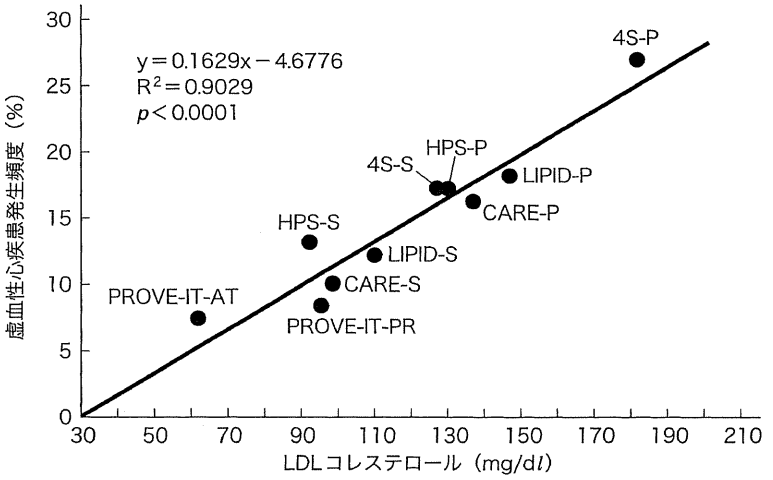


Fig. 1. 海外の二次予防試験における LDL-C 値と冠疾患イベント発症の関係

［文献 9）より引用］

スタチンによる試験にいたるまで，LDL-C 値下降がこれらの患者において総死亡を含めたイベント発生防止に有効であることが，数多くの試験で実証されたといえる[9]．また最近 CTT 共同研究から，2005 年の同じグループからの報告の続報として，1,000 人以上の被験者に 2 年以上の観察を行った 26 の前向き無作為割付試験の解析の結果から，早期に用いられたスタチンと，強力な LDL-C 値低下作用を有する後期スタチンによる LDL-C 値の下降の程度と心血管事故，血管再形成術，脳梗塞の発生リスクの関連を検証している．その結果，LDL-C 値が下降するほど心血管イベントが明らかに低下することが改めて示された[2]．わが国の 2007 年ガイドラインでは二次予防の管理目標は LDL-C＜100 mg/dl であり，「生活習慣改善とともに薬物治療を考慮する」としており，一次予防に比べて薬物治療を早期から考慮することを示しているが，海外の試験に基づく検証では，冠疾患再発率は試験中の LDL-C 値と正相関を示し，LDL-C 値を下降させるほど再発が抑えられること，再発率を完全に抑えるには LDL-C 値を 30 mg/dl 前後にすべきであるとしている（Fig. 1）[9]．

このような背景から，NCEP ATPIII においては虚血性心疾患の既往を有する者は高リスクに分類され，LDL-C の管理目標値は 100 mg/dl もしくは 70 mg/dl 未満としているが，重要なのはこのカテゴリーに含まれるのは，虚血性心疾患既往者以外に冠動脈疾患に相当するリスクとして，糖尿病，末梢動脈疾患，腹部大動脈瘤，頸動脈疾患，二つ以上の冠危険因子をもち，かつ 10 年リスク＞20% の者としていることである．このように，米国ではすでに，二次予防に相当する患者における冠疾患イベントの絶対リスクの成績に基づいて LDL-C の管理目標値を提示していると思われるが，わが国においては二次予防の絶対リスクのデータは J-LIT 研究などに限られているといわざるをえない．わが国の二次予防研究である MUSASHI-AMI では ST 上昇型急性心筋梗塞（AMI）で PCI 施行した 486 例において，AMI 発症後 96 時間以内のスタチン群とスタチン非投与群に無作為割付したが，スタチン群の血清 LDL-C 値は 100 mg/dl 前後が非投与群に比べ有意に低下し，24ヵ月間の観察期間中に複合心血管イベントリスクの 42% 下降が認められたとしている[10]．

対　　案

以上の状況に基づいて，以下に対案を述べる．

Table 2. わが国の血管内超音波(IVUS)を用いた二次予防試験

試験名	ESTABLISH*1 2004	JAPAN-ACS*2 2008	COSMOS*3 2009
介　入	atorvastatin 20 mg vs control (*n*= 24/24)	atorvastatin 20 mg vs pitavastatin 4 mg (*n*= 127/125)	rousvastatin 2.5～20 mg (*n*= 126)
対　象	急性冠症候群 緊急 PCI	急性冠症候群 緊急 PCI	安定虚血性心疾患 待機的 PCI
LDL-C 値 (mg/d*l*) (% 変化)	125→70(41.7%)	132→84(36.0%)	140→83(38.6%)
プラーク退縮%	13.1%	17.5%	5.1%

*1 Okazaki S et al：Circulation 110(9)：1061, 2004.
*2 Miyauchi K et al：Circ J 70(12)：1624, 2006.
*3 Takayama T et al：Circ J 71(2)：271, 2007.

1）一次予防における血清脂質管理基準の策定には，NIPPON DATA80 などの疫学研究による虚血性心疾患発症などの絶対リスクを考慮すべきである．

2）二次予防については，ATPIII と同様，LDL-C 値の管理目標を 70 mg/d*l* 未満にする．

3）また，糖尿病の合併は二次予防と同様に扱う．

両者の比較

1）一次予防における血清脂質管理には，最近の疫学研究成果の解析による絶対リスクの導入の可能性が出てきているが，NIPPON DATA 80 のデータには低 HDL-C 血症，CKD，冠動脈疾患の家族歴などのリスクが含まれておらず，これらのリスクをどのように NIPPON DATA 80 の成績に反映させるかが問題となる．

2）わが国の二次予防試験は限られているが，J-LIT 研究の二次予防では心臓死＋非致死性心筋梗塞の発症率は 4.97/1,000 人・年であった．J-LIT 研究は無作為割付試験ではなく，原則的にスタチン治療下にあったといえるが，これを配慮しても米国の 10 年間に 20% 以上という発症率に比べてはるかに低い．最近のわが国の二次予防試験では IVUS によるプラーク容積の退縮に関する研究が相次いで発表された(Table 2)．これらによると LDL-C 値を 70～80 mg/d*l* に下降させるとプラーク容積の退縮を認めるとしている．特記すべきこととしては，後述するように糖尿病を合併する患者においてはプラークの退縮が抑制されることが示されており，二次予防においても患者の個々にもつリスクの層別化が必要となると思われる．

3）糖尿病を虚血性心疾患既往と同様に扱うべきかについては，わが国の研究では十分な論拠が得られていないのが現況である．前述の J-LIT 研究では糖尿病は虚血性心疾患既往と同等のリスクとは判定できなかった．現在 JDCS などの前向き研究が進行しており，その成績から糖尿病のリスクとしての評価が行われると思われる．わが国では，しかしながら，糖尿病は増加するリスクとしてもっとも憂慮すべきであり，ガイドライン改訂に際しても十分な検討を行う必要があると考える．

ま と め

虚血性心疾患の一次ならびに二次予防における脂質管理目標につき，現状とその問題点，課題ならびに想定される提案と判定について述べた．一次予防については絶対リスクの評価をいかに考慮

するか，二次予防については個々の患者のもつリスクの層別化に基づいてLDL-Cの管理目標値をどこまで下げるべきかが今後の検討課題である．わが国におけるエビデンスに基づいたガイドラインが望まれる．

文　献

1) Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). JAMA **285** (19) : 2486, 2001
2) Baigent C et al : Efficacy and safety of more intensive lowering of LDL cholesterol : a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet **376** (9753) : 1670, 2010
3) 日本動脈硬化学会：動脈硬化性疾患予防ガイドライン2007年版，東京，2007
4) Shepherd J et al : Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia : West of Scotland Coronary Prevention Study Group. N Engl J Med **333** (20) : 1301, 1995
5) Ridker PM et al : Reduction in C-reactive protein and LDL cholesterol and cardiovascular event rates after initiation of rosuvastatin : a prospective study of the JUPITER trial. Lancet **373** (9670) : 1175, 2009
6) Nakamura H et al : Primary prevention of cardiovascular disease with pravastatin in Japan (MEGA Study) : a prospective randomised controlled trial. Lancet **368** (9542) : 1155, 2006
7) Taylor F et al : Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev, 2011 (1) : CD004816
8) Grundy SM et al : Implications of recent clinical trials for the National Cholesterol Education Program Adult Treatment Panel III Guidelines. J Am Coll Cardiol **44** (3) : 720, 2004
9) O'Keefe JH Jr et al : Optimal low-density lipoprotein is 50 to 70 mg/dl : lower is better and physiologically normal. J Am Coll Cardiol **43** (11) : 2142, 2004
10) Sakamoto T et al : Effects of early statin treatment on symptomatic heart failure and ischemic events after acute myocardial infarction in Japanese. Am J Cardiol **97** (8) : 1165, 2006

DNA Methylation Dynamics in Human Induced Pluripotent Stem Cells over Time

Koichiro Nishino, Masashi Toyoda, Mayu Yamazaki-Inoue, Yoshihiro Fukawatase, Emi Chikazawa, Hironari Sakaguchi, Hidenori Akutsu, Akihiro Umezawa*

Department of Reproductive Biology, National Institute for Child Health and Development, Tokyo, Japan

Abstract

Epigenetic reprogramming is a critical event in the generation of induced pluripotent stem cells (iPSCs). Here, we determined the DNA methylation profiles of 22 human iPSC lines derived from five different cell types (human endometrium, placental artery endothelium, amnion, fetal lung fibroblast, and menstrual blood cell) and five human embryonic stem cell (ESC) lines, and we followed the aberrant methylation sites in iPSCs for up to 42 weeks. The iPSCs exhibited distinct epigenetic differences from ESCs, which were caused by aberrant methylation at early passages. Multiple appearances and then disappearances of random aberrant methylation were detected throughout iPSC reprogramming. Continuous passaging of the iPSCs diminished the differences between iPSCs and ESCs, implying that iPSCs lose the characteristics inherited from the parent cells and adapt to very closely resemble ESCs over time. Human iPSCs were gradually reprogrammed through the "convergence" of aberrant hyper-methylation events that continuously appeared in a de novo manner. This iPS reprogramming consisted of stochastic de novo methylation and selection/fixation of methylation in an environment suitable for ESCs. Taken together, random methylation and convergence are driving forces for long-term reprogramming of iPSCs to ESCs.

Citation: Nishino K, Toyoda M, Yamazaki-Inoue M, Fukawatase Y, Chikazawa E, et al. (2011) DNA Methylation Dynamics in Human Induced Pluripotent Stem Cells over Time. PLoS Genet 7(5): e1002085. doi:10.1371/journal.pgen.1002085

Editor: John M. Greally, Albert Einstein College of Medicine, United States of America

Received December 2, 2010; **Accepted** April 1, 2011; **Published** May 26, 2011

Funding: This research was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan; by Ministry of Health, Labour, and Welfare Sciences (MHLW) research grants; by a Research Grant on Health Science focusing on Drug Innovation from the Japan Health Science Foundation; by the program for the promotion of Fundamental Studies in Health Science of the Pharmaceuticals and Medical Devices Agency; by a Grant for Child Health and Development from the MHLW; by the Intramural Research Grant (22-5) for Neurological and Psychiatric Disorders of NCNP; by the Research Grant (22-2-4) for cardiovascular disease of NCVC given to AU; by a grant from New Energy and Industrial Technology Development Organization (NEDO) in Japan given to HA; and by Grant-in-Aid for Young Scientist(B)(WAKATE- B 21790372) given to KN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: umezawa@1985.jukuin.keio.ac.jp

Introduction

DNA methylation is an important epigenetic modification and is a key component in normal differentiation, development and disease [1–3]. Expression of tissue-specific genes, such as *Oct-4* [4], *Nanog* [5], *Sry* (sex determining region on Y chromosome) [6] and *MyoD* [7], are induced by spatio-temporal demethylation during development. DNA methylation therefore specifically varies depending on tissue types and cell linage [2], indicating that information regarding cell type-specific DNA methylation profiles can enable the identification and validation of cell types. Transformation of iPSCs from somatic cells requires a process of epigenetic reprogramming promoted by transient ectopic expression of defined transcription factors expressed in ESCs [8–11]. Human iPSCs are considered to be powerful resources in regenerative medicine because of their potential of pluripotency and avoidance of rejection of their derivatives by the immune system, and for ethical issues as well [12]. Although iPSCs show pluripotency, they have different propensities for differentiation in mouse models [13]. Human iPSCs also exhibit donor cell-specific gene expression [14,15]. Moreover, iPSCs possess inherited DNA methylation states as epigenetic memories from parent cells [15–17], suggesting that these memories influence different propensities of the iPSCs. On the other hand, continuous passaging of mouse iPSCs reduces differences from each other in gene expression profiles [15]. Epigenome-wide analysis started to be used in this field [18,19], and differentially methylated regions have been identified among human iPSCs, their parent cells and ESCs [17,20]. Aberrant epigenetic reprogramming has recently been reported in human iPSCs [21,22]. However, these analyses were limited to the use of a small number of cells as a source for generation of iPS cells. Moreover, human iPSCs have only been analyzed at a single point of passage. Therefore, it has not been clarified whether human iPSCs generated from various types of cells are dissimilar from each other at different points during passage; how continuous passaging of human iPSCs influences the differences between iPSCs and ESCs; and how aberrant methylation in human iPSCs during passaging. To address these issues, we compared the epigenetic and transcriptional states of human iPSCs derived from five cell types of different origins during passage, and found random aberrant hyper-methylation at different points of adaptation into ESCs.

Results

Establishment of human iPSCs

Human iPSCs derived from fetal lung fibroblasts (MRC5), amnion (AM), endometrium (UtE), placental artery endothelium (PAE) and menstrual blood cells (Edom) were independently established in our laboratory by retroviral infection of 4 genes

Author Summary

iPSCs change to resemble ESCs via two phases: the transgene-dependent phase, in which the transcription factors act to transform somatic cells into pluripotent stem cells, and the transgene-independent phase, in which the transcription factors are silenced. In this study, we established human iPSCs derived from 5 different cell types by retroviral infection of the Yamanaka 4 factors, and we identified 8 novel epigenetic markers (*SALL4, EPHA1, PTPN6, RAB25, GBP4, LYST, SP100,* and *UBE1L*) by comprehensive DNA methylation analysis. The aberrant hyper-methylation in iPSCs occurred stochastically throughout the genome and decreased during the long-term iPSC reprogramming, suggesting that the aberrant stochastic hyper-methylation and their convergence are a direct cause of the transgene-independent phase of iPS reprogramming. These results favor the stochastic model of the Yamanaka model rather than the elite model. In addition, the stem cell–specific methylation states and the epigenetic difference between iPSCs and ESCs are useful indices for evaluating human iPSCs in therapeutic applications.

(*OCT-3/4*, *SOX2*, *c-MYC*, and *KLF4*) (Figure 1A, 1B and Table S1). These cells clearly showed human ES-like characters in terms of morphology; cell-surface antigens; gene expression of stem cell markers; teratoma formation in which these cells differentiated to various tissues including neural tissues (ectoderm), cartilage (mesoderm), and epithelial tissues (endoderm); growth (more than 20 passages); and DNA methylation patterns at *OCT-3/4* and *NANOG* promoter regions (Figures S1, S2, S3). Short tandem repeat (STR) analysis showed clonality between the respective iPSC lines and their parent cells (Table S2). Silencing of transgenes and normal karyotypes of iPSCs were also confirmed (Figure S4 and Table S3).

Analysis of DNA methylation profiles

To investigate the dynamics of DNA methylation in pluripotent stem cells, we examined 5 ESC lines (HUESCs) [23,24], 22 iPSC lines, their parent cells and 201B7, using Illumina's Infinium HumanMethylation27 BeadChip. In total, 24,273 CpG sites in 13,728 genes were analyzed, along with 33 human cell lines (Table S1). The iPSC line "201B7" was generated from human skin fibroblasts [8]. Quantitative scores of DNA methylation levels were obtained as β-values determined from the Illumina analysis, ranging from "0", for completely unmethylated, to "1", for completely methylated. We also performed genome-wide gene expression analysis using the Agilent Whole Human Genome Microarray chips. As assessed by unsupervised hierarchical clustering analysis and scatter plot of DNA methylation and gene expression data, human iPSCs could be clearly discriminated from their parent cells and were similar to ESCs (Figure 1C and Figure

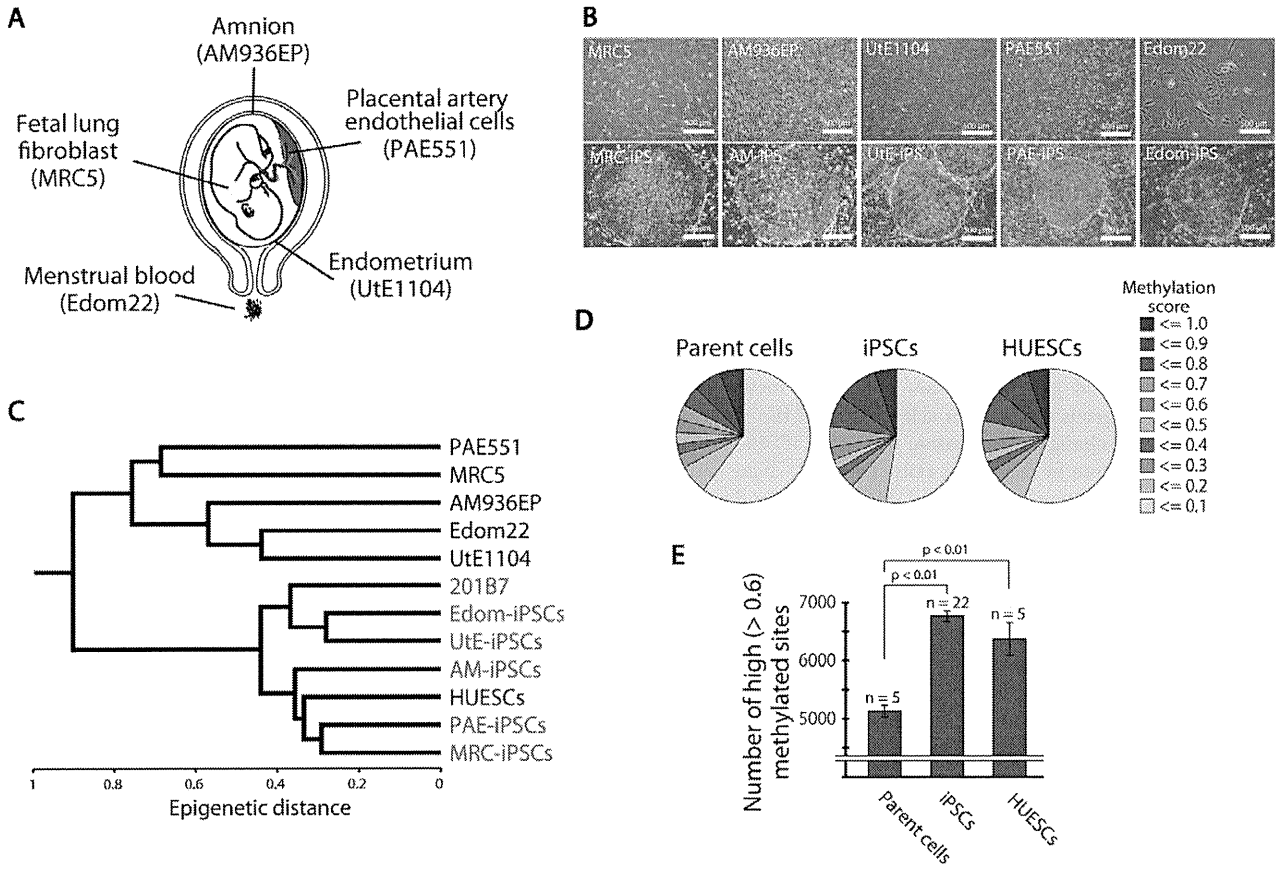


Figure 1. Pluripotent stem cells are significantly more hyper-methylated than their parent cells. (*A*) The human cell origins used for generation of iPSCs. (*B*) Morphology of the parent cells (upper panels) and iPSCs (lower panels). (*C*) Unsupervised hierarchical clustering analysis based on DNA methylation. (*D*) Distribution of 24,273 CpG sites with their methylation scores in the parent cells, iPSCs and ESCs. (*E*) The average number of high (>0.6) methylated CpG sites. The iPSCs have more highly methylated sites than the parent cells.
doi:10.1371/journal.pgen.1002085.g001

S5). The distribution of DNA methylation levels shows that the degree of global methylation in pluripotent stem cells was higher compared to the parent cells (Figure 1D, 1E), suggesting that a global gain of DNA methylation occurs during reprogramming.

Identification of stem cell-specific differentially methylated regions (DMRs)

For further analysis, we defined DMR as representing a CpG site whose score differed 0.3 points or more from the β-value between the two groups. By comparison among ESCs (average from 5 lines), iPSCs (average from 22 lines), and parent cells (average from 5 lines), about 90% of the CpG sites (17,572 sites) examined did not show differential methylation among ESCs, iPSCs and parent cells (Figure 2A), suggesting that only a small number of the CpG sites is affected during reprogramming. The number of the CpG sites has been reported to be larger by genome-wide analysis [21].

We then identified 220 sites that are pluripotent stem cell-specific DMRs (Figure 2A). The 174 sites (79.5%) of the stem cell-specific DMRs had significantly higher methylation levels in iPSCs/ESCs when compared to the parent cells (Figure 2B). Approximately 80% of the DMRs between the iPSCs and their parent cells changed to a "hyper-methylated" state from a "hypo-methylated" state in iPSCs. In contrast, 45 sites of the stem cell-specific DMRs are hypo-methylated in iPSCs/ESCs, compared with the parent cells. Gene ontology analysis indicates that the hypo-methylated stem cell-specific DMRs especially included genes related to mRNA transcription regulation (Figure 2B). Interestingly, the majority of the hypo-methylated stem cell-specific DMRs were located on CpG islands, whereas the majority of the hyper-methylated stem cell-specific DMRs were located on non-CpG islands (Figure 2C). No iPS-specific DMRs were detected. We extracted 3,123 sites that are differentially methylated in one or more parent-specific iPSCs, compared to their parent cells, because DMRs are dependent on parent cell types (Figure S6). These DMRs are here designated as stem cell-required DMRs. Distribution analysis of the stem cell-required DMRs revealed a dispersed pattern rather than specific localization on the genome (Figure S7A).

From the combined gene expression and DNA methylation data, we chose 27 genes in the stem cell-specific DMRs showing more than a 5-fold change in expression of human iPSCs/ESCs, as compared with those in the parent cells (Table S4). Nine genes with hypo-methylated stem cell-specific DMRs were found in the group "genes significantly expressed in iPSCs/ESCs," and 17 genes with hypo-methylated stem cell-specific DMRs belonged to the category "low expression or silenced in iPSCs/ESCs". In addition, the methylation state and gene expression in *EPHA1*, *PTPN6*, *RAB25*, *SALL4*, *GBP3*, *LYST*, *SP100* and *UBE1L* were confirmed by quantitative combined bisulfite restriction analysis (COBRA) [25] (Figure 2D), RT-PCR (Figure 2E) and bisulfite sequencing (Figure 2F).

We also extracted genes with stem cell-required DMRs exhibiting high expression or suppression in human iPSCs/ESCs (Tables S5, S6). Interestingly, gene ontology analysis of the genes with stem cell-required DMRs showed that genes in the transcription factor category were detected only in the hypo-methylated stem cell-required DMRs (Table S7). The top 20 transcription factor genes with hypo-methylated stem cell-required DMRs exhibiting high expression in human iPSCs are summarized in Table 1 and include *OCT-4/3* (also known as *POU5F1*), *SALL4*, *SOX8*, *ZIC5*, and *FOXD1*.

Aberrant and inherited methylation in iPSCs

Few changes in DNA methylation were detected between iPS and ES cells and these were not consistent among the different iPS lines (Figure 2A, Figures S6, S7). In further analyses, we compared the DNA methylation states of each iPSC line or each parent cell line with that of ESCs (averaged value) (Figure 3A). For the whole genome, the number of DMRs between ESCs and iPSCs (ES-iPS-DMRs) varied in the 22 iPSC lines (Figure 3B). A comprehensive analysis of methylation in ESCs and iPSCs identified 1,459 ES-iPS-DMRs covering 1,260 genes that were differentially methylated in one or more iPSC lines. ES-iPS-DMRs are composed of aberrant (iPS-specific) methylation sites, in comparison with ESCs and inherited methylation sites from the parent cells. The number of inherited sites as well as aberrant sites varied among iPSCs. Analysis of the ES-iPS-DMRs on each chromosome showed a characteristic distribution of the ES-iPS-DMRs on the X chromosome in XX-iPSCs (Figure 3B and Figure S8). Female XX-iPSCs demonstrate a tendency to carry a large number of ES-iPS-DMRs on the X chromosome, but male XY-iPSCs had few ES-iPS-DMRs on the X chromosome (Figure 3B, lower panel). While no ES-iPS-DMRs overlapped for all the iPSCs (Figure 2A), 20 ES-iPS-DMRs overlapped in more than 15 out of 22 lines (Figure 3C, inset). These 20 ES-iPS-DMRs include the genes for *MPG* (N-methylpurine-DNA glycosylase isoform b), *FZD10* (frizzled 10), *IREX2* (iroquois homeobox protein 2) and *ZNF248* (zinc finger protein 248), which are highly associated with aberrant methylation during reprogramming. Distribution analysis of the ES-iPS-DMRs across the genome did not show any specific localization (Figure S9). We further compared overlapping ES-iPS-DMRs in reference to a genome-wide methylation analysis [21], and found that 72 gene promoters overlapped between our data and that of Lister et al..

More than 70% of the ES-iPS-DMRs were hyper-methylated in each iPSC (Figure 3D), indicating that the iPSC genome is more methylated than the ESC genome. In addition, the majority of the ES-iPS-DMRs were located on CpG islands (Figure 3E), suggesting that aberrant methylation is biased towards CpG islands.

Effect of long-term culture on DNA methylation status in iPSCs

We investigated the effect of continuous passaging on the DNA methylation profile of human iPSCs. To address the effect, we subjected 7 iPSC lines to additional rounds of passaging under identical culture conditions, and obtained genomic DNA and RNA at passage 4 (P4) to P40 for DNA methylation and gene expression. The number of the ES-iPS-DMRs ranged from 80 in MRC-iPS-25 to 286 in UtE-iPS-11 at early passage (P10 to P20), whereas the number of the ES-iPS-DMRs dramatically decreased in all lines at late passage (P30 to P40) (Figure 4A, upper-left panel). The number of inherited and aberrant sites decreased to 30 and 70, respectively, at P30 to P40 (Figure 4A, upper-center and right panels). These decreases in the numbers of ES-iPS-DMRs indicate that iPSCs have become closer to ESCs in their DNA methylation profiles. In particular, XX-iPSC lines (AM-iPS-8, UtE-iPS-4 and -11, and Edom-iPS-2) showed decreases in the number of ES-iPS-DMRs with passaging. The XY-iPSC lines, such as MRC-iPS-25 and PAE-iPS-1, had only a small number of ES-iPS-DMRs. The number of ES-iPS-DMRs continued to decrease to approximately 100 ES-iPS-DMRs containing 30 inherited sites. Intriguingly, few ES-iPS-DMRs on the X chromosome were detected in XY-iPSCs throughout the passaging. In contrast, the number of ES-iPS-DMRs in XX-iPSCs ranged from 10 to 70 at the early passage (P4 to P20), and decreased to zero after P30 (Figure 4A, lower panels). We also

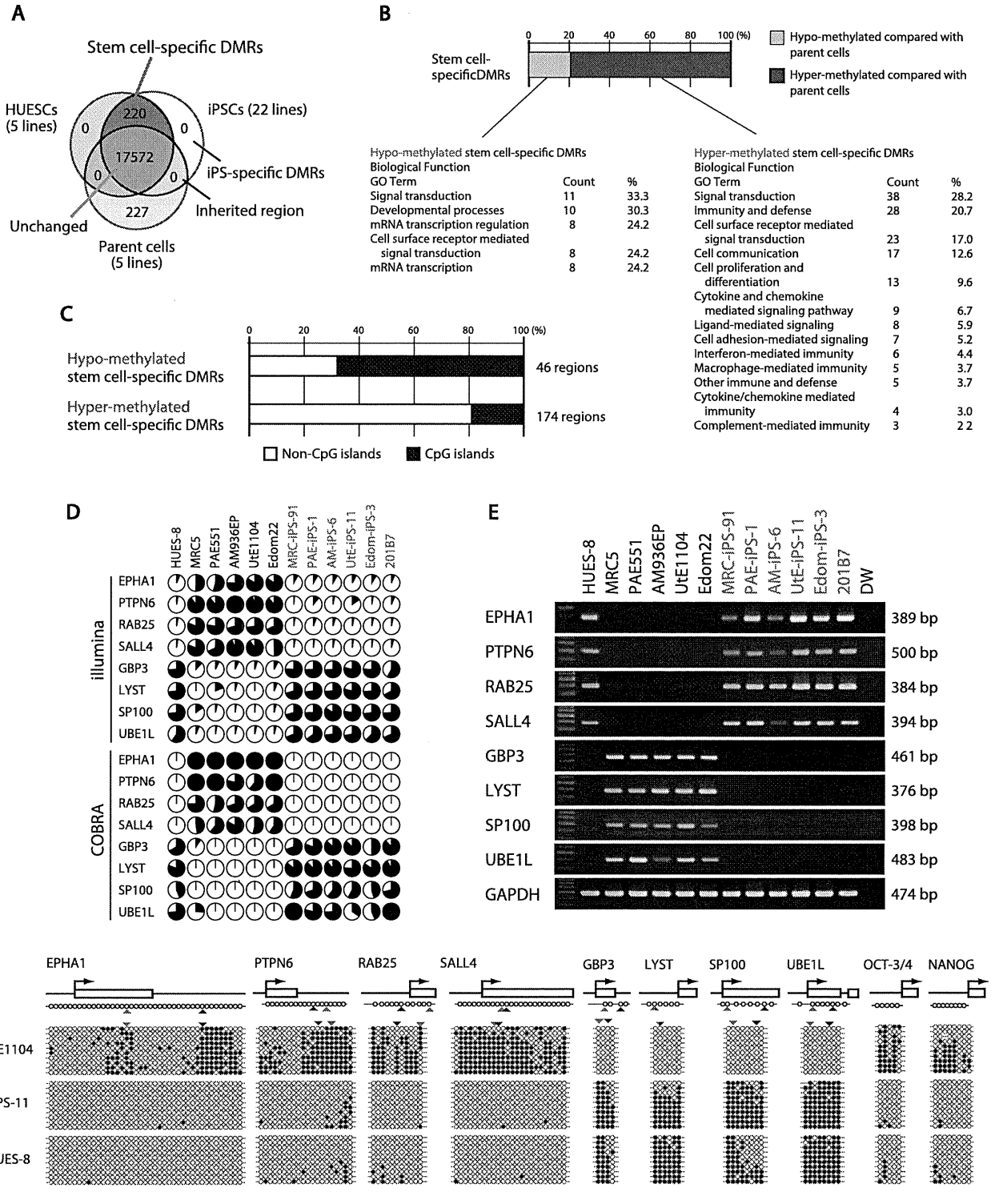


Figure 2. Defining stem cell-specific DMRs as novel epigenetic iPS markers. (*A*) Venn-like diagram showing overlapping CpG sites among ESCs, iPSCs and their parent cells. The 220 overlapping sites are stem cell-specific differentially methylated regions (DMRs). Notably, neither overlapping iPSCs-specific DMRs nor inherited regions in iPSCs from the parent cells were observed. (*B*) Proportion of the hyper- and hypo-methylated stem cell-specific DMRs and GO analysis. Approximately 80% of the regions were hyper-methylated in iPSCs, compared with that of the parent cells. (*C*) Proportion of the regions associated with CpG islands and non-CpG islands in the hypo-methylated stem cell-specific DMRs. The hypo-methylated regions were biased to CpG islands, whereas the hyper-methylated regions were biased to non-CpG islands. (*D*) DNA methylation levels in the 8 representative genes determined by Illumina Infinium HumanMethylation27 assay and Bio-COBRA. These 8 genes were defined as SS-DMRs with significant changes of expression and were described in Table S6. The relative amount of methylated and unmethylated DNA ratio is indicated as the black and white area, respectively, in the pie chart. (*E*) Expression of the 8 genes. Expression of the 8 genes had an inverse correlation with DNA methylation level. (*F*) Bisulfite sequencing analysis of the 8 genes in endometrial cells (UtE1104), UtE-iPS-11 and HUES-8 cells. (Top)

Schematic diagram of the genes. Arrows, open boxes and open circles represent transcription start site, first exon and position of CpG sites, respectively. (Bottom) Open and closed circles indicate unmethylated and methylated sites, respectively. Red and blue arrowheads represent the position of CpG sites in Infinium assay and COBRA assay, respectively.
doi:10.1371/journal.pgen.1002085.g002

investigated the effect of continuous passaging on the DNA methylation profile of the parent cells (UtE1104 and Edom22) (Figure 4B). The number of the DMRs between ESCs and parent cells (ES-parent-DMRs) increased with passaging. In addition, we also confirmed that the transgenes were silenced at each passage (Figure 4C and Figure S4), indicating that the decreasing number of the ES-iPS-DMRs in iPSCs occurred in the transgene-independent phase.

Comparative analysis of ES-iPS-DMRs dynamics

We then compared each ES-iPS-DMRs with passaging. The UtE-iPS-11 had 286 ES-iPS-DMRs at P13, 194 sites at P18, 110 sites at P31, and 55 sites at P39. The ES-iPS-DMRs detected at P13 decreased with passaging (blue bars in upper-left panel in Figure 5A). Interestingly, 66 de novo ES-iPS-DMRs appeared at P18, while at P13 these sites showed no differences between UtE-iPS-11 and ESCs (orange bars in upper-left panel in Figure 5A). These 66 ES-iPS-DMRs also decreased with passaging (P31 and P39). The 29 additional ES-iPS-DMRs at P31 also appeared and decreased with passaging (P39) (green bars in upper-left panel in Figure 5A) and 16 ES-iPS-DMRs at P39 (red bar in upper-left panel in Figure 5A) appeared. Rapid appearance and gradual disappearance of ES-iPS-DMRs was a recurring theme, but the number of newly-appearing ES-iPS-DMRs decreased with passaging (Figure 5A, upper-left panel). The same change in ES-iPS-DMRs occurred on the X chromosome, but the number of the ES-iPS-DMRs approached zero at early passages (Figure 5A, upper-center panel). Intriguingly, this change also occurred at inherited sites, which was contrary to our expectations. The inherited sites also repeatedly appeared and disappeared, and the number of newly-appearing inherited sites decreased with passaging (Figure 5A, upper-right panel). The term "inherited" is here used to mean the same methylation state found in iPSCs and their parent cells, but the "inherited" regions behaved like "aberrant" regions that had multiple appearances and disappearances. These multiple appearances/disappearances of ES-iPS-DMRs were observed in all iPSC lines regardless of parental cell type. The ES-parent-DMRs were also analyzed. The *de novo* ES-parent-DMRs appeared as well as the ES-iPS-DMRs, but did not decrease with passaging (Figure 5B).

Most ES-iPS-DMRs were hyper-methylated in iPSCs

ES-iPS-DMRs can be categorized into two groups: a, hyper-methylated and b, hypo-methylated sites in iPSCs, as compared

Table 1. List of the top 20 out of 82 transcription factor genes with hypo-methylated stem cell-required DMRs exhibiting "high" expression in human iPS cells.

		DNA methylation		
TargetID	**Gene name**	**HUESCs**	**iPSCs**	**Expression level**
cg13083810	POU5F1, POU domain; class 5; transcription factor 1 isoform 1	0.584	0.549	55543.9
cg06303238	SALL4, sal-like 4	0.032	0.026	29766.2
cg16990174	RYBP, RING1 and YY1 binding protein	0.076	0.119	10274.1
cg03589001	MORF4L1, MORF-related gene 15 isoform 2	0.176	0.173	7015.7
cg02204046	MYCN, v-myc myelocytomatosis viral related oncogene; neuroblastoma derived	0.022	0.027	5826.8
cg10705800	CITED4, Cbp/p300-interacting transactivator; with Glu/Asp-rich carboxy-terminal domain; 4	0.438	0.445	5342.2
cg21696393	SOX8, SRY (sex determining region Y)-box 8	0.074	0.061	1976.7
cg23131007	TCF12, transcription factor 12 isoform b	0.138	0.155	1930.7
cg18808261	SATB1, special AT-rich sequence binding protein 1	0.194	0.242	1634.4
cg15607672	OTX2, orthodenticle 2 isoform a	0.046	0.054	1227.5
cg05345286	MDFI, MyoD family inhibitor	0.023	0.040	1035.9
cg20909686	OVOL1, OVO-like 1 binding protein	0.215	0.204	991.0
cg26209676	ZNF581, zinc finger protein 581	0.113	0.196	916.1
cg05522383	PITX2, paired-like homeodomain transcription factor 2 isoform b	0.024	0.030	544.8
cg17675150	ZNF532, zinc finger protein 532	0.069	0.107	525.3
cg01510051	ZNF542, zinc finger protein 542	0.585	0.555	443.9
cg06154570	HEYL, hairy/enhancer-of-split related with YRPW motif-like	0.134	0.152	440.3
cg12556134	TGIF2, TGFB-induced factor 2	0.075	0.072	405.4
cg03663715	FOXD1, forkhead box D1	0.030	0.042	349.1
cg09721427	HHEX, hematopoietically expressed homeobox	0.077	0.101	206.9

"Expression level" is an average of raw data values in iPSCs from Gene Chip data.
doi:10.1371/journal.pgen.1002085.t001

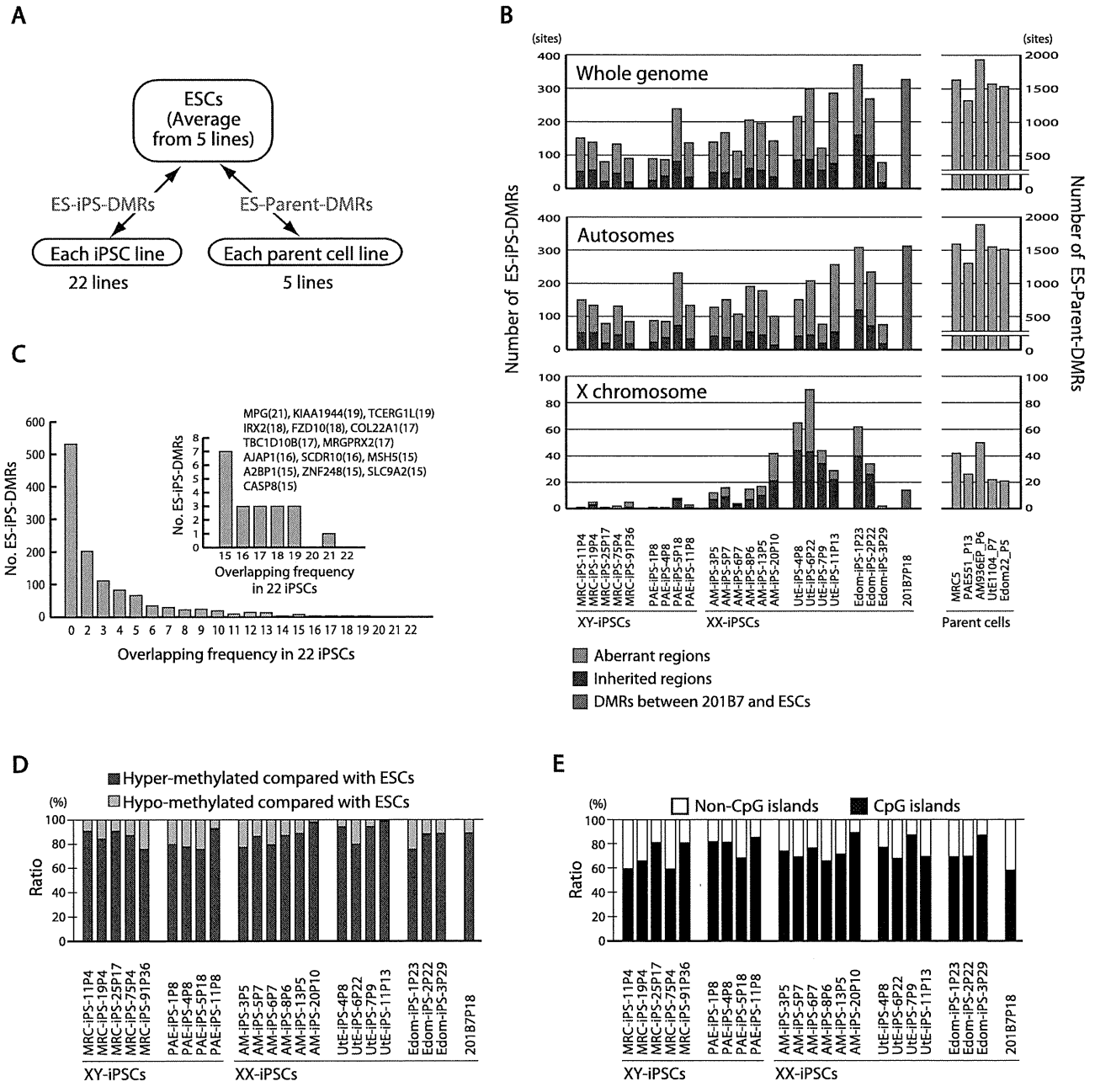
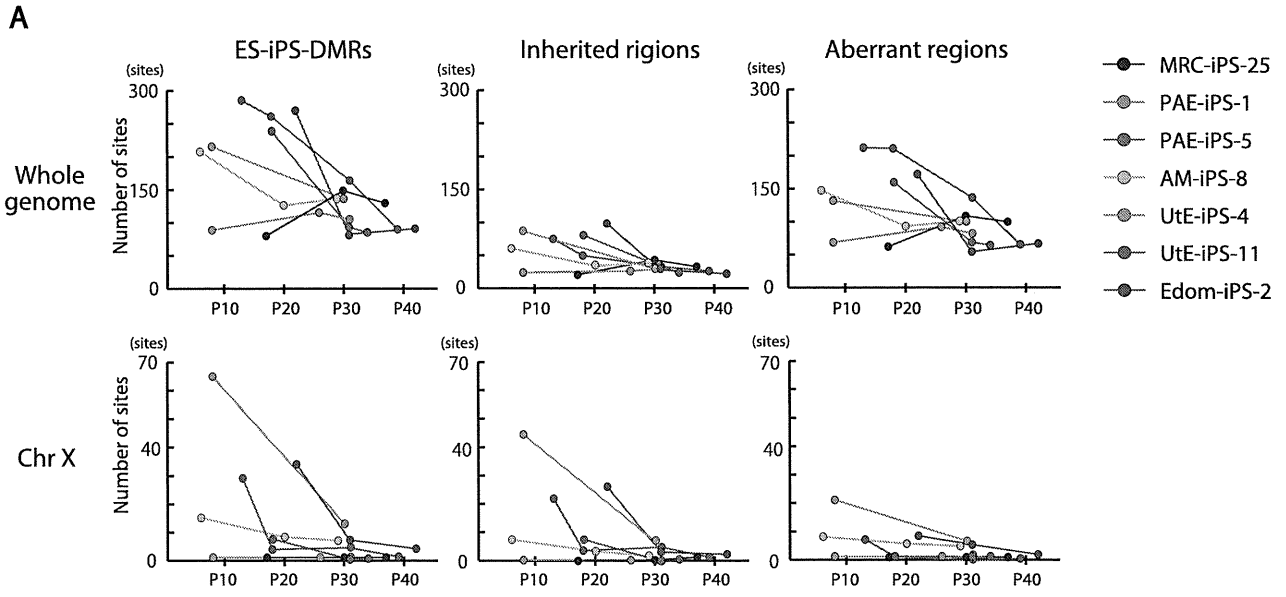
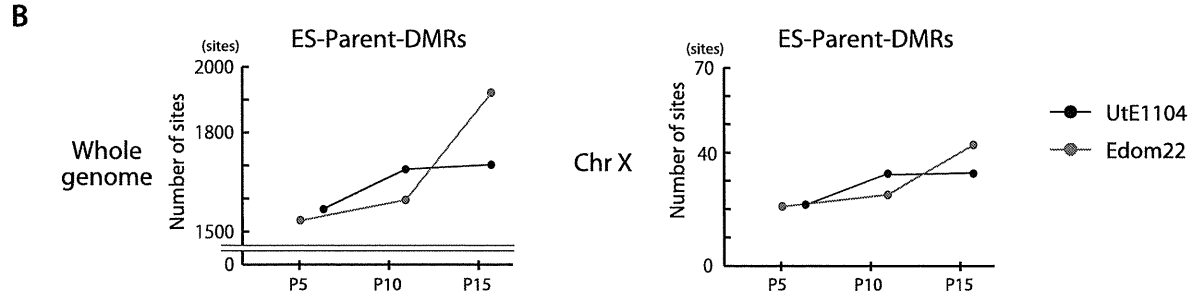


Figure 3. Aberrant methylation in human iPSCs. (*A*) Comparison of DNA methylation states of each iPSC line or each parent cell line with that of ESCs. The DMRs between ESCs and iPSCs are designated as ES-iPS-DMRs, and the DMRs between ESCs and parent cells are designated as ES-parent-DMRs. (*B*) The number of ES-iPS-DMRs and ES-parent-DMRs on whole genome (top), autosomes (middle) and X chromosome (bottom). Ratios of number of inherited regions in iPSCs from parent cells (blue) and aberrant regions in iPSCs that differ from ESCs and parent cells (red) in the ES-iPS-DMRs are shown in bars. Female iPSCs were demonstrated to carry high number of EiP-DMRs on X chromosome. (*C*) Number of overlapped ES-iPS-DMRs frequency in iPSCs. No overlapping ES-iPS-DMRs in all 22 iPSC lines. (Inlet) A small number of overlapping ES-iPS-DMRs of the frequency from 15 to 22. Overlapping frequency of each gene is indicated in parentheses. (*D*) Proportion of the hyper- and hypo-methylated ES-iPS-DMRs. More than 75% of the ES-iPS-DMRs were hyper-methylated in iPSCs. (*E*) Proportion of the ES-iPS-DMRs associated with CpG islands and non-CpG islands in ach iPSC line. ES-iPS-DMRs were biased to CpG islands.
doi:10.1371/journal.pgen.1002085.g003

with ESCs. ES-iPS-DMRs that disappeared at the last passage (P39) (blue bars in Figure 5) in both UtE-iPS-11 and Edom-iPS-2 were extracted, and each methylation score of the extracted ES-iPS-DMRs is shown (Figure 6, upper and middle panels). To compare methylation scores, a "difference value" was estimated by subtracting the scores of ESCs from those of each cell (Figure 6, lower panels). Positive and negative difference values indicate that these sites are hyper- and hypo-methylated, respectively, when compared with ESCs. Difference values of the ES-iPS-DMRs showing aberrant methylation states in iPSCs at the early passage approached zero with passaging. It should be noted that the almost all difference values became largely positive in iPSCs at


A
ES-iPS-DMRs
Inherited rigions
Aberrant regions
Whole genome
Chr X
Number of sites
(sites)
300
150
0
70
40
P10
P20
P30
P40
MRC-iPS-25
PAE-iPS-1
PAE-iPS-5
AM-iPS-8
UtE-iPS-4
UtE-iPS-11
Edom-iPS-2


B
ES-Parent-DMRs
Whole genome
Chr X
Number of sites
(sites)
2000
1800
1500
0
70
40
P5
P10
P15
UtE1104
Edom22

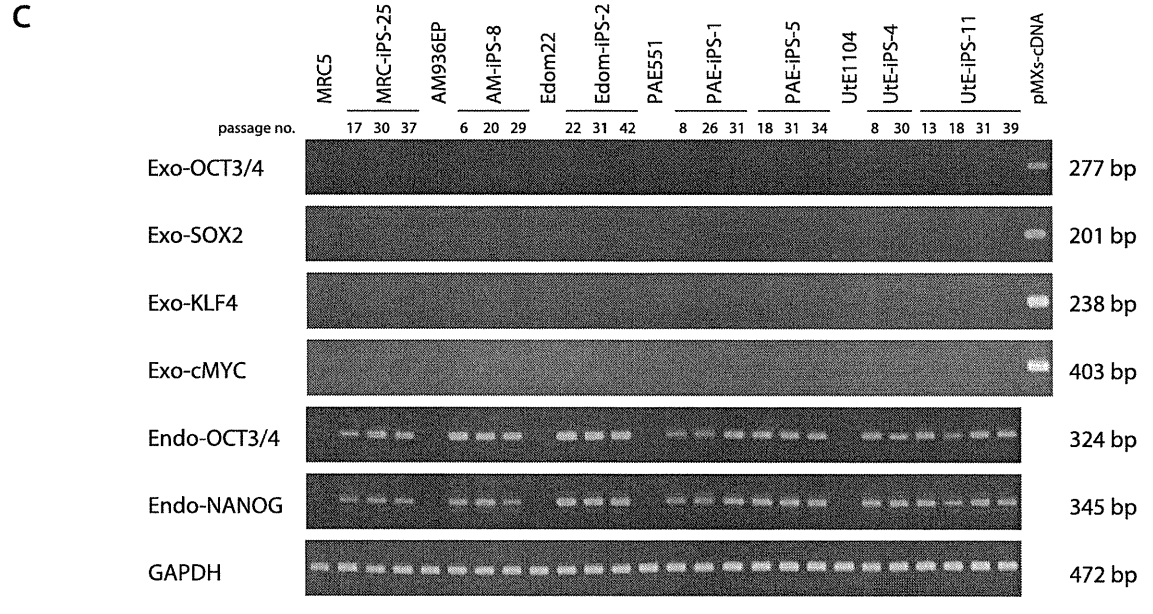

C
MRC5
MRC-iPS-25
AM936EP
AM-iPS-8
Edom22
Edom-iPS-2
PAE551
PAE-iPS-1
PAE-iPS-5
UtE1104
UtE-iPS-4
UtE-iPS-11
pMXs-cDNA
passage no.
17 30 37
6 20 29
22 31 42
8 26 31
18 31 34
8 30
13 18 31 39
Exo-OCT3/4
Exo-SOX2
Exo-KLF4
Exo-cMYC
Endo-OCT3/4
Endo-NANOG
GAPDH
277 bp
201 bp
238 bp
403 bp
324 bp
345 bp
472 bp

Figure 4. Effect of long-term cultivation on ES-iPS-DMRs. (*A*) Decrease in the number of the ES-iPS-DMRs with continuous passaging. Upper panels show change of the number of the ES-iPS-DMRs (left), the inherited regions (middle) and aberrant regions (right) on whole genome. Lower panels show change in the number of the ES-iPS-DMRs (left), inherited regions (middle) and aberrant regions (right) on X chromosome. The number of the ES-iPS-DMRs in XX-iPSCs approached zero with continuous passaging on X chromosome. In contrast, XY-iPSCs had few ES-iPS-DMRs on X chromosome throughout the passages. (*B*) The number of the ES-parent-DMRs with continuous passaging. (*C*) No expression of the transgenes in iPSCs at each passage was detected by RT-PCR.
doi:10.1371/journal.pgen.1002085.g004

early passage (P13 or P22), even though they were negative in the parent cells, and then approached zero upon further passaging. This transiently-induced hyper-methylation was observed at each passage in all iPSC lines examined. The observed transient hypermethylation patterns during iPS reprogramming did not correspond to methylated CpGs in the parental cells. However, this observation does not rule out that transient aberrant methylation could also be observed in some cases on sites that were methylated in the parental cells.

Discussion

Identification of novel epigenetic iPS markers

OCT-4/3 and *NANOG* have been used as epigenetic markers for iPSCs [8–10,26,27]. We previously showed candidate epigenetic markers by analyzing 6 iPS lines [17]. Here we identified 8 novel epigenetic markers more closely by defining 9 genes with the hypo-methylated stem cell-specific DMRs and significantly higher expression, and 17 genes with the hyper-methylated stem cell-specific DMRs and significantly lower expression in iPSCs/ESCs from 22 iPS lines. DNA methylation and expression of these genes, especially the 8 genes, *SALL4*, *EPHA1*, *PTPN6*, *RAB25*, *GBP4*, *LYST*, *SP100* and *UBE1L*, can now be used as epigenetic markers for pluripotent stem cells. Among these 8 genes, *SALL4* has been used as an expression marker, and is revealed for the first time as an epigenetic marker. These epigenetic changes during reprogramming can be detected by 3 different methods (Illumina assay, COBRA and bisulfite sequencing), and is evident, i.e. CpG sites are methylated or unmethylated in an all-or-none fashion. The identification of these novel epigenetic markers can be another tool for the validation of pluripotent stem cells that are iPSCs and ESCs.

The hypo-methylated stem cell-required DMRs may have an important role for reprogramming as do the stem cell-specific DMRs, because reprogramming is dependent on the type of parent cells. In fact, genes associated with the hypo-methylated stem cell-required DMRs include a large number of transcription factors that are involved in pluripotency. Establishment of the stem cell-required DMRs database in iPSCs derived from different types of parent cells can help to generate human iPSCs in a fast and easy manner. Hypo-methylated stem cell-specific regions have been reported to be abundant in CpG islands [28–30]. In this study, the hypo-methylated stem cell-specific DMRs were significantly biased towards CpG islands, whereas the hyper-methylated stem cell-specific DMRs were biased to non-CpG islands, suggesting that genes with CpG islands have a propensity to be demethylated during reprogramming towards pluripotent stem cells. The higher number of the hyper-methylated stem cell-specific DMRs in iPSCs indicates that the Yamanaka factors activate only limited numbers of stem cell-specific/associated genes through demethylation of the specific DMRs shown in this study on the genome in parallel with methylating most genes associated with tissue-specific function during reprogramming.

Multiple appearances/disappearances of aberrant hyper-methylation

Continuous passaging of iPSCs reduces differences among clones in gene expression profiles in mouse [15] and in human [31] cells. Here we detected multiple appearances and disappearances of aberrant hyper-methylation throughout iPSC reprogramming. Furthermore, human iPSCs were gradually reprogrammed through the "convergence" of periodic aberrant hyper-methylation upon continuous passaging (Figure 7). The term "convergence" is used here to mean that amplitude of aberrant hyper-methylation (or number of ES-iPS-DMRs) decreases. The decrease of aberrant methylation suggests that iPSCs lose the characteristics inherited from the parent cells and adapt to ESCs. This aberrant and stochastic hyper-methylation and their convergence may be a direct cause of the transgene-independent phases of iPS reprogramming [15]. Aberrant hyper-methylation, for which the mechanism remains unclear, can possibly be attributed, at least in part, to up-regulation of DNMT3B, a de novo methyltransferase, at the early stages of reprogramming.

Maintenance of an epigenetic memory of their parent cells at early passage of human iPSCs (Figure 4A) is consistent with recent reports involving mouse iPSCs [15–17]. However, most inherited sites from the parent cells in iPSCs were inconsistent among iPSC clones from the same parent cells on the genome, and these sites showed periodic aberrant hyper-methylation during passaging, as well as aberrant sites. Inherited methylation is non-synchronous and stochastic, much like aberrant methylation, rather than deterministic. The inherited sites thus comprise a portion of all aberrant methylation observed in the clones.

Mouse female iPSCs as well as mouse female ESCs carry two active X chromosomes [32], but inactivation of the X chromosome in human female ESCs is variable [22,33–35]. It has been reported recently that human female iPSCs show a variable state of X-inactivation as is seen in human female ESCs [22,36]. In this study, human iPSCs exhibited a dynamic epigenetic state on the X chromosome. The ES-iPS-DMRs on the X chromosome in XY-iPSCs were rare and the average number of ES-iPS-DMRs in XY-iPSCs was significantly lower than in XX-iPSCs, suggesting that iPSCs are prone to aberrant hyper-methylation on the inactive X chromosome. A recent report showed that X inactivation in human ESCs is sensitive to the level of oxygen through culture in vitro [35]. Therefore, analysis of aberrant methylation in iPSCs that are established and cultured in low oxygen condition would be help to understand physiological relevance of X inactivation and reprogramming.

Incomplete adaptation of iPSCs to ESCs

The number of passages for "convergence" of the aberrant hyper-methylation seems be dependent on parental cell types and their sex. Disappearance of iPSCs in culture within 10 passages is occasionally observed, regardless of the cell of origin. This instability may be due to an excess of aberrant hyper-methylation at early passages in addition to the "partial reprogramming" theory [15]. The late-passage iPSCs, like the early-passage iPSCs, retained the ability to differentiate into cell types found in all three germ layers. iPSCs showed reduced aberrant methylation during adaptation to ESCs; however, iPSCs retained approximately 100 aberrant sites on autosomes, implying that iPSCs do not become identical to ESCs, although they become very close. The remaining aberrant sites were inconsistent among iPSC clones

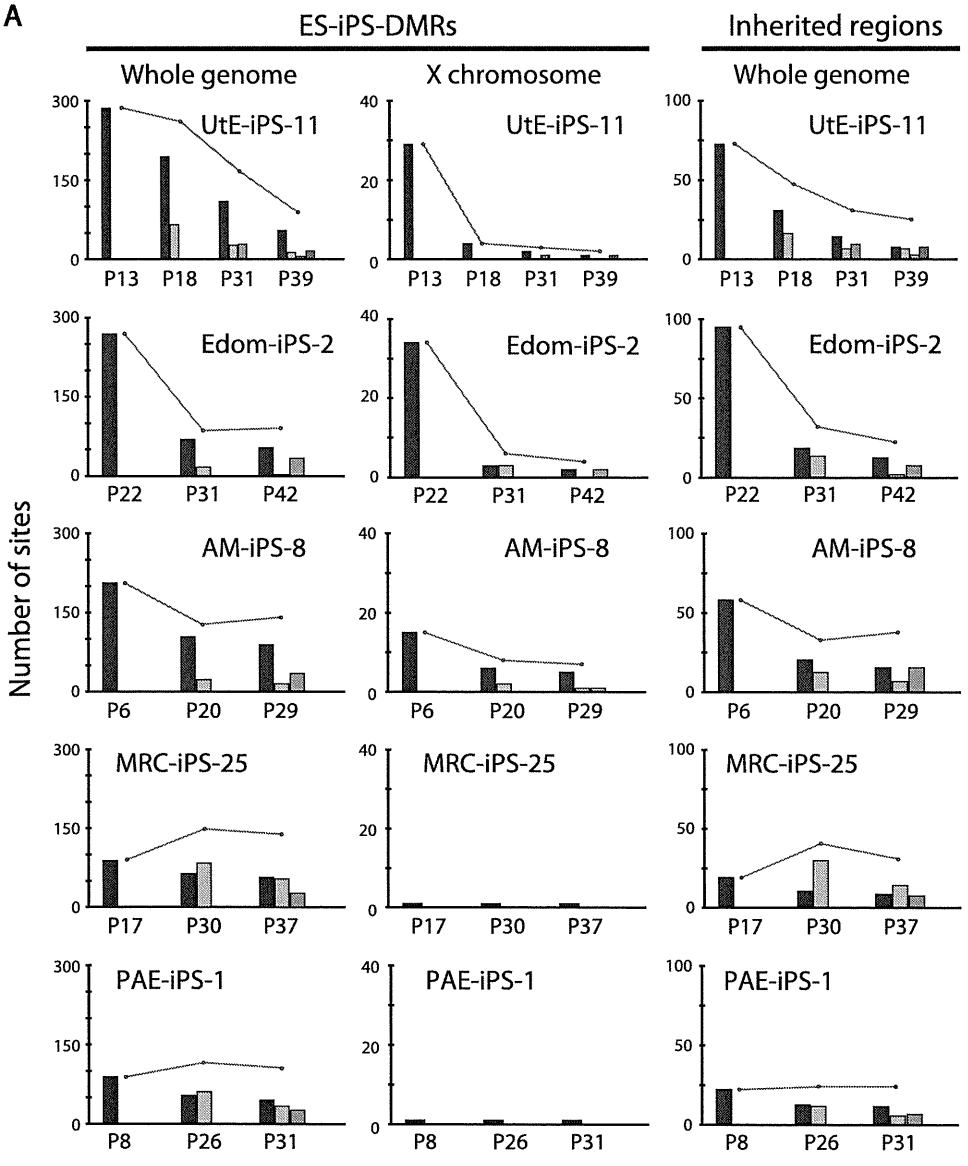

A
ES-iPS-DMRs
Inherited regions
Whole genome
X chromosome
Whole genome
Number of sites
UtE-iPS-11
Edom-iPS-2
AM-iPS-8
MRC-iPS-25
PAE-iPS-1
P13 P18 P31 P39
P22 P31 P42
P6 P20 P29
P17 P30 P37
P8 P26 P31
B
ES-Parent-DMRs
Whole genome
X chromosome
Number of sites
UtE1104
Edom22
P7 P11 P16
P5 P11 P16

Figure 5. Number of the ES-iPS-DMRs and ES-parent-DMRs with passaging. (*A*) Number of the ES-iPS-DMRs with passaging. Red line plots indicate total number of the ES-iPS-DMRs. Blue bars indicate the number of the ES-iPS-DMRs that appeared at the earliest passage. Orange, green and red bars indicate the number of the ES-iPS-DMRs that appeared secondarily at later passages. Appearance/disappearance of the ES-iPS-DMRs and inherited regions were repeated, but the number of newly-appeared ES-iPS-DMRs was decreased with passaging. (*B*) Number of the ES-parent-DMRs with passaging. Blue bars indicate the number of the ES-parent-DMRs at P5 (or P7). Orange and green bars indicate de novo ES-parent-DMRs at P11 and P16, respectively.
doi:10.1371/journal.pgen.1002085.g005

with different parent cell types, but the numbers were consistent among iPSC clones after a 42-week cultivation. The quantity (or number) of ES-iPS-DMRs would be another validation index for iPSC identity as well as quality analysis (or methylation ratio) of pluripotent stem cell-specific methylation.

Abnormalities of imprint genes, *MEG3* genes, and *H19* genes in human iPSCs

Genomic imprinting of *H19*, *IGF2* and *MEG3* has been reported to be unstable in human ESCs [37,38]. The *Dlk1-Dio3* genes were aberrantly silenced in most of the mouse iPSC lines. But mouse iPSCs without *MEG3* expression still have the ability to differentiate into cell type of three germ layers *in vitro* [39]. In humans, IG-DMR and MEG3-DMR are relevant to upd(14)pat-like and upd(14)mat-like phenotypes [40]. In this study, only *MEG3* and *H19*, out of 87 imprinted genes examined showed aberrant methylation in human iPSCs (Figure S10). Six out of 15 human iPSC lines were aberrantly methylated at MEG3-DMR. *MEG3* expression was silenced in those six lines regardless of their parent cell type, although all parent cells showed about 50% methylation at MEG3-DMR and expression of *MEG3* (Figure S10A, S10B). However, MEG3-negative iPS lines are almost indistinguishable from MEG3-positive iPS lines in DNA methylation and gene expression in human. Continuous passaging

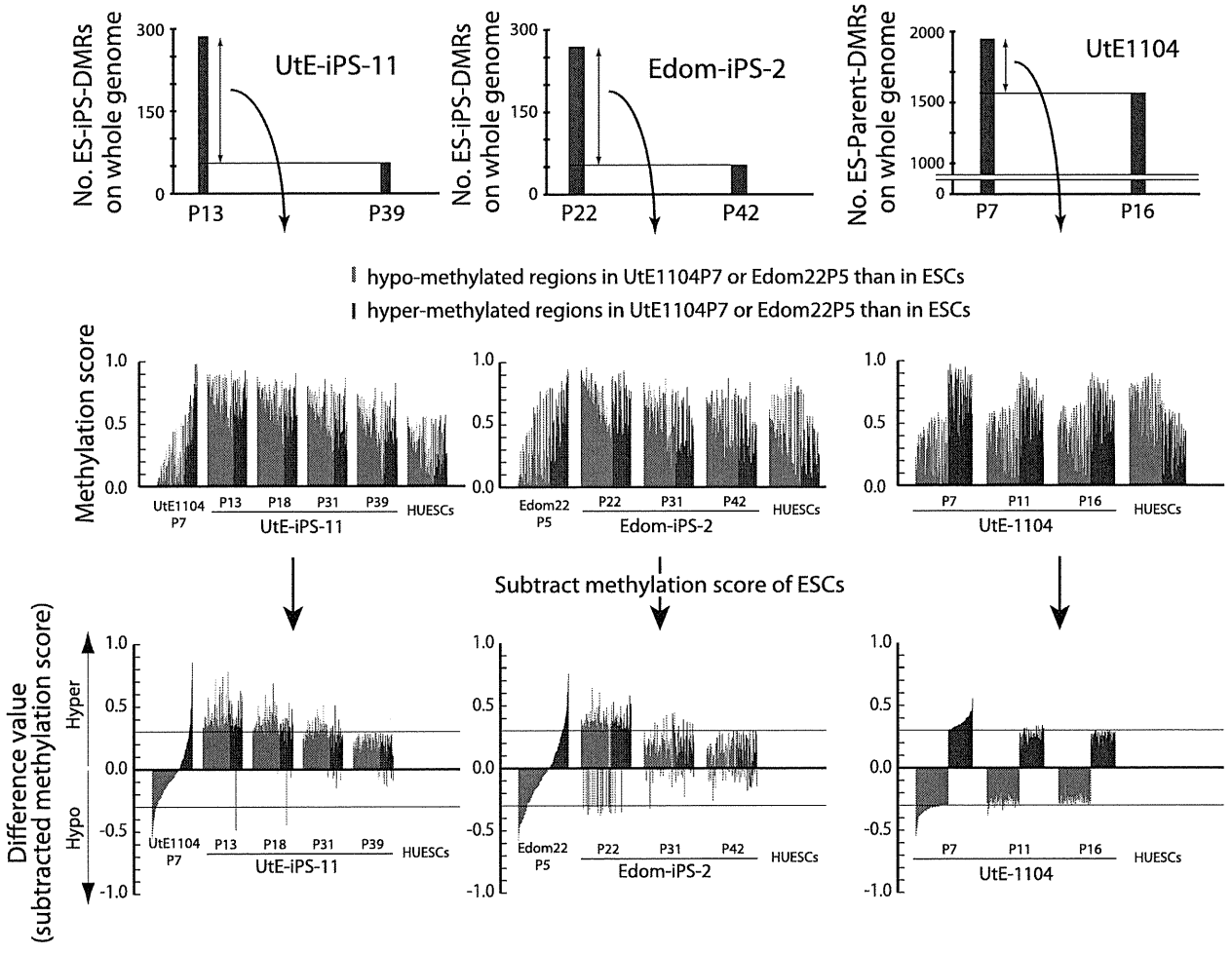


Figure 6. Hyper-methylation in the ES-iPS-DMRs and ES-parent-DMRs. ES-iPS-DMRs that disappeared in UtE-iPS-11 and Edom-iPS-2 at the latest passage (upper) were analyzed and the methylation score of each ES-iPS-DMR was plotted on bar graph (middle). To clearly compare methylation scores, difference value were estimated by subtracting the scores of ESCs from that of each sample (lower). Red and blue bars represent hypo- and hyper-methylated regions, respectively, in the parent cells, compared with ESCs. Notably, almost all the regions, even though their difference values were hypo-methylated in the parent cells, became hyper-methylated in iPSCs at the early passage, and then their methylation levels were adjusted to the level of ESCs with passaging, i.e. subtracted methylation score became close to zero. This transiently-induced hyper-methylation was not detected in parent cells.
doi:10.1371/journal.pgen.1002085.g006

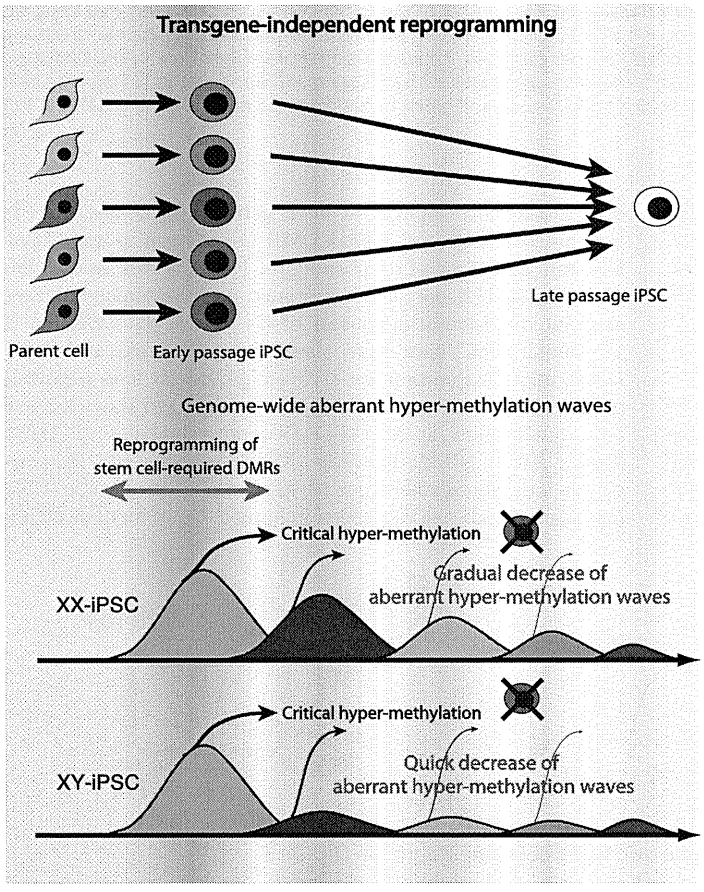


Figure 7. Model of mechanism for transgene-independent reprogramming. During reprogramming from somatic cells to iPSCs, the cells undergo dynamic change of methylation of SS-DMRs and genome. The cells with incomplete reprogramming or excessive hyper-methylation of the genome fail to maintain pluripotency at early passages. Human iPSCs are transgene-independently reprogrammed gradually through "convergence" of periodic aberrant hyper-methylation and become closer to ESCs upon continuous passaging. Due to the sensitivity to aberrant methylation on X chromosome, XY-iPSCs become close to ESCs faster than XX-iPSCs do.
doi:10.1371/journal.pgen.1002085.g007

did not resolve the aberrant hyper-methylation at MEG3-DMR, suggesting that these abnormalities occur at early passage and are fixed at later stages. In addition, aberrant hyper-methylation at *H19* in all iPSCs and ESCs was observed (Figure S10C), and *H19* was not expressed in all iPSCs and their parent cells.

We revealed that transgene-independent reprogramming is a convergence of periodic hyper-methylation. The aberrant hyper-methylation in iPSCs occurs stochastically throughout the genome. Early-stage iPSC clones with different propensities due to stochastic hyper-methylation may be used after selection of desirable phenotypes to treat a wide range of target diseases using cell-based therapy, and would thus have advantages for clinical use. In this sense, the number of ES-iPS-DMRs and methylation states of the stem cell-specific DMRs are useful epigenetic indices for evaluating human iPSCs in therapeutic applications.

Materials and Methods

Ethics statement

Human endometrium, amnion, placental artery endothelium and menstrual blood cells were collected by scraping tissues from surgical specimens, under signed informed consent, with ethical approval of the Institutional Review Board of the National Institute for Child Health and Development, Japan. Signed informed consent was obtained from donors, and the surgical specimens were irreversibly de-identified. All experiments handling human cells and tissues were performed in line with Tenets of the Declaration of Helsinki.

Human cell culture

Endometrium (UtE1104), amnion (AM936EP), placental artery endothelium (PAE551) and menstrual blood cell (Edom22) cell lines were independently established in our laboratory [41,42]. UtE1104, AM936EP, Edom22, and MRC-5 [43] cells were maintained in the POWEREDBY10 medium (MED SHIROTORI CO., Ltd, Tokyo, Japan). PAE551 cells were cultured in EGM-2MV BulletKit (Lonza, Walkersville, MD, USA) containing 5% FBS. Human iPSCs were generated in our laboratory, via procedures described by Yamanaka and colleagues [8] with slight modification [17,41,44–46]. The human cells were infected with retroviruses produced from the retroviral vector pMXs, which encodes the cDNA for human *OCT3/4*, *SOX2*, *c-MYC*, and *KLF4*. Human iPSCs were established from MRC-5, AM936EP, UtE1104, and PAE551, which were designated as MRC-iPSCs, AM-iPSCs, UtE-iPSCs and PAE-iPSCs [17,41,44–46]. Edom-iPSCs were established from Edom22 in this study. Human iPSCs were maintained on irradiated MEFs in 0222 medium (MED SHIROTORI CO., Ltd, Tokyo, Japan) supplemented with

10 ng/ml recombinant human basic fibroblast growth factor (bFGF, Wako Pure Chemical Industries, Ltd., Osaka, Japan). The 201B7 human iPSC line [8] that was generated from human skin fibroblasts by retroviral transfection with 4 transcription factors was also used. Frozen pellets of human ESCs (HUESCs) [23,24] were kindly gifted from Drs. C. Cowan and T. Tenzan (Harvard Stem Cell Institute, Harvard University, Cambridge, MA).

DNA methylation analysis

DNA methylation analysis was performed using the Illumina infinium assay with the HumanMethylation27 BeadChip (Illumina) and the BeadChip was scanned on a BeadArray Reader (Illumina), according to the manufacturer's instructions. Methylated and unmethylated signals were used to compute a β-value, which was a quantitative score of DNA methylation levels, ranging from "0", for completely unmethylated, to "1", for completely methylated. On the HumanMethylation27 BeadChip, oligonucleotides for 27,578 CpG sites covering more than 14,000 genes are mounted, mostly selected from promoter regions. CpG sites with ≥0.05 "Detection p value" (computed from the background based on negative controls) were eliminated from the data for further analysis, leaving 24,273 CpGs (13,728 genes) valid for use with the 51 samples tested. Average of methylation was calculated from HUESCs, MRC-iPSCs, AM-iPSCs, UtE-iPSCs, PAE-iPSCs and Edom-iPSCs, in which DMRs among each line in the each set were removed. Analyzed data sets (list of stem cell-specific DMRs and stem cell-required DMRs) can be obtained from http://www.nch.go.jp/reproduction/e/thdmds.html.

Gene expression analysis

Gene expression analysis was performed using the Agilent Whole Human Genome Microarray chips G4112F (Agilent, Santa Clara, CA), which contains over 41,000 probes. Raw data were normalized and analyzed by GeneSpringGX11 software (Silicon Genetics, Redwood City, CA). For RT-PCR, an aliquot of total RNA was reverse-transcribed using Random Hexamer primers. The cDNA template was amplified using specific primers for *EPHA1*, *PTPN6*, *RAB25*, *SALL4*, *GBP3*, *LYST*, *SP100*, *UBE1L*, OCT3/4 and NANOG. For detecting RNA derived from transgenes, specific primer sets, FY-11 and OCT3/4-SR, FY-11 and SOX2-SR, KLF4-SF and FY-12, cMYC-SF and FY-12, were used. Expression of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a control. Primers used in this study are summarized in Table S8.

Quantitative combined bisulfite restriction analysis (COBRA) and bisulfite sequencing

To confirm the DNA methylation state, bisulfite PCR-mediated restriction mapping (known as the COBRA method) was performed. Sodium bisulfite treatment of genomic DNA was carried out using EZ DNA Methylation-Gold kit (Zymo Research). PCR amplification was performed using BIOTAQ HS DNA polymerase (Bioline Ltd; London, UK) with specific primers for *EPHA1*, *PTPN6*, *RAB25*, *SALL4*, *GBP3*, *LYST*, *SP100*, and *UBE1L*. Primers used in this study are summarized in Table S8. After digestion with restriction enzymes, HpyCH4IV or Taq I, quantitative-COBRA coupled with the Shimadzu MCE-202 MultiNA Microchip Electrophoresis System (Shimadzu, Japan) was carried out for quantitative DNA methylation level. To determine the methylation state of individual CpG sites, the PCR product was gel extracted and subcloned into pGEM T Easy vector (Promega, Madison, WI), and then sequenced. The promoter regions of the *OCT3/4* and *NANOG* [41,44] were also amplified and sequenced. Methylation sites were visualized and quality control was carried out by the web-based tool, "QUMA" (http://quma.cdb.riken.jp/) [47].

Web tools

The following web tools were used in this study: NIA Array [48] (http://lgsun.grc.nia.nih.gov/ANOVA/) for hierarchical clustering, DAVID Bioinformatics Resources [49] (http://david.abcc.ncifcrf.gov/home.jsp), PANTHER Classification System [50] (http://www.pantherdb.org/).

Accession numbers

NCBI GEO: HumanMethylation27 BeadChip data and gene expression microarray data have been submitted under accession number GSE 20750, GSE24676 and GSE24677.

Supporting Information

Figure S1 Immunohistochemistry of stem cell-specific surface antigens, NANOG, OCT3/4, SOX2, SSEA-4 and TRA-1-60 in AM-iPSCs, MRC-iPSCs and Edom-iPSCs, and teratoma formation of those iPSCs by subcutaneous implantation into NOD/Scid mice. The iPSCs differentiated to various tissues including ectoderm (neural tissues and retinal pigment epithelium), mesoderm (cartilage) and endoderm (gut). Immunostaining and teratoma formation were carried out as previously described [41,44].
(PDF)

Figure S2 Immunohistochemistry of stem cell-specific surface antigens, NANOG, OCT3/4, SOX2, SSEA-4 and TRA-1-60 in PAE-iPSCs and UtE-iPSCs, and teratoma formation of those iPSCs by subcutaneous implantation into NOD/Scid mice. The iPSCs differentiated to various tissues including ectoderm (neural tissues and retinal pigment epithelium), mesoderm (cartilage) and endoderm (gut). Immunostaining and teratoma formation were carried out as previously described [41,44].
(PDF)

Figure S3 Bisulfite sequencing at the OCT3/4 and NANOG promoter regions in ESCs, iPSCs and their parent cells.
(PDF)

Figure S4 Expression of the transgenes in iPSCs. (A) RT-PCR for transgenes in 22 iPSC lines. No expression of the transgenes in each iPSC lines was detected. (B) Quantitative RT-PCR for the transgenes at each passage. Relative expression of each transgene normalized to GAPDH was calculated. P0(D2), RNA from UtE1104 cells that were infected with the retroviruses and were cultured for 2 days. No expression of the transgenes at each passage was detected.
(PDF)

Figure S5 (A) Unsupervised hierarchical clustering analysis based on DNA methylation (left) and gene expression (right) in each ESC line, iPSC line and their parent cell line. (B) Unsupervised hierarchical clustering analysis based on DNA methylation (left) and gene expression (right) of average of ESCs, iPSCs and parent cells. (C) Scatter plot of DNA methylation (left) and gene expression data (right) in ESCs, iPSCs and their parent cells.
(PDF)

Figure S6 (A) Venn-like diagram showing seven categories (aa-gg) overlapped CpG sites among ESCs, iPSCs and their parent cells. (B) Number of CpG sites involved in each seven category from the five ESCs-iPSCs-the parent cell sets. "Overlapped"

indicates a number of sites that overlap in all iPSCs examined. The 220 overlapping sites in "ee" are designated as stem cell-specific differentially methylated regions (DMRs) and 3,123 total sites in "ee" are designated as stem cell-required DMRs. Notably, no overlapping sites were observed in "bb" that is a category involved in iPSCs-specific DMRs and in "ff" that is a category involved in inherited regions in iPSCs from the parent cells.
(PDF)

Figure S7 (A) Distribution of stem cell-required DMRs on each chromosome (upper) and frequency on each chromosome (bottom). (B) The number of parent cell specific DMRs (left) and the number of iPSC derived from different parent cells specific DMRs (left).
(PDF)

Figure S8 The number of DMRs between ESCs and each iPSC line (ES-iPS-DMRs) on each chromosome. ES-iPS-DMRs between 201B7 (iPSCs from Yamanaka) and ESCs are shown for comparison.
(PDF)

Figure S9 Distribution of the ES-iPS-DMRs on each chromosome. Distribution of the EiP-DMRs overlapped in less than 9 lines (light blue bars), in more than 10 and less than 14 lines (blue bars), and in more than 15 lines (red bars) among 22 lines.
(PDF)

Figure S10 DNA methylation at human *MEG3* and *H19*. (A) DNA methylation at MEG3-DMR (CG7) and expression of MEG3. (Top) Schematic diagram of the MEG3 gene. The arrow, open boxes and open circles represent transcription start site, first exon and position of CpG sites, respectively. Red and blue arrowheads represent the position of CpG sites in Infinium assay and COBRA assay, respectively. DNA methylation scores of MEG3 were determined by Illumina Infinium HumanMethylation27 assay (upper bar graph) and Bio-COBRA (lower bar graph). (Bottom) Expression of MEG3 and GAPDH was determined by RT-PCR. Information of MEG3 primers for COBRA and RT-PCR is described by Kagami et al. [40]. (B) Bisulfite sequencing analysis of MEG3-DMRs (CG7). (C) Methylation scores of H19 were determined by Illumina Infinium HumanMethylation27 assay.
(PDF)

Table S1 List of human cells analyzed for a methylation state in this study.
(PDF)

Table S2 STR analysis of iPSCs.
(PDF)

Table S3 Karyotypic analysis of iPSCs.
(PDF)

Table S4 List of genes with stem cell-specific DMRs exhibiting significant changes in expression in human iPS cells.
(PDF)

Table S5 List of the top 100 genes with hypo-methylated stem cell-required DMRs exhibiting 'high' expression in human iPS cells.
(PDF)

Table S6 List of top 100 genes with hyper-methylated stem cell-required DMRs exhibiting suppression in human iPS cells.
(PDF)

Table S7 List of top 5 categories of GO Term in "Stem cell-required DMRs".
(PDF)

Table S8 Primer list.
(PDF)

Acknowledgments

We would like to express our sincere thanks to Drs. C. Cowan and T. Tenzan for HUESC lines; to Drs. K. Hata and K. Nakabayashi for COBRA; to Dr. H. Makino for establishing the AM936EP, UtE1104, PAE551, and Edom22 cells; to Mr. M. Machida for immunohistochemical analysis; to Ms. Y. Takahashi for bioinformatics analyses; and Dr C. Ketcham for critical proofreading.

Author Contributions

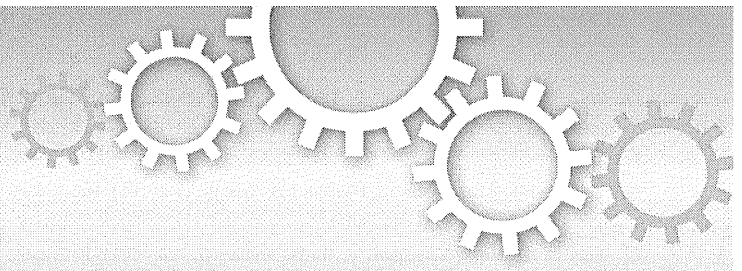
Conceived and designed the experiments: KN AU. Performed the experiments: KN MT MY-I. Analyzed the data: KN. Contributed reagents/materials/analysis tools: KN MT MY-I YF EC HS HA. Wrote the paper: KN AU.

References

1. Li E (2002) Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3: 662–673.
2. Reik W (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447: 425–432.
3. Feng S, Jacobsen SE, Reik W (2010) Epigenetic reprogramming in plant and animal development. Science 330: 622–627.
4. Hattori N, Nishino K, Ko YG, Ohgane J, Tanaka S, et al. (2004) Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. J Biol Chem 279: 17063–17069.
5. Hattori N, Imao Y, Nishino K, Ohgane J, Yagi S, et al. (2007) Epigenetic regulation of Nanog gene in embryonic stem and trophoblast stem cells. Genes Cells 12: 387–396.
6. Nishino K, Hattori N, Tanaka S, Shiota K (2004) DNA methylation-mediated control of Sry gene expression in mouse gonadal development. J Biol Chem 279: 22306–22313.
7. Zingg JM, Pedraza-Alva G, Jost JP (1994) MyoD1 promoter autoregulation is mediated by two proximal E-boxes. Nucleic Acids Res 22: 2234–2241.
8. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131: 861–872.
9. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. Science 318: 1917–1920.
10. Park IH, Zhao R, West JA, Yabuuchi A, Huo H, et al. (2008) Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451: 141–146.
11. Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, et al. (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 458: 766–770.
12. Park IH, Arora N, Huo H, Maherali N, Ahfeldt T, et al. (2008) Disease-specific induced pluripotent stem cells. Cell 134: 877–886.
13. Miura K, Okada Y, Aoi T, Okada A, Takahashi K, et al. (2009) Variation in the safety of induced pluripotent stem cell lines. Nat Biotechnol 27: 743–745.
14. Ghosh Z, Wilson KD, Wu Y, Hu S, Quertermous T, et al. (2010) Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. PLoS ONE 5: e8975. doi:10.1371/journal.pone.0008975.
15. Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, et al. (2010) Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat Biotechnol 28: 848–855.
16. Kim K, Doi A, Wen B, Ng K, Zhao R, et al. (2010) Epigenetic memory in induced pluripotent stem cells. Nature.
17. Nishino K, Toyoda M, Yamazaki-Inoue M, Makino H, Fukawatase Y, et al. (2010) Defining Hypo-Methylated Regions of Stem Cell-Specific Promoters in Human iPS Cells Derived from Extra-Embryonic Amnions and Lung Fibroblasts. PLoS ONE 5: e13017. doi:10.1371/journal.pone.0013017.
18. Fazzari MJ, Greally JM (2004) Epigenomics: beyond CpG islands. Nat Rev Genet 5: 446–455.
19. Fazzari MJ, Greally JM (2010) Introduction to epigenomics and epigenome-wide analysis. Methods Mol Biol 620: 243–265.

20. Doi A, Park IH, Wen B, Murakami P, Aryee MJ, 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.
21. Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, et al. (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471: 68–73.
22. Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, et al. (2011) Reference Maps of Human ES and iPS Cell Variation Enable High-Throughput Characterization of Pluripotent Cell Lines. Cell 144: 439–452.
23. Cowan CA, Klimanskaya I, McMahon J, Atienza J, Witmyer J, et al. (2004) Derivation of embryonic stem-cell lines from human blastocysts. N Engl J Med 350: 1353–1356.
24. Osafune K, Caron L, Borowiak M, Martinez RJ, Fitz-Gerald CS, et al. (2008) Marked differences in differentiation propensity among human embryonic stem cell lines. Nat Biotechnol 26: 313–315.
25. Brena RM, Auer H, Kornacker K, Plass C (2006) Quantification of DNA methylation in electrofluidics chips (Bio-COBRA). Nat Protoc 1: 52–58.
26. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126: 663–676.
27. Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, et al. (2008) Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 26: 1269–1275.
28. Fouse SD, Shen Y, Pellegrini M, Cole S, Meissner A, et al. (2008) Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. Cell Stem Cell 2: 160–169.
29. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, et al. (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454: 766–770.
30. Sato S, Yagi S, Arai Y, Hirabayashi K, Hattori N, et al. (2010) Genome-wide DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) residing in mouse pluripotent stem cells. Genes Cells 15: 607–618.
31. 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.
32. Maherali N, Sridharan R, Xie W, Utikal J, Eminli S, et al. (2007) Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell 1: 55–70.
33. Hall LL, Byron M, Butler J, Becker KA, Nelson A, et al. (2008) X-inactivation reveals epigenetic anomalies in most hESC but identifies sublines that initiate as expected. J Cell Physiol 216: 445–452.
34. Shen Y, Matsuno Y, Fouse SD, Rao N, Root S, et al. (2008) X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. Proc Natl Acad Sci U S A 105: 4709–4714.
35. Lengner CJ, Gimelbrant AA, Erwin JA, Cheng AW, Guenther MG, et al. (2010) Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell 141: 872–883.
36. Tchieu J, Kuoy E, Chin MH, Trinh H, Patterson M, et al. (2010) Female human iPSCs retain an inactive X chromosome. Cell Stem Cell 7: 329–342.
37. Rugg-Gunn PJ, Ferguson-Smith AC, Pedersen RA (2005) Epigenetic status of human embryonic stem cells. Nat Genet 37: 585–587.
38. Rugg-Gunn PJ, Ferguson-Smith AC, Pedersen RA (2007) Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines. Hum Mol Genet 16 Spec No. 2: R243–251.
39. Stadtfeld M, Apostolou E, Akutsu H, Fukuda A, Follett P, et al. (2010) Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. Nature 465: 175–181.
40. Kagami M, Sekita Y, Nishimura G, Irie M, Kato F, et al. (2008) Deletions and epimutations affecting the human 14q32.2 imprinted region in individuals with paternal and maternal upd(14)-like phenotypes. Nat Genet 40: 237–242.
41. Nagata S, Toyoda M, Yamaguchi S, Hirano K, Makino H, et al. (2009) Efficient reprogramming of human and mouse primary extra-embryonic cells to pluripotent stem cells. Genes Cells 14: 1395–1404.
42. Cui CH, Uyama T, Miyado K, Terai M, Kyo S, et al. (2007) Menstrual blood-derived cells confer human dystrophin expression in the murine model of Duchenne muscular dystrophy via cell fusion and myogenic transdifferentiation. Mol Biol Cell 18: 1586–1594.
43. Jacobs JP, Jones CM, Baille JP (1970) Characteristics of a human diploid cell designated MRC-5. Nature 227: 168–170.
44. Makino H, Toyoda M, Matsumoto K, Saito H, Nishino K, et al. (2009) Mesenchymal to embryonic incomplete transition of human cells by chimeric OCT4/3 (POU5F1) with physiological co-activator EWS. Exp Cell Res 315: 2727–2740.
45. Saito S, Onuma Y, Ito Y, Tateno H, Toyoda M, et al. (2010) Potential linkages between the inner and outer cellular states of human induced pluripotent stem cells. BMC Bioinformatics, in press.
46. Toyoda M, Yamazaki-Inoue M, Itakura Y, Kuno A, Ogawa T, et al. (2010) Lectin microarray analysis of pluripotent and multipotent stem cells. Genes Cells, in press.
47. Kumaki Y, Oda M, Okano M (2008) QUMA: quantification tool for methylation analysis. Nucleic Acids Res 36: W170–175.
48. Sharov AA, Dudekula DB, Ko MS (2005) A web-based tool for principal component and significance analysis of microarray data. Bioinformatics 21: 2548–2549.
49. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44–57.
50. Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, et al. (2005) The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res 33: D284–288.

SCIENTIFIC REPORTS



OPEN

SUBJECT AREAS:
BIOLOGICAL SCIENCES
CELL BIOLOGY
MEDICAL RESEARCH
DEVELOPMENTAL BIOLOGY

Received
8 April 2011

Accepted
4 August 2011

Published
19 August 2011

Correspondence and requests for materials should be addressed to K.M. (kmiyado@nch.go.jp)

* These authors equally contributed to this work

β-catenin is a molecular switch that regulates transition of cell-cell adhesion to fusion

Youki Takezawa[1]*, Keiichi Yoshida[1]*, Kenji Miyado[1], Masahiro Sato[2], Akihiro Nakamura[1], Natsuko Kawano[1], Keiichi Sakakibara[1], Takahiko Kondo[1], Yuichirou Harada[1], Naoko Ohnami[1], Seiya Kanai[1], Mami Miyado[1], Hidekazu Saito[3], Yuji Takahashi[3], Hidenori Akutsu[1] & Akihiro Umezawa[1]

[1]Department of Reproductive Biology, National Center for Child Health and Development, 2-10-1 Okura, Setagaya, Tokyo 157-8535, Japan, [2]Section of Gene Expression Regulation, Frontier Science Research Center, Kagoshima University, 1-21-20 Korimoto, Kagoshima, Kagoshima 890-0065, Japan, [3]Division of Reproductive Medicine, Department of Perinatal Medicine and Maternal Care, National Center for Child Health and Development, 2-10-1 Okura, Setagaya, Tokyo 157-8535, Japan.

When a sperm and an oocyte unite upon fertilization, their cell membranes adhere and fuse, but little is known about the factors regulating sperm-oocyte adhesion. Here we explored the role of β-catenin in sperm-oocyte adhesion. Biochemical analysis revealed that E-cadherin and β-catenin formed a complex in oocytes and also in sperm. Sperm-oocyte adhesion was impaired when *β-catenin*-deficient oocytes were inseminated with sperm. Furthermore, expression of β-catenin decreased from the sperm head and the site of an oocyte to which a sperm adheres after completion of sperm-oocyte adhesion. UBE1-41, an inhibitor of ubiquitin-activating enzyme 1, inhibited the degradation of β-catenin, and reduced the fusing ability of wild-type (but not *β-catenin*-deficient) oocytes. These results indicate that β-catenin is not only involved in membrane adhesion, but also in the transition to membrane fusion upon fertilization.

An oocyte fuses to only one sperm at fertilization, which results in the creation of a single cell with two nuclei that undergoes a series of complex processes (Supplementary Fig. S1a)[1]. After the sperm detaches the cumulus cells, the somatic cells surrounding oocytes, from the oocytes by enzymatic activities, the sperm adheres to the zona pellucida (ZP), the oocyte extracellular matrix. The sperm then penetrates the ZP and adheres to the oocyte cell membrane. At this time, fusion occurs between sperm and oocyte.

CD9[2] and Izumo1[3] belong to the tetraspan protein family (tetraspanin) and immunoglobulin superfamily, respectively, and play a crucial role in sperm-oocyte fusion[3–5]. Both *CD9*-deficient oocytes and *Izumo1*-deficient sperm are unable to fuse to their wild-type partner's cells, but retain adhesive activity[3,4]. A couple of these findings suggest that the molecular event underlying membrane adhesion is different from that underlying membrane fusion. The mechanism of membrane fusion has been explored by us[6] and others[7,8], but little is known about the factors regulating the adhesion of a sperm to an oocyte membrane.

Three proteins, β-catenin, α-catenin and E-cadherin, are a well-known functional set mediating intercellular junctions, which are called adherens junctions and typically served as a lateral connector between epithelial cells[9]. Besides epithelial cells, these three proteins are co-expressed in non-epithelial female germ cells, such as immature oocytes and fully-grown oocytes (to which only one sperm can adhere upon fertilization)[10]. Typically, β-catenin directly binds to the cytoplasmic domain of E-cadherin and connects to the adherens junctional complex with actin, a major component of microfilaments[11]. The β-catenin bound to E-cadherin is involved in intercellular adhesion, while E-cadherin-free β-catenin functions as a transcriptional factor driving the Wnt signaling pathway that regulates embryonic morphogenesis[12].

In mouse oocytes, the presence of both β-catenin and E-cadherin has been reported[10], but it remains unclear whether these two proteins are essential during the process of fertilization. In this study, we explored the possible role of β-catenin in sperm-oocyte adhesion, one of the important steps leading to mammalian fertilization.

Results

Subcellular localization of actin and its possible function in sperm-oocyte adhesion. To identify candidate proteins involved in sperm-oocyte membrane adhesion, we first examined immunocytochemically whether two

cytoskeletal proteins, actin and tubulin, would exhibit cortical localization in ovulated oocytes. These two monoclonal antibodies (mAbs) are raised against the conserved domain of actin isoforms or β-tubulin. Confocal microscopic observation demonstrated that actin was asymmetrically localized in the oocyte: namely, metaphase II-arrested chromosomes enclosed by the actin were observed in one side of the oocyte cytoplasm (Supplementary Fig. S1b, c). In addition, actin was found to be concentrated on the cortical surface of the oocyte and appeared to exist as orderly arranged spots beneath the oocyte cell membrane (Supplementary Fig. S1d). Since tetraspanin CD9 is known to be present on the microvilli that regularly line the cell surface of an oocyte[8], oocytes were subjected to double staining with CD9 and actin mAbs. Neither protein was co-localized, and the actin-rich cortical area was clearly separated from the CD9-rich area (Supplementary Fig. S1d, e), implying the presence of at least two types of membranous structures in the oocyte cell membrane, as suggested previously[13]. Since CD9 plays an important role in sperm-oocyte fusion, but not sperm-oocyte adhesion[4], it was hypothesized that the actin-rich membranous structure on the cell surface of an oocyte may be involved in sperm-oocyte membrane 'adhesion'.

E-cadherin/β-catenin complex formation in both oocytes and sperm. Since E-cadherin/β-catenin complex has been known to bind to actin, by which cell-cell membrane adhesion is regulated[9], we considered that this E-cadherin/β-catenin complex may play a role as a regulator of sperm-oocyte membrane adhesion. To assess the problem, we first examined the possible interaction between E-cadherin/β-catenin complex and actin on an oocyte using immunocytochemical methods. In Fig. 1, subcellular localization of α- and β-catenins and E-cadherin is shown and data on the

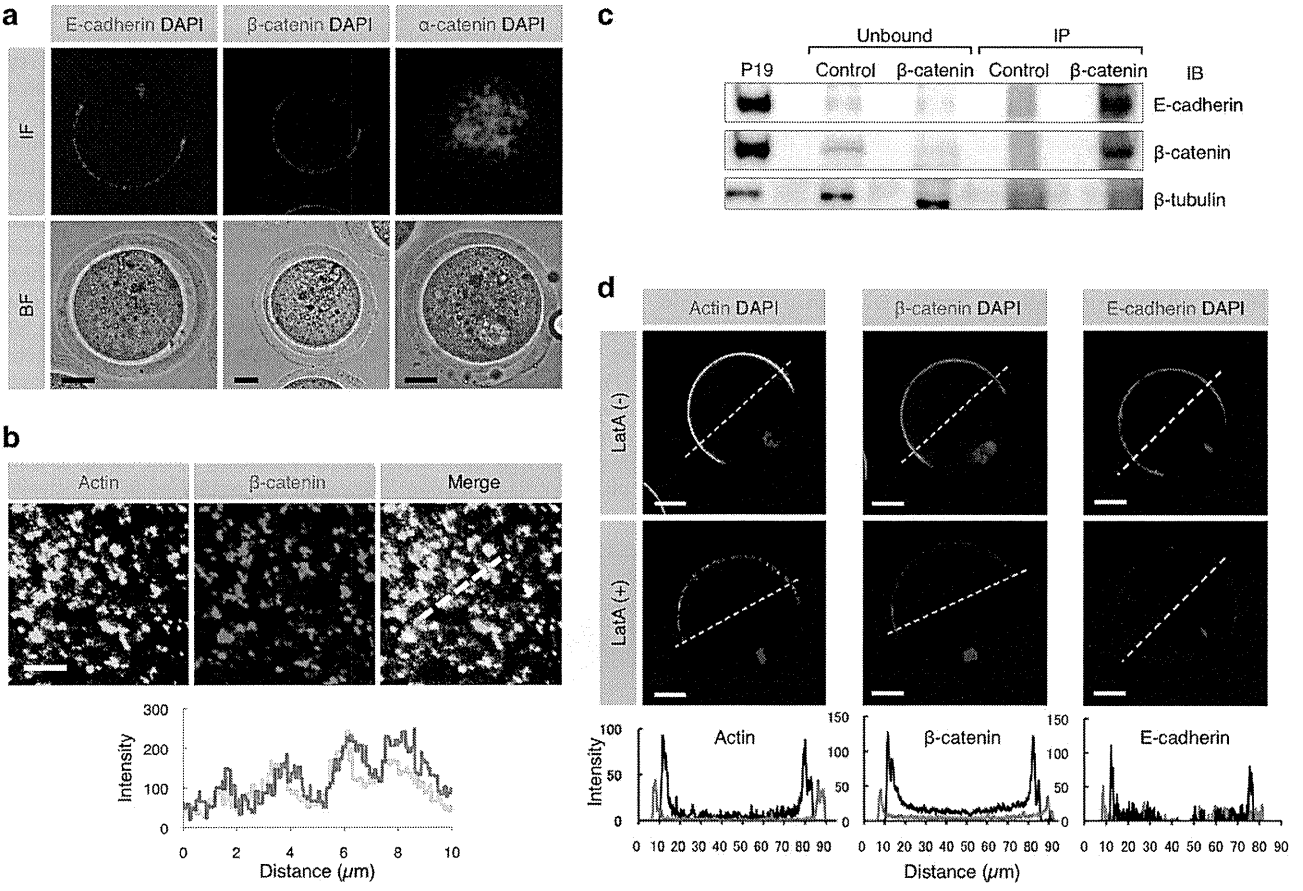


Figure 1 | Expression of E-cadherin and β-catenin and localization of E-cadherin/β-catenin complex in oocytes. (a) Localization of E-cadherin, β-catenin and α-catenin in ovulated oocytes. Similar distribution pattern of E-cadherin and β-catenin on an oocyte suggests complex formation between these two proteins. IF, immunofluorescence; BF, bright field. Scale bars: 20 μm. (b) Localization of β- and γ-actin isoforms and β-catenin beneath the oocyte cell membrane and their fluorescent intensities. The route scanned on the membrane was indicated as a dotted line. Fluorescence intensities for each protein were measured and graphed based on the 3D image, as described in the Experimental Procedures. Red and green lines in the lower panel indicate intensities of β-catenin and actin, respectively. Scale bar: 5 μm. (c) Biochemical evidence for the presence of E-cadherin/β-catenin complex in oocytes. The extract from 905 oocytes immunoprecipitated (IP) by anti-β-catenin mAb and mouse IgG purified from preimmune serum (Control) was subjected to immunoblotting with anti-E-cadherin, anti-β-catenin or anti-β-tubulin mAb. Extracts from mouse embryonic carcinoma cell line P19[35] were also subjected to immunoprecipitation with anti-β-catenin mAb and the resulting immunoprecipitates were reacted with each mAb as a positive control. Note that the extract (IP) immunoprecipitated by anti-anti-β-catenin mAb was reactive with both anti-β-catenin and anti-E-cadherin mAbs, but the extract (Unbound) that was not immunoprecipitated by anti-β-catenin mAb failed to bind to both antibodies. On the other hand, the β-tubulin was detectable in the Ab-unbound (but not Ab-IP) fractions. (d) Disassembly of β-catenin, E-cadherin and actin induced by latrunculin A (latA) treatment. Oocytes were doubly immunostained with anti-β- and γ-actin isoforms mAbs and DAPI (shown as 'Actin DAPI') or with anti-β-catenin mAb or with anti-E-cadherin mAb and DAPI (shown as 'β-catenin DAPI' or 'E-cadherin DAPI'). In the lower panels, the fluorescence intensities measured after being traced along dotted lines in the figures of the upper panels are shown. The intensities of latA-treated oocytes are indicated by red lines, while those of the latA-untreated oocytes are shown by black lines. Scale bar: 20 μm.

comparison between their localization and the localization of actin is also shown. Regarding the immunoreactivity of E-cadherin to an oocyte, a mAb that recognizes an N-terminal extracellular region of E-cadherin was used. Immunocytochemical staining demonstrated that E-cadherin was localized on the cell membrane (microvillar region) of an oocyte that has not been treated with permeabilization (Fig. 1a). β-catenin was detected beneath the oocyte cell membrane, and its localization pattern appeared to be similar to that of E-cadherin (Fig. 1a). In contrast, α-catenin was present in the oocyte cytoplasm (Fig. 1a). When the distribution of β-catenin on an oocyte was compared to that of actin, both proteins were found to be co-localized (Fig. 1b).

Secondly, we assessed the possible formation of β-catenin and E-cadherin complex using an immunoprecipitation method. A cell extract of mouse oocytes (n = 905) was immunoprecipitated with anti-E-cadherin mAb, and the resulting precipitate was then reacted with anti-β-catenin mAb. As a result, the cell extract immunoprecipitated with anti-E-cadherin mAb reacted with the anti-β-catenin mAb (Fig. 1c), indicating the presence of β-catenin and E-cadherin complex in an oocyte.

Thirdly, we assessed the effect of latrunculin A (latA), an inhibitor of actin polymerization, on the formation of β-catenin and E-cadherin complex. When oocytes were treated with 10 μM latA for 1 h at 37°C, actin immunoreactivity was reduced along with decreased immunoreactivity to β-catenin and E-cadherin (Fig. 1d). These results suggest that the β-catenin/E-cadherin complex formed in the oocyte cell membrane is closely associated with actin.

Since the β-catenin/E-cadherin complex is known to play a role in cell-cell adhesion via its homophilic interaction with E-cadherin[14], we predicted that this protein complex would also be produced in sperm. To test this hypothesis, epididymal capacitated sperm were collected from 10-week-old males and subjected to Western blotting (Fig. 2a) and immunoprecipitation (Fig. 2b) analyses. Western blotting revealed that both E-cadherin and β-catenin were detected in the sperm collected (Fig. 2a); however, N-cadherin was not detectable in those samples (Fig. 2a), although its expression has been reported in mouse oocytes[15]. Immunoprecipitation analysis also revealed the presence of the E-cadherin/β-catenin complex in sperm. The sperm extracts immunoprecipitated with anti-β-catenin mAb were reactive with anti-E-cadherin mAb, and those immunoprecipitated with anti-E-cadherin mAb were reactive with anti-β-catenin mAb (Fig. 2b). To confirm this further, immunocytochemical staining was performed for the isolated sperm. Staining of unpermeabilized sperm with anti-E-cadherin mAb demonstrated that E-cadherin was broadly expressed on the cell membrane of sperm from the head region to the mid-piece as well as part of the tail (Fig. 2c, d). Staining of permeabilized sperm with anti-β-catenin mAb revealed that the expression of β-catenin was localized beneath the sperm cell membrane and, notably, its localization pattern was similar to that of E-cadherin (Fig. 2c, d).

In mammals, both sperm-oocyte fusion and adhesion have been believed to occur in a specific region of the sperm head, called an equatorial segment (ES)[1]. Therefore, it is reasonable to consider that factor(s) regulating sperm-oocyte adhesion should exist in this segment. Since in the sperm head of the Asian musk shrew, *Suncus murinus*, ES is recessed within the waist of the sperm nucleus[16], it is easy to detect proteins localized in this segment. When immunocytochemical staining of the permeabilized shrew capacitated sperm was performed using anti-E-cadherin mAb, E-cadherin was expressed on the ES, the mid-piece and part of the tail (Supplementary Fig. S2a). Staining with anti-β-catenin mAb revealed the expression of β-catenin specifically localized on the ES of the sperm head (Supplementary Fig. S2b). Notably, its localization pattern was similar to that of E-cadherin in the shrew sperm (Supplementary Fig. S2a *vs.* Fig. S2b) and also to that of E-cadherin in the mouse sperm (Fig. 2c, d). Furthermore, some β-catenin molecules were released from the acrosomes of the shrew sperm (Supplementary Fig. S2b). Since the ES is recessed in the acrosome of the shrew sperm[16], β-catenin may have accumulated specifically in the ES upon completion of ES formation. These collected results led to a conclusion

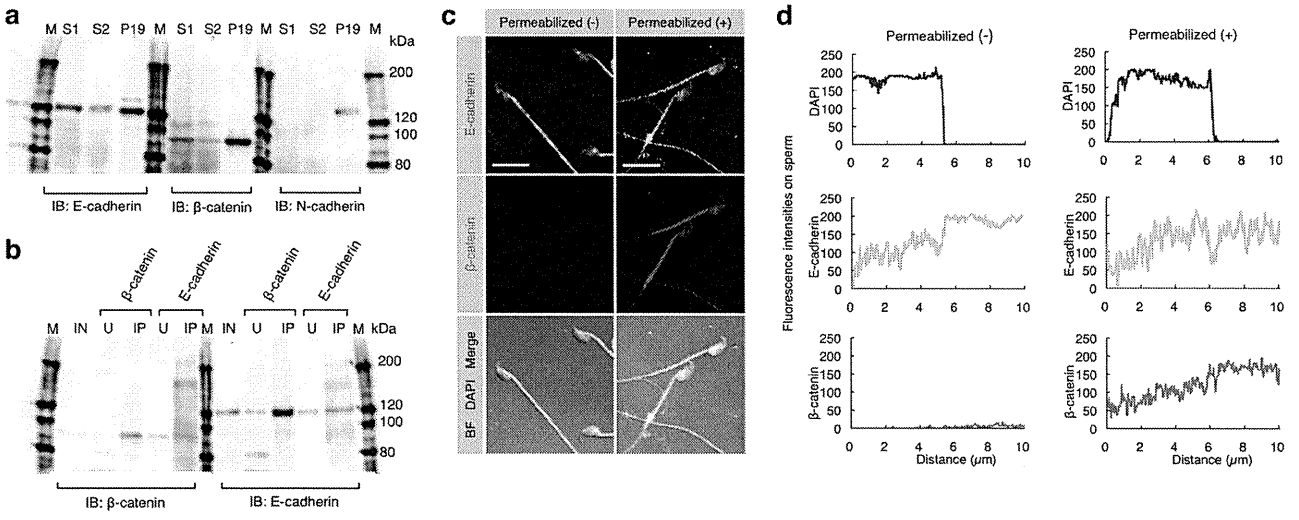


Figure 2 | Expression of E-cadherin and β-catenin and their interaction in sperm. (a) Expression of E-cadherin and β-catenin in epididymal sperm. E-cadherin and β-catenin, but not N-cadherin, in the sperm were detected by immunoblotting (IB). Sperm (S1 and S2) were collected from the epididymis of two males and used for IB. Extracts from P19 cells were used as a positive control. IB was performed using anti-E-cadherin, anti-β-catenin or anti-N-cadherin mAbs. M, molecular weight markers. (b) Interaction between E-cadherin and β-catenin in sperm. Extracts from the sperm (used as input sample (IN)) were immunoprecipitated by anti-E-cadherin or anti-β-catenin mAb. The precipitates (IP) and unbound extracts (U) were immunoblotted (IB) with anti-E-cadherin or anti-β-catenin mAb. M, molecular weight markers. (c) Localization of E-cadherin and β-catenin in sperm. Unpermeabilized or permeabilized sperm were doubly immunostained with anti-E-cadherin (ECCD-2) and anti-β-catenin mAbs, and their nuclei were stained with DAPI. ECCD-2, which recognizes an epitope in the N-terminal extracellular region of E-cadherin, bound to E-cadherin without permeabilization pretreatment. Scale bar: 5 μm. (d) The fluorescence intensity profiles of E-cadherin and β-catenin in sperm shown in (c). Fluorescence intensities were measured after being traced on the sperm along dotted lines shown at the bottom of the panels in (c).