenogenic iPSCs were reported previously (Bar-Nur et al., 2011; Pick et al., 2009; Stelzer et al., 2011). See the Supplemental Experimental Procedures for full culture conditions.

DNA and mRNA Extraction

Total genomic DNA was extracted using a genomic DNA extraction kit (RBC). Total RNA (DNase-treated) was extracted using the RNAeasy Mini Kit (QIAGEN).

DNA Expression and Methylation Microarray Analyses

RNA and DNA were subjected to either Human Gene 1.0 ST microarray platform (Affymetrix) or to HumanMethylation27 BeadChip (Illumina) analysis, respectively. See the Supplemental Experimental Procedures for further details.

ACCESSION NUMBERS

Microarray expression data are available in the Gene Expression Omnibus (GEO) database under the accession numbers GSE27362 and GSE29880. Illumina methylation data have been deposited in GEO under the accession number GSE52238.



SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, two figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.stemcr.2013.11.007>.

ACKNOWLEDGMENTS

The authors thank Uri Ben-david, Yonatan Stelzer, Daniel Ronen, Tomer Halevy, and Uri Weissbein for critically reading this manuscript and Tamar Golan-Lev for her assistance with the graphical design of the figures. We thank Yonatan Stelzer for his assistance with this work. N.B. is the Herbert Cohn Chair in Cancer Research. This research was partially supported by the Legacy Heritage Biomedical Science Partnership Program of the Israel Science Foundation (grant 269/12) and the Juvenile Diabetes Research Foundation (grant 1-2011-555). The research leading to these results has also received partial funding from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115439 (StemBANCC), resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies in kind contribution. This publication reflects only the author's views and neither the IMI JU nor EFPIA nor the European Commission are liable for any use that may be made of the information contained therein. For more information, go to <http://www.stembancc.org/>.

Received: July 18, 2013

Revised: October 31, 2013

Accepted: November 13, 2013

Published: December 12, 2013

REFERENCES

- Bar-Nur, O., Russ, H.A., Efrat, S., and Benvenisty, N. (2011). Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell* 9, 17–23.
- Barrero, M.J., Berdasco, M., Paramonov, I., Bilic, J., Vitaloni, M., Esteller, M., and Izpisua Belmonte, J.C. (2012). DNA hypermethylation in somatic cells correlates with higher reprogramming efficiency. *Stem Cells* 30, 1696–1702.
- Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A.J., Socci, N.D., and Scandura, J.M. (2011). DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS ONE* 6, e14524.
- Brunner, A.L., Johnson, D.S., Kim, S.W., Valouev, A., Reddy, T.E., Neff, N.F., Anton, E., Medina, C., Nguyen, L., Chiao, E., et al. (2009). Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Res.* 19, 1044–1056.
- Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet.* 10, 295–304.
- Deng, J., Shoemaker, R., Xie, B., Gore, A., LeProust, E.M., Antosiewicz-Bourget, J., Egli, D., Maherali, N., Park, I.H., Yu, J., et al. (2009). Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat. Biotechnol.* 27, 353–360.
- Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., et al. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* 41, 1350–1353.
- Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and 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, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., et al. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* 6, 479–491.
- Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I., et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* 467, 285–290.
- Le Carrou, T., Assou, S., Tondeur, S., Lhermitte, L., Lamb, N., Reme, T., Pantesco, V., Hamamah, S., Klein, B., and De Vos, J. (2010). Amazonia!: an online resource to Google and visualize public human whole genome expression data. *Open Bioinform. J.* 4, 5–10.
- Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* 20, 311–316.
- Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471, 68–73.
- Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivaachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454, 766–770.
- Morita, S., Takahashi, R.U., Yamashita, R., Toyoda, A., Horii, T., Kimura, M., Fujiyama, A., Nakai, K., Tajima, S., Matoba, R., et al. (2012). Genome-Wide Analysis of DNA Methylation and Expression of MicroRNAs in Breast Cancer Cells. *Int. J. Mol. Sci.* 13, 8259–8272.
- Nagae, G., Isagawa, T., Shiraki, N., Fujita, T., Yamamoto, S., Tsutsumi, S., Nonaka, A., Yoshida, S., Matsusaka, K., Midorikawa, Y., et al. (2011). Tissue-specific demethylation in CpG-poor promoters during cellular differentiation. *Hum. Mol. Genet.* 20, 2710–2721.
- Nishino, K., Toyoda, M., Yamazaki-Inoue, M., Fukawatase, Y., Chikazawa, E., Sakaguchi, H., Akutsu, H., and Umezawa, A. (2011). DNA methylation dynamics in human induced pluripotent stem cells over time. *PLoS Genet.* 7, e1002085.
- Pick, M., Stelzer, Y., Bar-Nur, O., Mayshar, Y., Eden, A., and Benvenisty, N. (2009). Clone- and gene-specific aberrations of parental imprinting in human induced pluripotent stem cells. *Stem Cells* 27, 2686–2690.



- Plath, K., and Lowry, W.E. (2011). Progress in understanding reprogramming to the induced pluripotent state. *Nat. Rev. Genet.* *12*, 253–265.
- Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* *28*, 848–855.
- Stelzer, Y., Yanuka, O., and Benvenisty, N. (2011). Global analysis of parental imprinting in human parthenogenetic induced pluripotent stem cells. *Nat. Struct. Mol. Biol.* *18*, 735–741.
- 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.
- Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* *448*, 196–199.
- Urbach, A., Bar-Nur, O., Daley, G.Q., and Benvenisty, N. (2010). Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell Stem Cell* *6*, 407–411.
- Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* *39*, 457–466.
- Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D., et al. (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* *153*, 1134–1148.
- Ziller, M.J., Müller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C.B., Bernstein, B.E., Lengauer, T., et al. (2011). Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet.* *7*, e1002389.