RESOURCE

nature structural & molecular biology

Global analysis of parental imprinting in human parthenogenetic induced pluripotent stem cells

Yonatan Stelzer, Ofra Yanuka & Nissim Benvenisty

To study the role of parental imprinting in human embryogenesis, we generated parthenogenetic human induced pluripotent stem cells (iPSCs) with a homozygote diploid karyotype. Global gene expression and DNA methylation analyses of the parthenogenetic cells enabled the identification of the entire repertoire of paternally expressed genes. We thus demonstrated that the gene *U5D*, encoding a variant of the U5 small RNA component of the spliceosome, is an imprinted gene. Introduction of the *U5D* gene into parthenogenetic cells partially corrected their molecular phenotype. Our analysis also uncovered multiple miRNAs existing as imprinted clustered transcripts, whose putative targets we then studied further. Examination of the consequences of parthenogenesis on human development identified marked effects on the differentiation of extraembryonic trophectoderm and embryonic liver and muscle tissues. This analysis suggests that distinct regulatory imprinted small RNAs and their targets have substantial roles in human development.

Human iPSCs (HiPSCs) hold great promise for regenerative medicine and cell therapy and are a valuable tool for the study of genetic diseases[1–3]. Here we aimed to conduct a comprehensive study of human genomic parental imprinting using HiPSCs. Imprinted genes are expressed in a monoallelic fashion according to their parent-of-origin signature and are known for their roles in human development and disease. Aberrant expression of imprinted genes may lead to serious developmental defects. Most notable are Prader-Willi and Angelman syndromes, which are two distinct disorders that develop from either maternal or paternal uniparental disomy of the same chromosomal region. By reconstituting mouse embryos comprising two paternal chromosomal sets (androgenotes) or two maternal chromosomal sets (parthenogenotes or gynogenotes), it was initially shown that maternal and paternal chromosomes provide certain distinct contributions and that both are needed for correct embryonic development[4,5]. Therefore, a parthenogenetic cell, which lacks the paternal genome, is considered a valuable system for studying the influence of paternal imprinting on early embryogenesis. Recently, several independent studies have demonstrated the generation of human parthenogenetic embryonic stem cell (PgHESC) lines[6–8]. Nevertheless, to date no comprehensive study of human genomic imprinting has been conducted using such cell lines. Moreover, in contrast to standard human embryonic stem cells (HESCs), which are derived from surplus blastocysts, the generation of PgHESCs requires the use of oocytes and is thus ethically forbidden in many countries. It appears that cellular reprogramming of naturally occurring parthenogenetic benign tumors may provide a reasonable alternative. To assess this possibility, here we have generated parthenogenetic induced pluripotent stem cells and analyzed their global gene expression and differentiation.

RESULTS

Human parthenogenetic induced pluripotent stem cells

Human parthenogenetic embryos may develop into benign ovarian teratomas. To generate parthenogenetic HiPSC (PgHiPSC) lines, diploid parthenogenetic teratoma cell lines derived from two female patients were infected with pseudoretroviruses expressing the four reprogramming factors *OCT4* (also called *POU5F1*), *SOX2*, *KLF4* and *c-MYC* (*MYC*)[1]. Cells with a typical embryonic stem cell–like morphology appeared within 2–3 weeks after infection (**Supplementary Fig. 1a**) with an efficiency of ~0.01% (approximately 100 colonies out of 10^6 fibroblast cells), similar to the efficiency of generating wild-type HiPSC (WT-HiPSC). Three PgHiPSC lines were generated from each of the ovarian teratoma cell lines and were named PgHiPSC-A-11, PgHiPSC-A-20 and PgHiPSC-A-26, and PgHiPSC-B-34, PgHiPSC-B-36 and PgHiPSC-B-41, respectively. These cell lines express pluripotency markers such as alkaline phosphatase, OCT4 and Tra-1-60 (**Supplementary Fig. 1b**,**c** and **Fig. 1a**, respectively). Moreover, by analyzing the global gene expression profile of the cells, we found that they lost the expression of the differentiation-related markers and acquired expression of pluripotency-associated genes, giving them a gene expression profile closely resembling that of pluripotent cells (**Figs. 1** and **2a**,**b** and **Supplementary Fig. 1d**). Pluripotency is evaluated on the basis of competence to differentiate into the three embryonic germ layers. We therefore performed *in vitro* and *in vivo* differentiation assays by generating embryoid bodies and teratomas, respectively. Histology together with gene expression analysis showed differentiation of the cells into the three embryonic germ layers (**Supplementary Fig. 2a**,**b**).

Stem Cell Unit, Department of Genetics, Silberman Institute of Life Sciences, The Hebrew University, Jerusalem, Israel. Correspondence should be addressed to N.B. (nissimb@cc.huji.ac.il).

Received 22 December 2010; accepted 24 February 2011; published online 15 May 2011; doi:10.1038/nsmb.2050

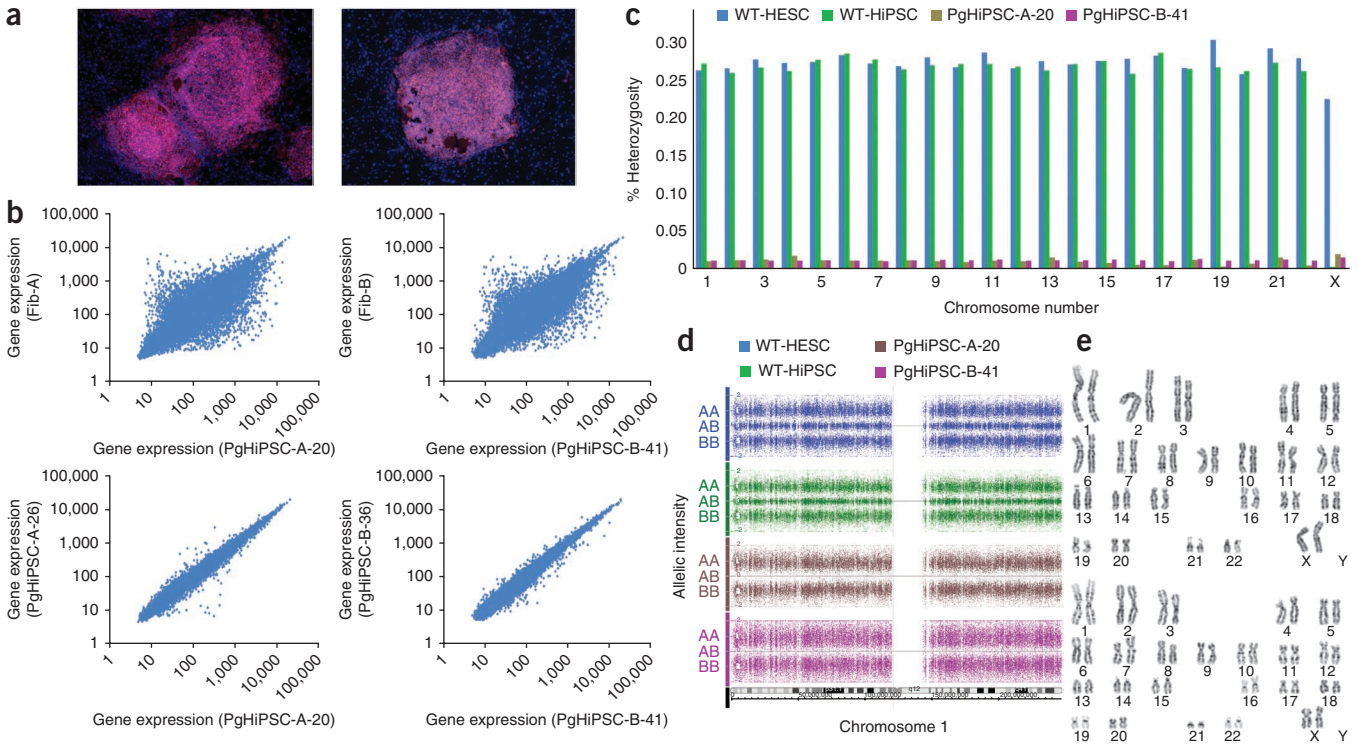

Figure 1 Characterization of parthenogenetic HiPSCs showing normal diploid homozygote genome. (**a**) Immunostaining of undifferentiated PgHiPSC-A-26 (left) and PgHiPSC-B-41 (right) lines for TRA-1-60 pluripotency marker. (**b**) Scatter-plot analysis demonstrating differences in expression signature between PgHiPSC lines and their parental fibroblasts, and similar expression signature between different established PgHiPSC lines. (**c**) Percentage of SNP heterozygote calls; *y* axis represents the percentage of heterozygosity. Chromosomes are ordered on the *x* axis. (**d**) Calls of probe intensity on chromosome number 1; *y* axis varies from heterozygote call (AB) to homozygote calls (AA or BB). (**e**) G-band karyotyping of two human parthenogenetic induced pluripotent stem cell clones, PgHiPSC-A-20 and PgHiPSC-B-41, that were established from two different patients.

PgHiPSCs comprise a complete homozygote diploid karyotype

Parthenogenesis occurs naturally in females as a result of the generation of diploid oocytes. Parthenogenesis may result from a defect in the first or second meiosis or from reduplication of the haploid genome (**Supplementary Fig. 3**). It was recently shown that the recombination signature of PgHESCs depends on the stage in which meiosis is arrested[9,10]. Here, we used high-density single-nucleotide polymorphism (SNP) analysis to verify the degree of heterozygosity in the PgHiPSC lines that were derived from the two patients. Notably, genotype analysis showed complete homozygosity of both the PgHiPSC-A and PgHiPSC-B lines, in contrast to the normal frequency of heterozygosity (27–29%) in wild-type control HiPSC and HESC lines (**Fig. 1c**,**d**). Because the evaluation of the genomic integrity of the teratomas and their derived PgHiPSC lines indicated that they had a normal diploid karyotype (**Fig. 1e**), we conclude that each of the teratoma tumor cell lines originated from a reduplication event in a mature ovum[11] (**Supplementary Fig. 3**).

Genome-wide gene expression analysis of PgHiPSCs

We have recently shown that although HiPSCs are derived from somatic cells, they predominantly maintain monoallelic expression of their imprinted genes[12]. Thus induced pluripotent stem cells that are derived from parthenogenetic ovarian teratomas should maintain their parthenogenetic nature. We therefore analyzed the global gene expression of the six PgHiPSC clones and both of the ovarian teratomas from which they were derived. As a control, we analyzed gene expression in normal HESC and HiPSC lines. Comparing the expression signature between PgHiPSC and HiPSC lines, we found a high correlation between the expression profiles, indicative of successful reprogramming (**Figs. 1b** and **2a** and **Supplementary Fig. 1d**). It is evident that pluripotency markers are upregulated in the reprogrammed clones relative to their parental cells, both in the normal and the parthenogenetic HiPSCs, whereas the somatic-specific genes are downregulated (**Fig. 2b**).

In spite of this general similarity between parthenogenetic and normal HiPSCs, close examination revealed that the gene expression profiles of these two cell types are distinct in several respects. Specifically, many genes known to be paternally expressed were downregulated in PgHiPSCs compared to WT-HiPSCs (**Fig. 2a**,**b**). To further analyze parental imprinting in PgHiPSCs, we examined the expression profile of all the known human paternally expressed genes (**Supplementary Fig. 4a** and **Supplementary Table 1**). We were able to divide these genes into three groups: (i) genes that are expressed in wild-type pluripotent stem cells and are downregulated in the parthenogenetic cells; (ii) genes that are expressed at similar levels in the normal and parthenogenetic stem cells; and (iii) genes that are not expressed in any of the pluripotent stem cells. Our analysis demonstrated that nearly 60% of the known paternally expressed genes are downregulated in the PgHiPSCs, whereas the majority of the remaining genes are not expressed at all in the pluripotent stem cells. Notably, our results did not confirm most of the putative paternally expressed genes recently suggested in a large bioinformatics survey[13] (**Supplementary Table 1**).

Of the 33 genes that were most substantially downregulated (by a factor of >5) in our analysis, 30 were known paternally expressed genes. Among the three genes that had not previously

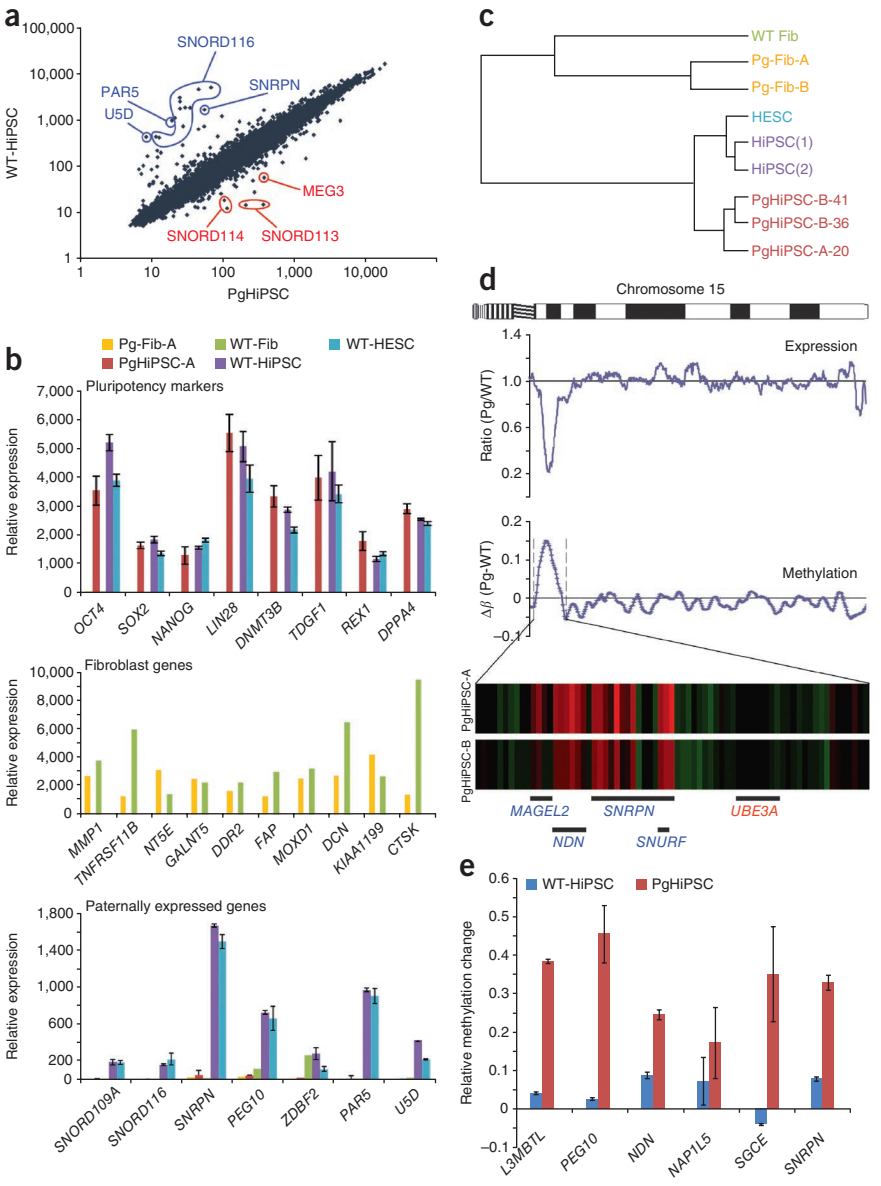

Figure 2 Global expression and DNA methylation analysis of imprinted genes in parthenogenetic HiPSCs. (**a**) Scatter-plot analysis of global gene expression in the PgHiPSC-A-20 versus the WT-HiPSC line. Values are displayed on a logarithmic scale. Labeled genes set as examples for down- and upregulation of known paternally (blue lines) and maternally (red lines) expressed genes, respectively. Notably, multiple points cluster representing different variants of *SNORD116*. (**b**) Expression levels of three PgHiPSC-A (average of clone nos. 11, 20 and 26), the fibroblasts from which they were derived (Pg-Fib-A), WT-HiPSC clones (average of two clones), their fibroblasts and three WT-HESC lines. The two histograms (top panels) demonstrate the reprogramming of the parental fibroblasts, consisting of pluripotency marker activation and silencing of fibroblast-specific genes. Histogram below represents the expression signature of selected paternally expressed genes. Shown are averages of gene expression and s.d. (**c**) Dendrogram hierarchical clustering based on global DNA methylation analysis in WT-HiPSCs, PgHiPSCs and their parental fibroblasts. Data consists of 25,578 CpG sites, which correlate to 14,495 genes. (**d**) Gene expression (top) and DNA methylation (middle) profiles; shown is a moving-average plot along chromosome 15 in the parthenogenetic and WT-HiPSCs. Bottom, DNA methylation heat map in various maternally (red) or paternally (blue) expressed genes. (**e**) Differences in DNA methylation levels (±s.d.) of various imprinted genes in PgHiPSCs (red column) and WT-HiPSCs (blue column) compared to the levels in HESC lines. Genes were sorted according to β values representing hemimethylation state (β 0.4–0.6) in the WT-HiPSCs and hypermethylation state ($\beta > 0.7$) in the PgHiPSCs.

been identified as imprinted genes, we showed that two (*TCN2* and *CHAC1*) were expressed in a biallelic fashion, whereas one (*U5D*, also called *RNU5D*) was expressed in a monoallelic fashion (see below). Out of the 13 genes that were downregulated less substantially (by a factor of 3–5) in the parthenogenetic cell lines, 6 are paternally expressed genes, and the rest were not previously suggested as putative imprinted genes. Notably, 12 genes were upregulated in the parthenogenetic cells (by a factor of >3), of which 5 are known maternally expressed genes (**Supplementary Table 1**). Among these genes, *MEG3*, the human homolog of mouse *Gtl2*, is overexpressed in PgHiPSCs in comparison to WT-HiPSCs and HESCs (**Fig. 2a**). Curiously, in contrast to recent findings in mice[14], we found that *MEG3* is not aberrantly silenced in most HESCs and HiPSCs and is highly expressed in the parthenogenetic cells. Because our PgHiPSC gene expression data could identify paternally expressed genes, we used our data to search for a genomic cluster of such genes. On the basis of differential expression between wild-type and parthenogenetic cells along the human chromosomes, we identified the Prader-Willi–Angelman region on chromosome 15q11.2 as the most prominent paternally expressed cluster (**Supplementary Fig. 4b**).

Genome-wide DNA methylation analysis of PgHiPSCs

Imprinted genes acquire their DNA methylation patterns in germ cells[15], and this signature is involved in regulating the parent-of-origin expression later in development. Because parthenogenetic embryos lack the paternal genome, it is expected that they will show an altered DNA methylation profile. Although allele-specific chromatin conformations also mark imprinted genes, DNA methylation is considered the hallmark mechanism regulating parental imprinting. We therefore performed a wide-scale DNA methylation analysis using Illumina's Infinium methylation assay, which allowed us to study a substantial portion of DNA methylation sites in the genome. As expected, we observed demethylation at sites located in the promoters and near transcription start sites of known pluripotency-related genes (**Supplementary Fig. 2c**). Moreover, when we used DNA methylation profiles to perform an unbiased clustering for all samples, we observed an evident epigenetic signature for all of the pluripotent cells that differed from those of their parental fibroblasts. Thus, we demonstrate that reprogramming also occurs at the DNA methylation level (**Fig. 2c**). Additionally, the PgHiPSCs derived from the two different

parthenogenetic cell lines were clustered together, further supporting the differences between the parthenogenetic and WT-HiPSCs seen in the expression data (**Fig. 2c**). We next plotted the differences in DNA methylation score (β value, ranging from 1 (highly methylated) to 0 (not methylated)) between PgHiPSCs and normal HiPSCs on a global scale. This analysis allows the identification of known differentially methylated regions (DMRs), the most prominent of which are located on chromosome 15 (the Prader-Willi–Angelman region, **Fig. 2d**) and chromosome 14 (the DLK1–DIO3 region, **Supplementary Fig. 5**). In these regions, the differential DNA methylation in the parthenogenetic cells as compared to wild-type cells correlates with differential gene expression. Paternally expressed genes may appear hemimethylated (score between 0.4 and 0.6) in HESCs and HiPSCs and methylated (score 0.75–1) in PgHiPSCs. This analysis identified nine genes, of which six are known to be paternally expressed (*L3MBTL*, *PEG10*, *NDN*, *NAP1L5*, *SGCE*, *SNRPN*; **Fig. 2e**) and three (*FBLN7*, *SIGLEC6* and *C1orf118*) had not previously been suggested to be paternally expressed. Notably, among the differentially methylated genes, we could identify even paternally expressed genes that are not expressed in pluripotent stem cells (*L3MBTL*, *NAP1L5* and *SGCE*).

snRNA variant of U5 spliceosome is a novel imprinted gene

Among the most strongly downregulated genes in the parthenogenetic cell lines was the gene for U5D, a small nuclear RNA (snRNA) variant of the U5 complex[16] (**Fig. 2a**). We further confirmed by RT-PCR that the expression of the U5D variant is indeed downregulated in the parthenogenetic cell lines (**Fig. 3a**). To determine whether *U5D* is an imprinted gene, we identified a SNP within the genomic DNA transcribed region in two different normal pluripotent stem cell lines and demonstrated monoallelic expression at the RNA level in both HESCs and HiPSCs (**Fig. 3b**). These independent assays substantiated the fact that *U5D* is imprinted and expressed exclusively from the paternal allele during human embryogenesis. The U5D variant is transcribed as a discrete unit to form its highly conserved secondary structure[16]. The gene is located at chromosome 1p34.1 and, like some other imprinted genes, is a stand-alone imprinted gene[17]. Notably, by

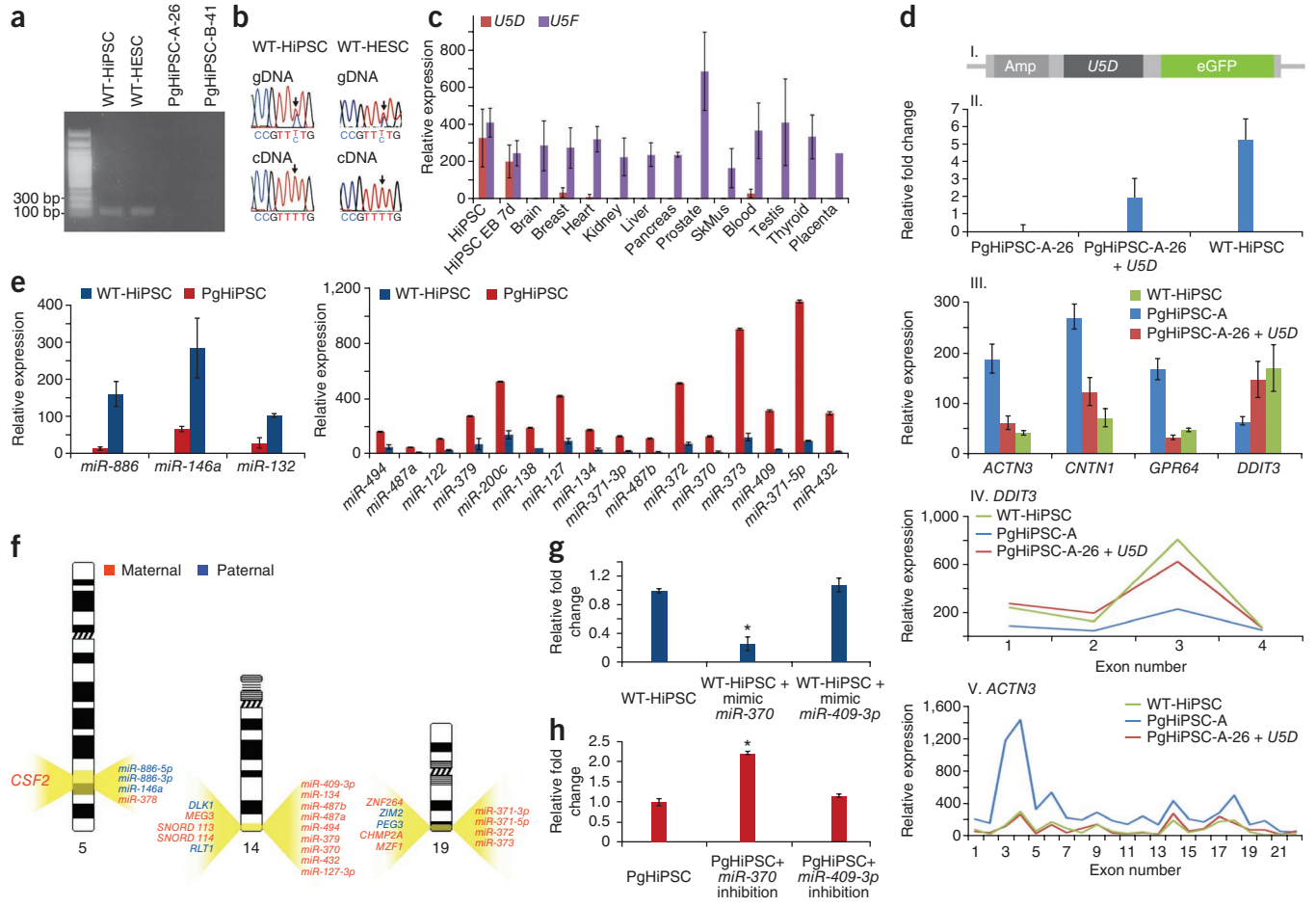

Figure 3 Small noncoding imprinted RNAs in parthenogenetic HiPSCs. (**a**) RT-PCR analysis of the expression of *U5D* variant in PgHiPSC-A-26, PgHiPSC-B-41, WT-HESCs and HiPSCs. (**b**) Genomic and cDNA sequencing of the transcribed region of *U5D* in WT-HiPSCs and HESCs. (**c**) Gene expression distribution (±s.d.) of two U5 snRNA variants in 12 cell types. (**d**) I, schematic diagram of the *U5D* overexpression vector. II, expression analysis (±s.d.) of *U5D* variant induced to PgHiPSCs. III, effect of *U5D* induction on expression of nonimprinted differentially expressed genes in PgHiPSCs (±s.d.). IV,V, *DDIT3* and *ACTN3* exon-level analyses, respectively, after transfection of PgHiPSCs with *U5D*; *x* axis indicates exon number. (**e**) Average expression levels of annotated human miRNA. Left histograms, miRNAs that are downregulated in PgHiPSCs compared to WT-HiPSCs; right histograms, miRNAs that are upregulated. Shown are average expression levels ± s.d. Parthenogenetic samples are composed of three clones from two cell lines (PgHiPSC-A-20/26 and PgHiPSC-B-41); WT-HiPSCs are composed of two different clones. (**f**) Genomic organization of clusters of differentially expressed miRNAs with respect to the known imprinted mRNAs that reside in the same locus. (**g**) qRT-PCR of the relative fold change (±s.e.m.) of CHAC1 levels following overexpression of *miRNA-370* and *miRNA-409-3p* in the WT-HiPSCs. (**h**) qRT-PCR of the relative fold change (±s.e.m.) of *CHAC1* expression following inhibition of miRNA-370 and miRNA-409-3p in the PgHiPS cells. These results demonstrate that *miRNA-370*, but not *miRNA-409*, affects the expression of *CHAC1*. $P < 10^{-5}$.

analyzing the expression patterns of the different U5 snRNA variants in 12 different tissues, we found that U5D, unlike other U5 variants, is predominantly expressed in pluripotent stem cells (**Fig. 3c**). It was previously shown that the U5 small nuclear ribonucleoprotein participates in pre-mRNA splicing through specific recognition of the 3′ splice site[18], and it has been suggested to have an essential role in normal development and cell growth[19]. To examine a possible role of U5D in regulating gene expression in parthenogenetic cells, we introduced the *U5D* gene into PgHiPSCs (**Fig. 3d**, I). Transfected parthenogenetic cells showed higher U5D expression than did PgHiPSCs (**Fig. 3d**, II). Intriguingly, as a result of U5D expression, several genes that showed aberrant expression in the parthenogenetic cells (by factors of >3) returned to expression levels similar to those of normal cells (**Fig. 3d**, III). To evaluate the possible role of U5D in the post-transcriptional regulation of the corrected genes, we performed an exon-level expression analysis to determine any change in splice variants (**Fig. 3d**, IV and V). These results indicated that the expression levels of several affected genes may have changed due to alternative splicing at specific exons. Alternative splicing may explain the changes in the detection levels of the complete transcript. This could be the result either of a reduction in the number of responding probes making up the transcript in the microarray, or of the destabilization of the entire transcript, for example by the introduction of premature stop codons causing nonsense-mediated decay[20].

Expression analysis of miRNAs in PgHiPSCs

Imprinted noncoding RNAs have been found in many imprinted clusters and are suggested to play a role in regulating gene expression by acting *in cis* and *in trans*. Most of the identified imprinted noncoding RNAs are C/D box-containing small nuclear RNAs (SNORDs). To date, only a few human miRNAs have been suggested to be imprinted, primarily on the basis of high homology to mouse imprinted miRNAs[21,22]. An effective assay for detecting imprinted miRNAs is to compare their expression signature in uniparental and disomic cells. We therefore set out to examine the expression of miRNAs in PgHESCs. We extracted total RNA from the WT-HiPSCs and PgHiPSCs and analyzed it with a microarray platform specific for the detection of human miRNAs. Notably, we observed a marked differential expression of various miRNAs between the two cells types by both DNA microarray and qRT-PCR analysis (**Fig. 3e**, **Supplementary Fig. 6a** and **Supplementary Table 1**). Thus, we identified three miRNAs that were downregulated (by a factor of >3) in the parthenogenetic cells, suggesting that they may be paternally expressed. Conversely, we identified 17 miRNAs that are upregulated (by a factor of >3) in the parthenogenetic cells and are therefore candidates to be maternally expressed. Among them is the embryonic stem cell–specific miRNA-371–miRNA-372–miRNA-373 cluster[23], which was suggested to be involved in testicular germ cell tumors[24] (**Fig. 3e**).

Notably, when we examined the genomic position of the 20 newly identified putative imprinted miRNAs, we found that they are localized in five genomic regions. Among these, three are clusters of imprinted miRNAs (**Fig. 3f**); one of these is a well-characterized cluster located in chromosome 14q32, and the other two, located on chromosomes 5q33.1–3 and 19q13.41, are previously unknown clusters of parentally imprinted miRNAs. Close examination of gene expression, together with the DNA methylation patterns, identified the imprinted cluster of *DLK1–DIO3*, which was recently shown to have a major role in the developmental capacity of mouse iPSCs[14], as the largest miRNA imprinted cluster in the human genome (**Fig. 3f** and **Supplementary Fig. 5**).

Because miRNAs are known to affect gene expression, we next examined the possibility that several of the differentially expressed genes that are not known imprinted genes are regulated by the differentially expressed miRNAs. We used target-prediction programs to find a possible match between the 3′ untranslated region of the mRNA genes and the corresponding imprinted miRNAs. Notably, we found that four of these genes (*TCN2*, *CHAC1*, *INHBE* and *LRRC17*) have a complete seed match with miRNAs that are upregulated in the parthenogenetic cell lines (**Supplementary Figs. 6b** and **7a**). To further examine the possibility that miRNA regulates these four genes, we overexpressed miRNA-370 and miRNA-409 in WT-iPS cells or, alternatively, downregulated these miRNAs in the PgHiPSCs (**Supplementary Fig. 7b**,**c**). We next examined the relative levels of expression of the four putative targets and showed that miRNA-370 regulates the expression of *CHAC1* and to a lesser extent also *INHBE*. Accordingly, *CHAC1* is differentially expressed in the parthenogenetic cells, although it is transcribed in a biallelic fashion (**Supplementary Fig. 7d**,**e**). Notably, *CHAC1* is downregulated by overexpression of miRNA-370 and upregulated when miRNA-370 is inhibited (**Fig. 3g** and **h**, respectively). Therefore, *CHAC1* seems to be downregulated in the parthenogenetic cells as a result of activation by an imprinted miRNA.

Differentiation capacity of the PgHiPSCs

Mouse parthenogenetic embryonic stem cells show an impaired developmental potential *in vivo*, mainly in the trophectoderm but also in some embryonic tissues[4,5,25–28]. Human parthenogenetic fertilization commonly ends in benign ovarian teratoma tumors, and viable, sick offspring have occasionally been reported, in cases of chimerism[29]. Nevertheless, to date, no documentation regarding the developmental potential of human parthenogenetic embryos is recorded. Thus, it is of great interest to examine the developmental potential of human parthenogenesis by studying the differentiation capacity of PgHiPSCs. To address this issue, we have differentiated PgHiPSCs both *in vitro* and *in vivo*, analyzing the expression signature of PgHiPSCs in 20-d-old embryoid bodies and in teratomas formed under the kidney capsule of NOD-SCID mice. We quantified the level of perturbation for eight different tissues by clustering the 200 most exclusively expressed genes in each tissue and computing the average difference in the ratio of gene expression between WT-HiPSCs and PgHiPSCs (**Fig. 4a**). Notably, our results indicated that the endoderm lineage is consistently affected in both *in vivo* and *in vitro* assays. Furthermore, the tissue-specific gene analysis demonstrates that the liver is the most prominent tissue impaired in human parthenogenesis, which is suggested by a specific downregulation of genes that characterize the hepatic endoderm (**Fig. 4b** and **Supplementary Table 2**). In addition, genes specific to skeletal muscle and to the kidney were also affected in the parthenogenetic cells. It has been suggested that liver and muscle tissues are affected in parthenogenetic mice; however, this effect was not as prominent as that on the placenta.

In our analysis, the placenta-specific genes were also affected in the differentiation of PgHiPSCs into embryoid bodies as well as into teratomas (**Fig. 4a**). It was previously shown that BMP4 can promote differentiation of HESCs into the trophectoderm[30]. We therefore analyzed the gene expression of the 100 most exclusively expressed trophectoderm genes in PgHiPSC and WT-HiPSC controls following 7 d of treatment with 50 ng ml^{-1} of BMP4 (**Fig. 4c** and **Supplementary Table 3**). Evidently, among the top 100 trophectoderm specific genes, nearly 50% are substantially downregulated (by a factor of >1.5). Among the most strongly downregulated genes are several encoding chorionic gonadotrophins (*CGB*, *CGB1*, *CGB5* and *CGA*), which are

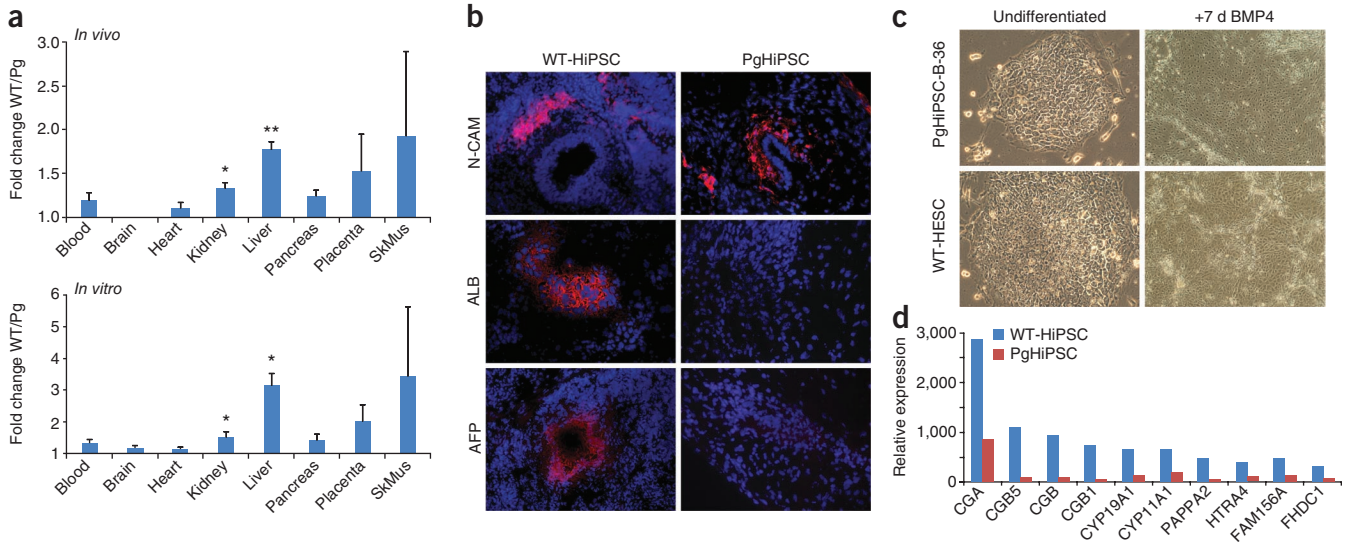

Figure 4 Differentiation potential of human parthenogenesis. (**a**) PgHiPSCs and WT-HiPSCs were differentiated *in vitro* into embryoid bodies and *in vivo* into teratomas. For each of the eight tissues, the gene expression of its 200 most exclusively cell-specific genes were compared between PgHiPSCs and WT-HiPSCs. Average fold difference ratio between PgHiPSCs and WT-HiPSCs was calculated for expressing genes only. Top, *in vivo* differentiation (gene expression of 30-d-old teratomas); bottom, *in vitro* differentiation (gene expression of 20-d-old embryoid bodies). Shown are average relative expression levels ± s.d. $*P < 0.05$, $**P < 10^{-6}$. (**b**) Immunostaining for representative hepatic-endoderm markers in both WT-HiPSCs and PgHiPSCs. Shown are differential expression of albumin (ALB) and α-fetoprotein (AFP) and similar expression of the neural marker N-CAM (as control). (**c**) Typical morphology of pluripotent cells (left) and morphology after 7 d of treatment with 50 ng ml^{-1} of BMP4 (right). (**d**) Expression levels of ten placental-specific genes showing substantial downregulation in the PgHiPSC cells.

the hallmark characteristic of trophectoderm differentiation (**Fig. 4d** and **Supplementary Fig. 8a**). Thus, our results indicate that, much like mouse parthenogenotes, PgHiPSCs show restricted growth potential of the extraembryonic trophoblast.

We next asked whether manipulation of U5D expression in the PgHiPSCs could affect the perturbed differentiation phenotype of these cells. Because of the unique structure of U5D and its biogenesis[16,31,32], we manipulated its levels by overexpression, rather than downregulation. Thus, a *U5D*-containing vector was stably introduced into PgHiPSCs (**Supplementary Fig. 8b**), and the cells were differentiated *in vitro* into 20-d-old embryoid bodies. Notably, by similarly quantifying the levels of perturbation for the eight different tissues, we could detect a substantial upregulation of hepatic-endoderm genes. Thus, instead of the downregulation in liver-enriched genes by a factor of 3.15 observed in PgHiPSCs, we detected a downregulation of only a factor of 1.18 (**Supplementary Fig. 8c**, $P < 10^{-5}$) after U5D expression in the parthenogenetic cells. Therefore, we now suggest that U5D may directly or indirectly regulate the differentiation of early hepatic endoderm.

DISCUSSION

Aberrant expression of imprinted genes results in human disorders with developmental anomalies[33]. In the mouse, parthenogenetic embryos were used to study imprinted genes and their effects on development. Female ovarian teratomas are benign tumors composed of parthenogenetic cells from which parthenogenetic fibroblasts can be derived. However, many imprinted genes are not expressed at all in both wild-type and parthenogenetic fibroblasts (**Fig. 2b**), so that this system is not very informative for characterizing the regulation of imprinted genes during development. We therefore set out to derive HiPSCs from parthenogenetic fibroblasts in order to reveal genes that are parentally imprinted during human embryogenesis. Our experimental system correctly identified most of the known imprinted genes, while enabling the identification of previously unknown imprinted genes.

Recent studies have emphasized the importance of miRNA regulation in HESCs and in early human development[34,35]. Our data suggests that imprinting is also important in the regulation of miRNAs, revealing a new dimension in the regulation of imprinting and parthenogenesis. By comparing the expression of miRNAs in WT-HiPSCs to that in PgHiPSCs, we have identified three previously unrecognized paternally expressed miRNAs whose expression is downregulated in the parthenogenetic cells. Intriguingly, we could also identify maternally expressed miRNAs that are upregulated in the PgHiPSCs. It is assumed that maternally expressed genes would show approximately doubled expression levels in PgHiPSCs compared to WT-HiPSCs because they harbor two maternal alleles. However, in our system several of the maternally expressed genes are upregulated by more than three-fold, suggesting that another mechanism may be involved that regulates the maternally expressed genes. One possible explanation of this phenomenon is regulation of some maternally expressed miRNAs by overlapping antisense transcripts that are expressed from the opposite strand of the paternal allele, thus regulating the corresponding miRNAs *in trans*. Such interrelationships between sense and antisense transcripts were previously suggested to regulate miRNA-127 (ref. 21).

The involvement of miRNAs in parthenogenesis suggests another layer of control of parental imprinting. Thus, alleles from a specific parent may either directly regulate cell differentiation or indirectly control the biological processes by inhibiting the expression of target genes. This involvement of imprinted regulatory noncoding RNAs is also evidenced in our discovery that U5D, a key RNA component of the splicing machinery, is imprinted and expressed exclusively from the paternal allele in human pluripotent stem cells. We therefore suggest that a set of imprinted noncoding small RNAs are responsible for part of the phenotype of parthenogenesis through their regulation

of the expression of their target genes. It will therefore be of great interest to study these complex interrelationships and their link to development and disease.

The availability of human parthenogenetic embryonic-like cells should enable researchers to further analyze the genes and processes that are regulated by the maternal and paternal genomes and thus improve understanding of the epigenetic processes that are involved in this phenomenon.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/nsmb/.

Accession code. NCBI Gene Expression Omnibus: all datasets are deposited under accession number GSE27362.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank M. Pick for her help with generating parthenogenetic iPSCs and O. Bar-Nur for providing WT-HiPSCs; U. Ben-David for his technical help with the teratoma formation experiments; T. Golan-Lev for her assistance with the graphic design and immunostaining assays; and A. Eden, D. Kitsberg, Y. Mayshar, D. Ronen and N. Sharon for critical reading of the manuscript. N.B. is the Herbert Cohn Chair in Cancer Research. This research was supported partially by the Israel Science Foundation–Morasha Foundation (grant no. 943/09).

AUTHOR CONTRIBUTIONS

Y.S.: conception and design, collection and assembly of data, data analysis and interpretation, and manuscript writing; O.Y.: conception and design, and collection and assembly of data; N.B.: conception and design, financial support, data analysis and interpretation, manuscript writing and final approval of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/nsmb/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
2. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
3. Park, I.H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
4. McGrath, J. & Solter, D. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37**, 179–183 (1984).
5. Surani, M.A. & Barton, S.C. Development of gynogenetic eggs in the mouse: implications for parthenogenetic embryos. *Science* **222**, 1034–1036 (1983).
6. Revazova, E.S. *et al.* Patient-specific stem cell lines derived from human parthenogenetic blastocysts. *Cloning Stem Cells* **9**, 432–449 (2007).
7. Mai, Q. *et al.* Derivation of human embryonic stem cell lines from parthenogenetic blastocysts. *Cell Res.* **17**, 1008–1019 (2007).
8. Brevini, T.A. *et al.* Cell lines derived from human parthenogenetic embryos can display aberrant centriole distribution and altered expression levels of mitotic spindle check-point transcripts. *Stem Cell Rev.* **5**, 340–352 (2009).
9. Kim, K. *et al.* Histocompatible embryonic stem cells by parthenogenesis. *Science* **315**, 482–486 (2007).
10. Kim, K. *et al.* Recombination signatures distinguish embryonic stem cells derived by parthenogenesis and somatic cell nuclear transfer. *Cell Stem Cell* **1**, 346–352 (2007).
11. Surti, U., Hoffner, L., Chakravarti, A. & Ferrell, R.E. Genetics and biology of human ovarian teratomas. I. Cytogenetic analysis and mechanism of origin. *Am. J. Hum. Genet.* **47**, 635–643 (1990).
12. Pick, M. *et al.* Clone- and gene-specific aberrations of parental imprinting in human induced pluripotent stem cells. *Stem Cells* **27**, 2686–2690 (2009).
13. Luedi, P.P. *et al.* Computational and experimental identification of novel human imprinted genes. *Genome Res.* **17**, 1723–1730 (2007).
14. Stadtfeld, M. *et al.* Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465**, 175–181 (2010).
15. Tucker, K.L. *et al.* Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. *Genes Dev.* **10**, 1008–1020 (1996).
16. Sontheimer, E.J. & Steitz, J.A. Three novel functional variants of human U5 small nuclear RNA. *Mol. Cell. Biol.* **12**, 734–746 (1992).
17. Edwards, C.A. & Ferguson-Smith, A.C. Mechanisms regulating imprinted genes in clusters. *Curr. Opin. Cell Biol.* **19**, 281–289 (2007).
18. Chabot, B., Black, D.L., LeMaster, D.M. & Steitz, J.A. The 3′ splice site of pre-messenger RNA is recognized by a small nuclear ribonucleoprotein. *Science* **230**, 1344–1349 (1985).
19. Patterson, B. & Guthrie, C. An essential yeast snRNA with a U5-like domain is required for splicing in vivo. *Cell* **49**, 613–624 (1987).
20. Hentze, M.W. & Kulozik, A.E. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell* **96**, 307–310 (1999).
21. Seitz, H. *et al.* Imprinted microRNA genes transcribed antisense to a reciprocally imprinted retrotransposon-like gene. *Nat. Genet.* **34**, 261–262 (2003).
22. Seitz, H. *et al.* A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res.* **14**, 1741–1748 (2004).
23. Suh, M.R. *et al.* Human embryonic stem cells express a unique set of microRNAs. *Dev. Biol.* **270**, 488–498 (2004).
24. Voorhoeve, P.M. *et al.* A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell* **124**, 1169–1181 (2006).
25. Fundele, R.H. *et al.* Temporal and spatial selection against parthenogenetic cells during development of fetal chimeras. *Development* **108**, 203–211 (1990).
26. Thomson, J.A. & Solter, D. The developmental fate of androgenetic, parthenogenetic, and gynogenetic cells in chimeric gastrulating mouse embryos. *Genes Dev.* **2**, 1344–1351 (1988).
27. Thomson, J.A. & Solter, D. Chimeras between parthenogenetic or androgenetic blastomeres and normal embryos: allocation to the inner cell mass and trophectoderm. *Dev. Biol.* **131**, 580–583 (1989).
28. Nagy, A., Sass, M. & Markkula, M. Systematic non-uniform distribution of parthenogenetic cells in adult mouse chimaeras. *Development* **106**, 321–324 (1989).
29. Strain, L., Warner, J.P., Johnston, T. & Bonthron, D.T. A human parthenogenetic chimaera. *Nat. Genet.* **11**, 164–169 (1995).
30. Xu, R.H. *et al.* BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.* **20**, 1261–1264 (2002).
31. Will, C.L. & Luhrmann, R. Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.* **13**, 290–301 (2001).
32. Karijolich, J. & Yu, Y.T. Spliceosomal snRNA modifications and their function. *RNA Biol.* **7**, 192–204 (2010).
33. Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nat. Rev. Genet.* **2**, 21–32 (2001).
34. Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S. & Hannon, G.J. Characterization of Dicer-deficient murine embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **102**, 12135–12140 (2005).
35. Kanellopoulou, C. *et al.* Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* **19**, 489–501 (2005).

ONLINE METHODS

Cell culture. Ovarian teratoma primary fibroblasts of cell lines GM01304 (Fib-A) and GM01306 (Fib-B) were purchased from the Coriell Institute and cultured in MEM (Sigma M5650) medium supplemented with 15% FCS, 2 mM L-glutamine, 50 units ml^{-1} penicillin and 50 $\mu g\ ml^{-1}$ streptomycin. PgHiPSCs cultures were grown in HESC medium containing Knockout Dulbecco's Modified Eagle's Medium (Gibco) supplemented with 15% Knockout Serum Replacement (Gibco), 2 mM L-glutamine, 0.1 mM nonessential amino acids, 0.1 mM β-mercaptoethanol and 4–8 ng ml^{-1} basic Fibroblast Growth Factor (bFGF) (Cytolab). All cells were maintained in a humidified incubator at 37 °C and 5% CO_2. For embryoid body formation, a six-well plate of semiconfluent PgHiPSCs was harvested using trypsin, and cell clumps were resuspended in HESC medium without bFGF, allowed to aggregate and transferred to one well of a six-well plate (Greiner). After 20 d, embryoid body RNA was isolated and analyzed. Monolayer differentiation was performed by culturing the PgHiPSCs in monolayer deprived of mouse embryonic fibroblasts (MEFs) and using MEM medium without basic-FGF.

Production of parthenogenetic induced pluripotent stem cells. For production of viral particles, 293T cells were transfected using Fugene 6 (Roche) with 4.5 μg of pMXs retroviral vectors containing human *OCT4*, *SOX2*, *KLF4* or *c-MYC* and 4.5 μg of PCL-Ampho plasmid. At 24 h after transfection, the culture was replaced with fresh medium, and 48 h after transfection, the supernatant was collected, filtered through a 0.45-μM cellulose acetate filter (Whatman) and supplemented with 4 $\mu g\ ml^{-1}$ of Polybrene (Sigma-Aldrich). 600,000 fibroblast cells from each patient were infected with viral particles encoding the four factors[36]. On day 4 after infection, the cells were transferred to HESC growth conditions, supported by mitomycin-C–treated mouse embryonic feeders. Morphological changes began to occur about 12 d after infection. Colonies with embryonic stem cell–like morphology were manually expanded.

Alkaline phosphatase staining and immunocytochemistry. Alkaline phosphatase staining was performed according to the instructions of the Leukocyte Alkaline Phosphatase kit (Sigma). For immunocytochemistry staining, cells were fixed with PBS containing 4%(w/v) paraformaldehyde for 10 min at room temperature. After washing with PBS, the cells were blocked for 1 h with PBS containing 2% (w/v) bovine serum albumin (BSA; Sigma), and 0.1% (v/v) Triton X-100 (Sigma). For albumin staining, cells were blocked for 1 h with PBS containing 2% (w/v) skim milk (Difco) and 0.1%(v/v) Triton X-100. Staining with primary antibodies was performed for 1 h at room temperature with antibodies (diluted in blocking buffer as indicated) to OCT4 (1:150, Santa Cruz), TRA-1-60 (1:200, Santa Cruz), NCAM (1:50, BioLegend), ALB (1:50, Santa Cruz) and α-fetoprotein (1:50, Santa Cruz). Secondary antibodies used were cyanine 3 (Cy3)-conjugated rabbit anti-mouse IgM or rabbit anti-mouse IgG (1:200, Jackson Immunoresearch Laboratories) following staining with Hoechst 33342 to detect the cell's nucleus (Invitrogen).

Teratoma formation. Each of the PgHiPSC lines was grown to semi-confluency in a 10-cm tissue culture plate and harvested by trypsin treatment. The cells were then injected under the kidney capsule of NOD-SCID mice (Jackson Laboratory). Six weeks after injection, tumors were dissected and cryopreserved in O.C.T. compound (Sakura Finetek). Several 1-μm slices were then cut and stained with hematoxylin and eosin.

Genomic DNA and RNA isolation and reverse transcription. Total genomic DNA was extracted using genomic DNA extraction kit (Real Biotech Corporation), and RNA (DNase treated) was purified with Total RNA Extraction kit (RBC). One microgram of total RNA was used for reverse transcription reaction using ImProm-II reverse transcriptase (Promega). For sequencing and quantitative experiments, PCRs were performed with ReadyMix (Sigma); for overexpression experiments, PCR reactions used Herculase II Fusion DNA polymerase (Agilent Technologies). Quantitative real-time PCR was performed with 1 μg of RNA reverse transcribed to cDNA and TaqMan Universal Master Mix or SYBR Green qPCR Supermix (see primer list below; Applied Biosystems) and analyzed with the 7300 real-time PCR system (Applied Biosystems). A full description of primer sequences and annealing temperature can be found in **Supplementary Table 4**.

DNA microarray analysis. Total RNA was extracted according to the manufacturer's protocol (Affymetrix). RNA was subjected to Human Gene 1.0 ST microarray platform (Affymetrix) analysis; washing and scanning were performed according to the manufacturer's protocol. Arrays were analyzed using Robust Multichip Analysis (RMA) in the Affymetrix Expression Console. Hierarchical clustering was performed with Partek Genomics Suite version 6.3 (Partek, http://www.partek.com).

SNP analysis. DNA was isolated from the PgHiPSC and WT-HiPSC lines and analyzed with the Genome-Wide Human SNP Array 6.0 (Affymetrix) according to the manufacturer's protocol (Affymetrix). Genotyping and allele differences were identified using the Genotyping Console (Affymetrix), with the 270 samples of the Human HapMap Project used as baseline.

Array-based DNA methylation analysis. The HumanMethylation27 BeadChip comprises 27,578 array probes CpG sites representing 14,495 genes with an average of two probes per gene. 500 ng of genomic DNA from each sample was bisulfite converted using the EZ-96 DNA Methylation Kit (Zymo Research Corporation) according to the manufacturer's recommendations. After bisulfite conversion, each sample was whole genome amplified and enzymatically fragmented, and about 200 ng DNA was applied to BeadChips (Illumina). DNA methylation levels were scored as β values using Illumina's Genome studio software. The β value is calculated as the ratio of the methylated signal over the total fluorescent signal on a scale of 0–1. Thus, the β value indicates the observed methylation of a CpG site relative to the site's full methylation potential[37]. We also obtained the detection P values, which indicate DNA methylation measurement quality. Measurements with $P < 0.05$ are considered to have signal intensity above background. Measurements with $P > 0.05$ were masked.

Introduction of *U5D* into the parthenogenetic iPSCs. The *U5D* overexpression sequence was amplified using PCR (see **Supplementary Table 4** for a full primer list). The PCR product was then subcloned into the pGEM vector (Promega) and infused into the N1-EGFP construct. For experiments involving overexpression, TransIT-LT1 (Mirus) transfection reagent was used to transiently transfect PgHiPS cell. Between 12% and 14% GFP-positive cells were sorted, and RNA was isolated from the purified population. For the differentiation experiments, after transfection, cells were treated with 40 $mg\ \mu l^{-1}$ of G418. Colonies resistant to the antibiotic were manually expanded.

Inhibition and overexpression of miRNA in the PgHiPS and WT-HiPS cells. We used Dharmacon constructs to modify the expression of miRNAs in the parthenogenetic and WT-HiPS cells. miRIDIAN microRNA Hairpin Inhibitor was transfected into the PgHiPS cells, using the transfection reagent Lipofectamine 2000 (Invitrogen). Alternatively, meridian microRNA mimic was transfected into WT-HiPS cells. At 48 h after transfection, cells were harvested and RNA was extracted from the cells.

miRNA expression analysis. Total RNA was extracted using MirVana miRNA isolation kit (Ambion) according to the manufacturer's protocol. RNA was subjected to Human GeneChip miRNA array platform analysis (Affymetrix); washing and scanning were performed according to the manufacturer's protocol. Arrays were analyzed using miRNA QC Tool. For specific miRNA expression, 1 μg of RNA was reversed transcribed using TaqMan MicroRNA Reverse Transcription Kit. Quantitative real-time PCR was performed with TaqMan Universal Master Mix or SYBR Green qPCR Supermix (see **Supplementary Table 4** for a full primer list).

36. Park, I.H., Lerou, P.H., Zhao, R., Huo, H. & Daley, G.Q. Generation of human induced pluripotent stem cells. *Nat. Protoc.* **3**, 1180–1186 (2008).
37. Le Bibikova, M. *et al*. Genome-wide DNA methylation profiling using Infinium assay. *Epigenomics* **1**, 177–200 (2009).

doi:10.1038/nsmb.2050