

Supplementary Information

The microRNA cluster C19MC confers differentiation potential into trophoblast lineages upon human pluripotent stem cells

Norio Kobayashi¹, Hiroaki Okae^{1,*}, Hitoshi Hiura², Naoto Kubota³, Eri H Kobayashi¹, Shun Shibata¹, Akira Oike¹, Takeshi Hori⁴, Chie Kikutake³, Hirotaka Hamada¹, Hirokazu Kaji⁴, Mikita Suyama³, Marie-Line Bortolin-Cavaillé⁵, Jérôme Cavaillé⁵, Takahiro Arima^{1,*}

¹Department of Informative Genetics, Environment and Genome Research Center, Tohoku University Graduate School of Medicine, Sendai 980-8575, Japan

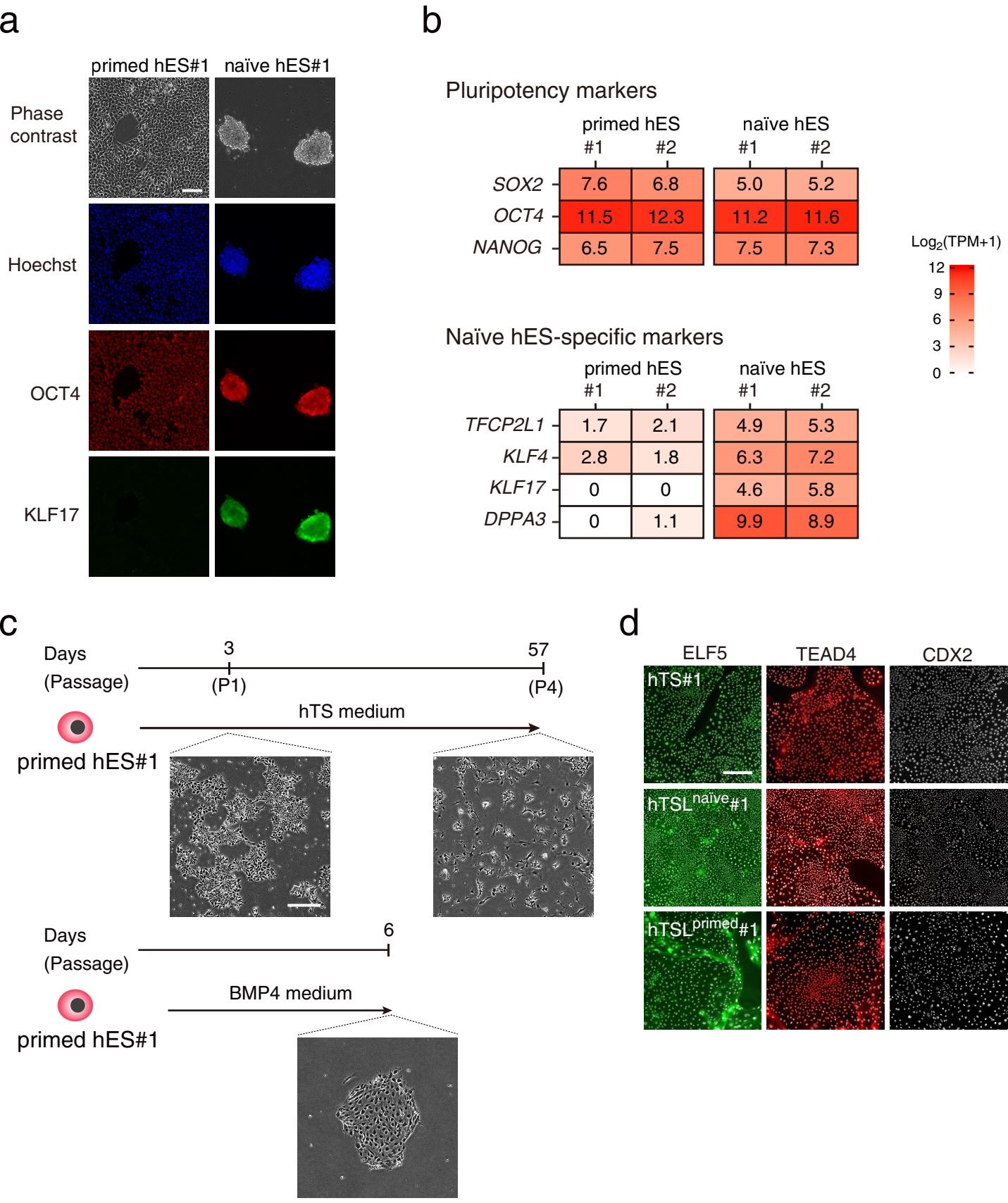
²Department of Bioscience, Faculty of Life Science, Tokyo University of Agriculture, Tokyo 156-8502, Japan

³Division of Bioinformatics, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan

⁴Department of Biomechanics, Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, Tokyo 101-0062, Japan

⁵Molecular, Cellular and Developmental biology department (MCD), Centre de Biologie Intégrative (CBI), University of Toulouse, CNRS, UPS, 31062 Toulouse, France

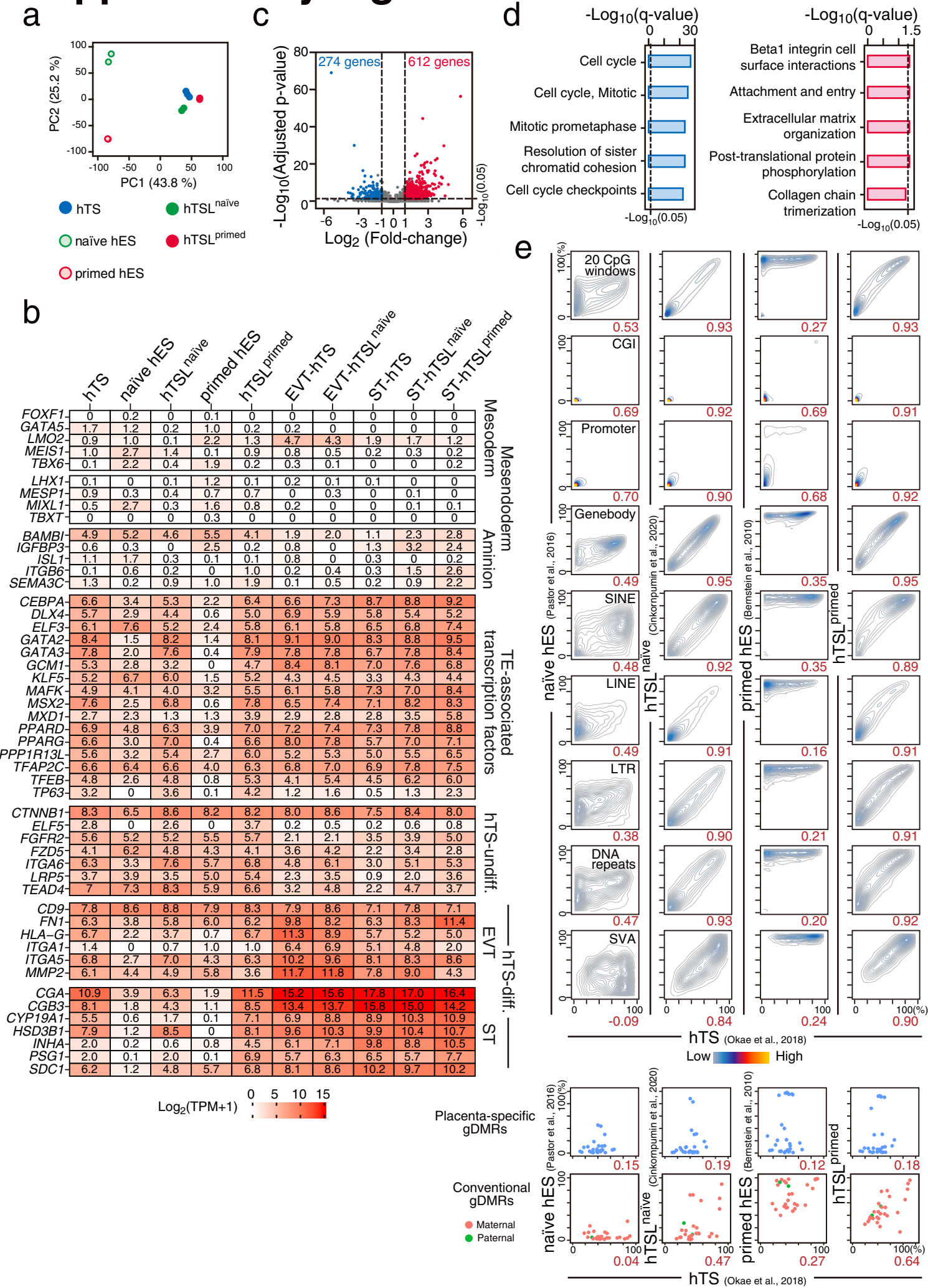
Supplementary Fig. 1



Supplementary Figure 1. Characterization of primed and naïve hES cells and schematic representation of hTSL cell derivation.

(a) Immunostaining of a pluripotency marker (OCT4) and a naïve hES cell-specific marker (KLF17) in primed and naïve hES cells. Nuclei were stained with Hoechst 33258. The scale bar indicates 100 μm . Similar results were obtained with two independent cell lines. (b) Expression levels of pluripotency markers and naïve hES cell-specific markers in primed and naïve hES cells. The data were obtained by RNA-seq. (c) Optimization of the duration of BMP4 treatment. When BMP4 treatment was omitted (upper panel), only cells with mesenchymal morphology proliferated. When primed hES cells were treated with BMP4 for six days (lower panel), cells with epithelial morphology appeared. However, these cells could not survive in hTS medium. The scale bar indicates 300 μm . Similar results were obtained with two independent cell lines. (d) Immunostaining of ELF5, TEAD4, and CDX2 in hTS and hTSL cells. Nuclei were stained with Hoechst 33258 (shown in gray). The scale bar indicates 200 μm . Similar results were obtained with two independent cell lines.

Supplementary Fig. 2



Supplementary Figure 2. Expression levels of lineage markers and DNA methylation levels of various genomic features in hTSL cells.

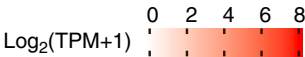
(a) Principal component analysis (PCA) of hES, hTS, and hTSL cells. Only 8,658 differentially expressed genes (TPM > 4 in at least one cell type, adjusted p-value < 0.05, fold change > 2) were analyzed. (b) Expression levels of lineage markers in hTSL cells and their differentiated derivatives. We performed RNA-seq and analyzed mesoderm and mesendoderm markers¹, amnion markers^{2,3}, TE-associated transcription factors⁴, and markers for undifferentiated and differentiated hTS cells⁵. (c) Volcano plot comparing hTSL^{primed} cells with hTS and hTSL^{naïve} cells. Gene expression levels in hTSL^{primed} cells were compared to those in hTS and hTSL^{naïve} cells, and expressed as log2 fold change. The down- and up-regulated genes are shown in blue and red, respectively (TPM > 4 in at least one cell type, adjusted p-value < 0.05, fold change > 2). (d) Pathways enriched among the 274 down-regulated and 612 up-regulated genes (blue and red dots in Supplementary Fig. 2c). The top five pathways are shown with q-values. (e) Comparisons of DNA methylation levels of various genomic features between hES, hTS, and hTSL cells. Global DNA methylation patterns were compared using 20 CpG windows and represented as contour plots (top). Contour plots for CGI, promoters, gene bodies, and retrotransposons (SINE, LINE, LTR, DNA repeats, and SVA) are also represented (top). DNA methylation levels of placenta-specific and conventional gDMRs are displayed as scatter plots (bottom). Red numbers indicate Pearson correlation coefficients. Publicly available WGBS data were used for hTS cells⁵, naïve hES cells⁶, hTSL^{naïve} cells², and primed hES cells⁷.

Supplementary Fig. 3

a

Imprinted genes associated with
placenta-specific gDMRs

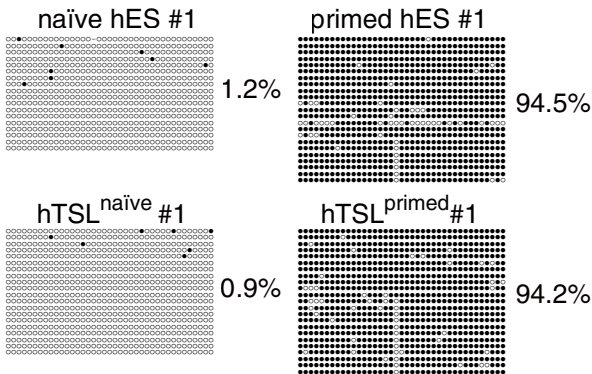
	hTS	naïve hES	hTSL naïve	primed hES	hTSL primed
ACCS	0.6	0.7	0.2	1.3	2.5
AGBL3	0.6	0.5	1.0	1.1	0.9
AIM1	4.3	1.3	5.2	2.9	5.9
C19MC	NA	NA	NA	NA	NA
CMTM3	5.2	3.5	5.8	5.4	5.9
DNMT1	6.4	4.3	7.6	6.3	7.8
GLIS3	1.5	0	2.1	0.1	2.5
HECW1	0	0.1	0	0.9	0.1
JMJD1C	2.4	3.3	3.5	5.4	3.3
KLHDC10	3.3	0.9	3.3	3.1	3.8
LIN28B	5.0	5.9	7.2	5.9	4.2
MCCC1	4.9	4.1	5.0	4.4	5.9
N4BP2L1	0.4	1.0	0.4	1.8	1.3
PCK2	5.2	5.8	6.2	6.7	3.9
PDE4D	2.8	1.6	4.7	0.9	3.9
PROSER2-AS1	3.1	0.9	0.1	0	0
RGMA	3.8	2.0	4.5	4.8	3.4
SLC4A7	2.1	4.0	3.7	5.1	2.1
ST8SIA1	0.5	0	0.3	0.3	0
VSTM5	2.6	0.2	1.8	0.4	4.4
VWA2	0	0	0.1	1.3	0.9
ZFAT	6.5	2.1	2.7	2.5	2.6
ZNF396	0.6	0.2	1.0	0.9	1.6



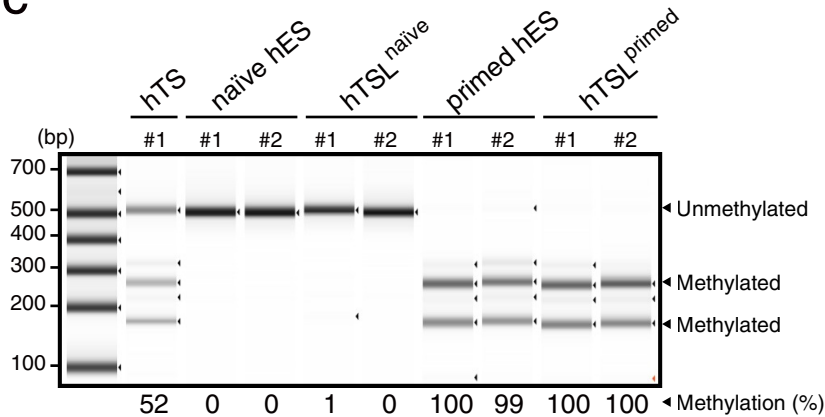
Imprinted genes associated with
conventional gDMRs

	hTS	naïve hES	hTSL naïve	primed hES	hTSL primed
Maternal					
BLCAP	6.0	4.9	6.1	5.7	7.1
CXORF56	4.7	5.2	5.4	4.6	4.8
DIRAS3	0.7	0	0.8	0	0.6
FAM50B	2.6	1.8	0.9	2.3	1.2
GNAS	9.3	9.3	9.5	10.1	9.9
GRB10	5.2	4.2	4.9	5.6	6.2
HTR5A	0	0.1	0	0	0
IGF1R	3.3	2.7	4.1	4.7	4.0
INPP5F	4.9	6.5	6.1	6.9	5.3
L3MBTL	2.3	1.4	1.8	2.4	2.7
HM13	7.6	7.0	7.5	6.4	7.8
MEST	6.7	7.0	7.2	7.7	7.5
NAP1L5	2.7	1.6	1.4	1.7	2.7
NHP2L1	7.9	8.4	8.2	7.8	7.5
PEG10	9.5	3.4	9.7	6.2	11.2
PEG3	6.4	2.8	6.8	2.7	7.9
PLAGL1	3.1	0.1	3.3	0.7	4.0
PPIEL	1.5	1.6	1.5	1.7	1.3
RB1	3.0	0.3	4.2	3.1	3.1
SNURF	2.6	6.8	7.8	10.1	5.9
TRAPPC9	3.6	2.5	3.5	3.1	4.0
ZNF331	4.2	3.6	3.7	4.1	4.7
ZNF597	0.9	2.2	0.3	1.4	0.8
KCNQ1OT1	0.1	0.1	0	0	0.1
CDKN1C	9.8	6.1	6.4	5.6	10.8
KvDMR1					
Paternal					
H19	13.4	7.4	13.7	2.0	14.3
IGF2	8.9	1.5	0.1	1.3	10.9
IGF2-AS	2.4	0.1	0	0.1	4.0
INS-IGF2	0	0	0	0	0
MEG3	0.1	9.0	0.1	5.2	0.1
DLK1	0	2.7	0	0.7	0
IG-DMR					

b



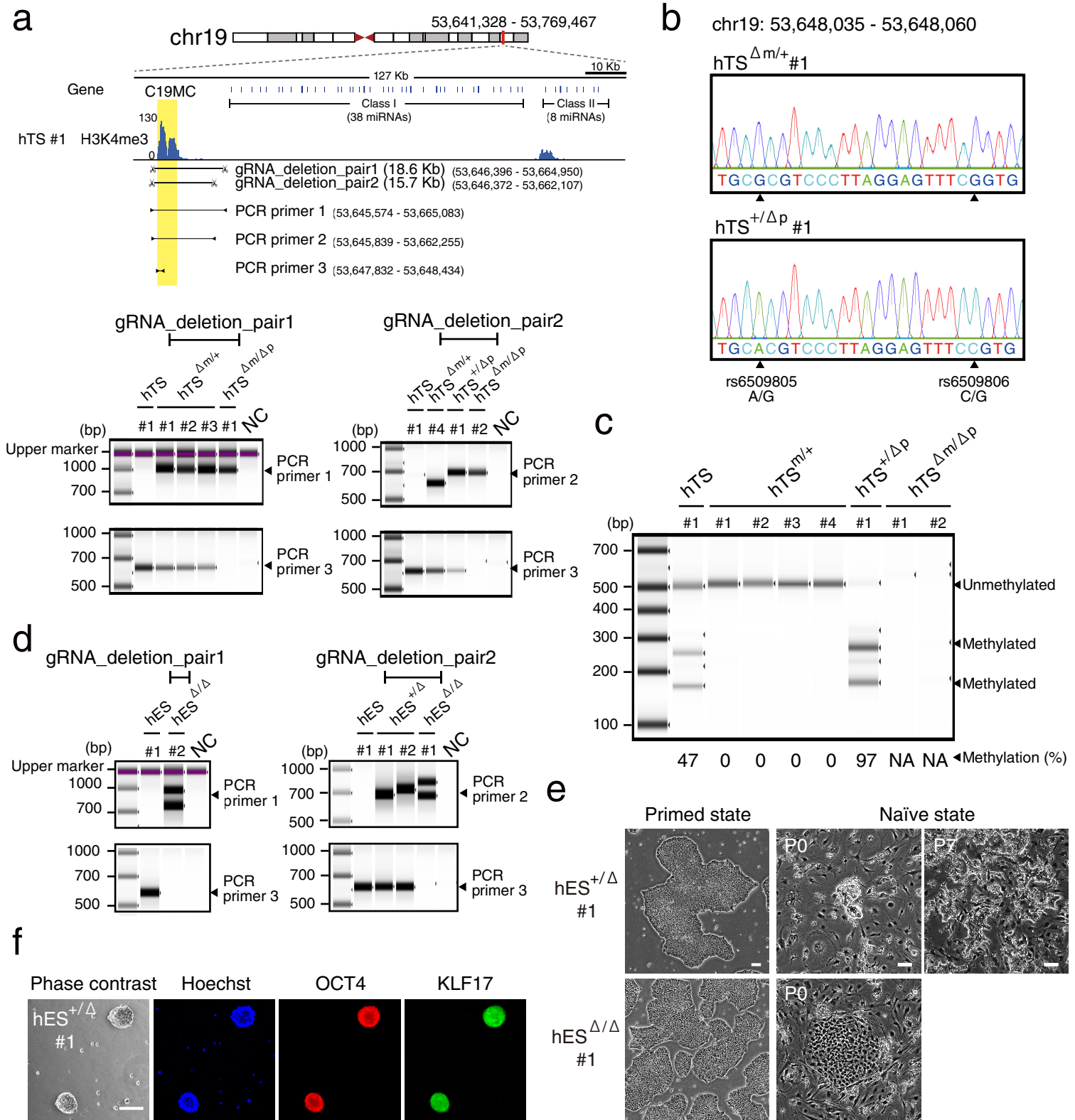
c



Supplementary Figure 3. Expression levels of imprinted genes in hTSL cells.

(a) Heatmap representation of expression levels of imprinted genes in hES, hTS, and hTSL cells. Note that more than one imprinted gene is shown for the KvDMR1, *H19* DMR, and IG-DMR. (b) DNA methylation analysis of the C19MC DMR by bisulfite sequencing. Black and white circles indicate methylated and unmethylated CpGs, respectively. Methylation levels are shown on the right. (c) DNA methylation analysis of the C19MC DMR by COBRA. The bisulfite-PCR amplicons were digested with HhaI. The C19MC DMR is imprinted in hTS#1 cells⁵ and these cells were used as control.

Supplementary Fig. 4

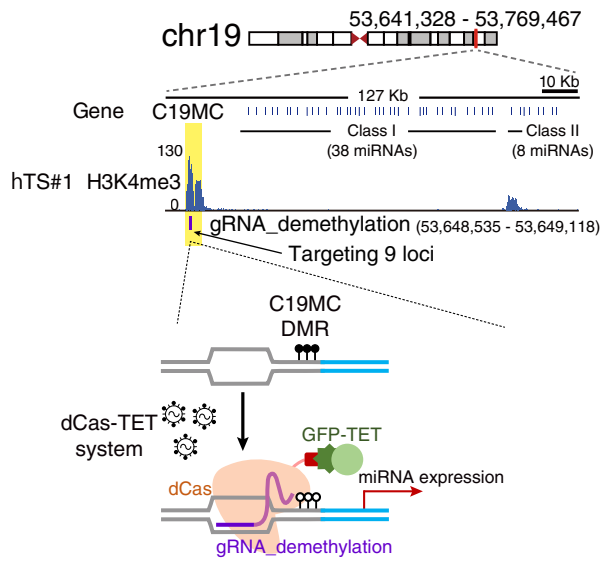


Supplementary Figure 4. Confirmation of the deletion of the C19MC DMR.

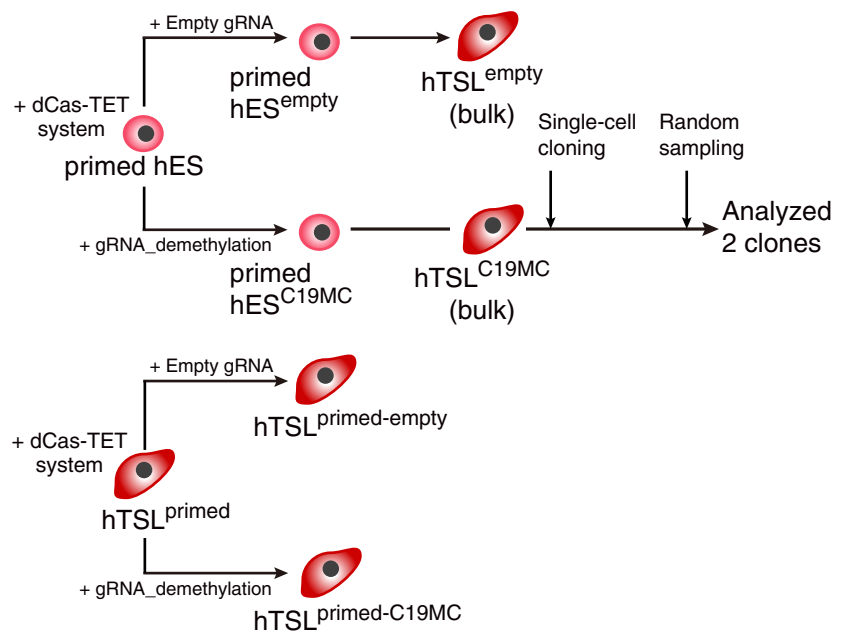
(a) Confirmation of the deletion of the C19MC DMR by genomic PCR. The C19MC DMR was deleted in hTS cells using two pairs of gRNAs (gRNA_deletion_pair1 and 2). Deletions were detected by genomic PCR. The position of three PCR primer sets is indicated. Three maternal and one homozygous deletion clone were obtained using gRNA_deletion_pair1. One maternal, one paternal, and one homozygous deletion clone were isolated using gRNA_deletion_pair2. (b) Determination of the deleted alleles using Sanger sequencing in heterozygous deletion clones. Genomic DNA was amplified using the PCR primer set 3. SNPs that were used to distinguish between maternal and paternal alleles are indicated by arrowheads. Four hTS^{Δm/+} and one hTS^{+/Δp} cell lines were analyzed, and representative results are shown. (c) DNA methylation analysis of the C19MC DMR by COBRA. The bisulfite-PCR amplicons were digested with HhaI. Genetically unmodified hTS cells were used as control. NA: not applicable. (d) Confirmation of the deletion of the C19MC DMR in primed hES cells by genomic PCR. The C19MC DMR was deleted in primed hES cells using two pairs of gRNAs (gRNA_deletion_pair1 and 2). We obtained two heterozygous (hES^{+/Δ}) and two homozygous (hES^{Δ/Δ}) deletion clones. (e) Derivation of naïve hES cells from primed hES^{+/Δ} and hES^{Δ/Δ} cells using 5i/L/A medium. Naïve hES cells were successfully derived from primed hES^{+/Δ} cells. However, primed hES^{Δ/Δ} cells formed flattened colonies at P0, and these colonies failed to proliferate after passaging. Similar results were obtained with each of the two independent hES^{+/Δ} and hES^{Δ/Δ} clones. The scale bar indicates 100 μm. (f) Immunostaining of a pluripotency marker (OCT4) and a naïve hES cell-specific marker (KLF17) in naïve hES^{+/Δ} cells. Nuclei were stained with Hoechst 33258. The scale bar indicates 100 μm. Similar results were obtained with two independent hES^{+/Δ} clones.

Supplementary Fig. 5

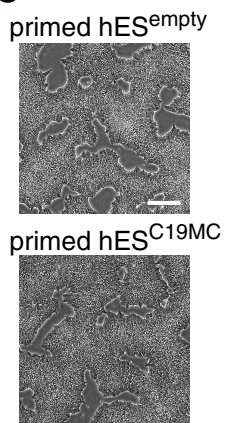
a



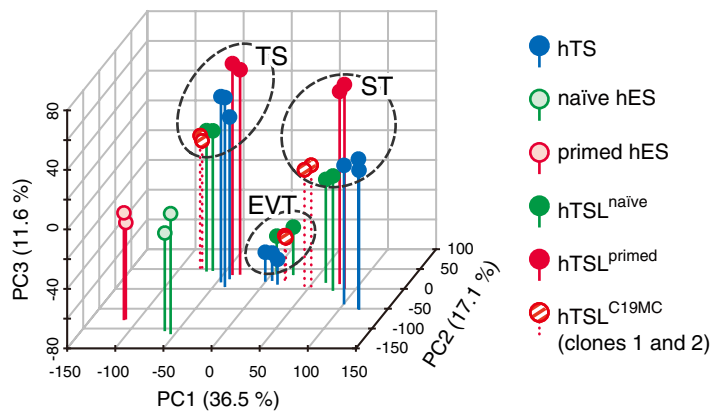
b



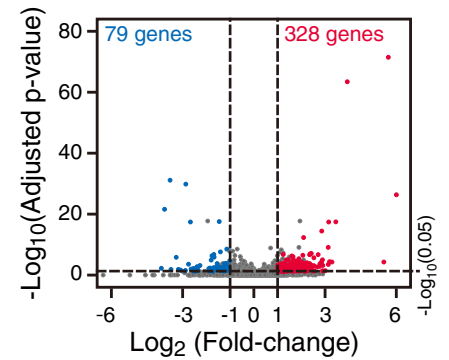
c



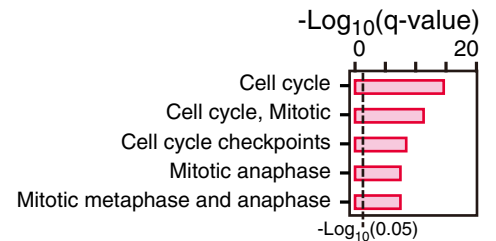
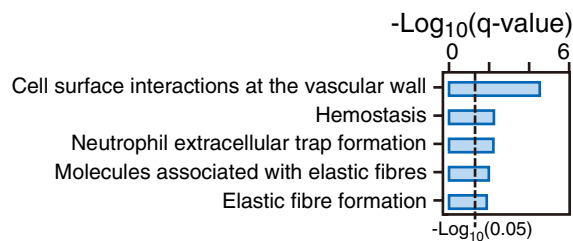
d



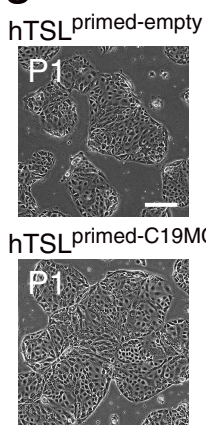
e



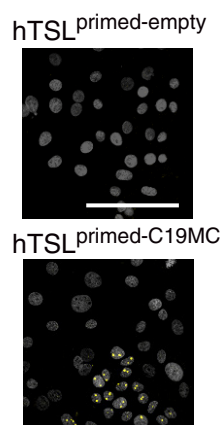
f



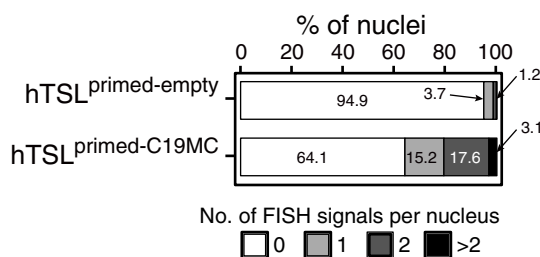
g



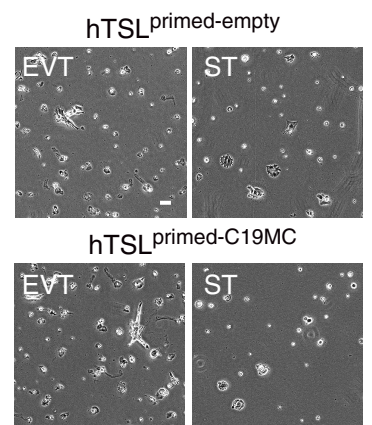
h



i



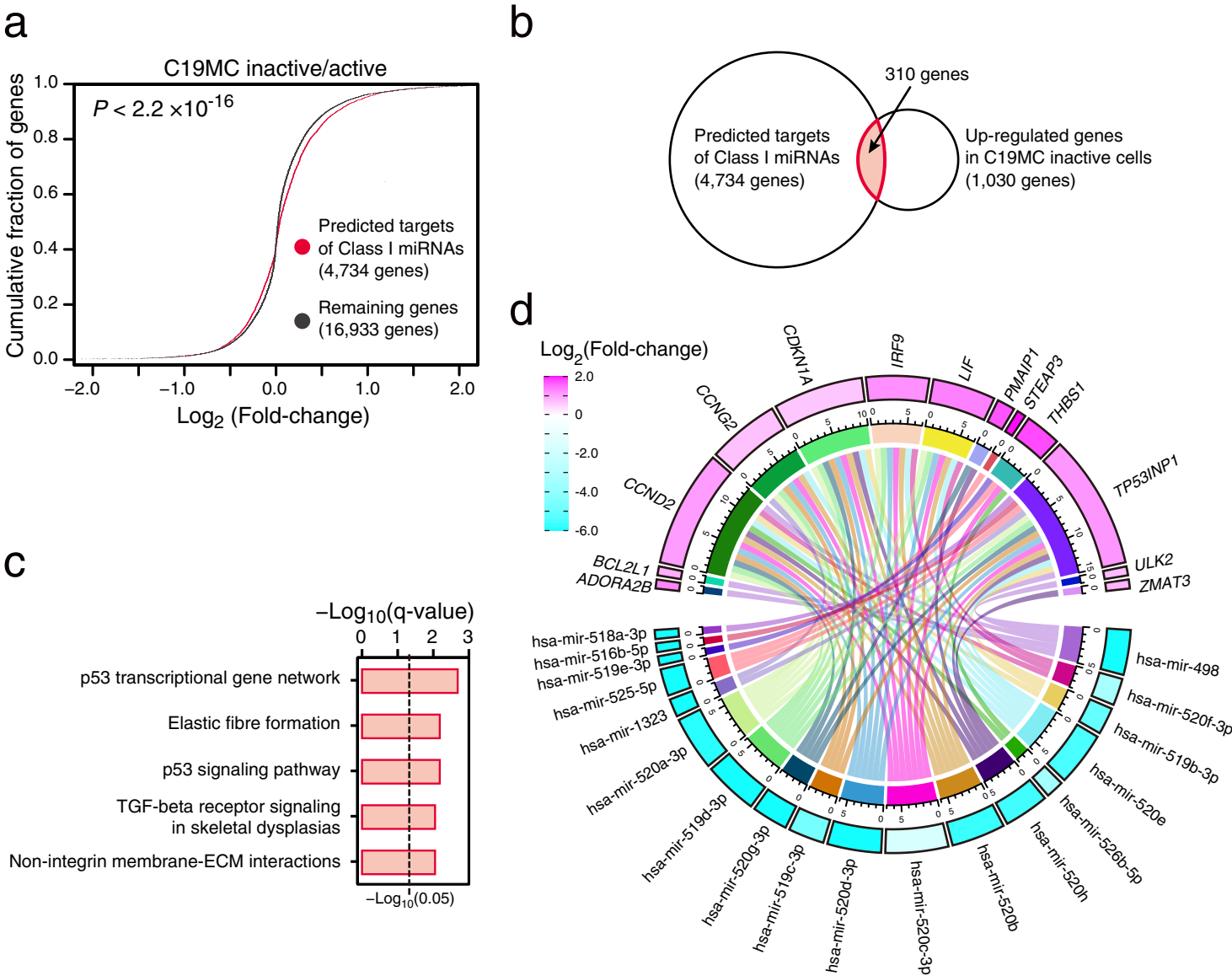
j



Supplementary Figure 5. Reactivation of C19MC in primed hES and hTSL^{primed} cells.

(a) Schematic representation of C19MC reactivation using the dCas-TET system. The catalytically inactive Cas9 (dCas) is fused to a peptide repeat that can recruit scFv-fused GFP and the catalytic domain of TET1. We designed a gRNA (gRNA_{demethylation}) targeting the short tandem repeats within the C19MC DMR. The C19MC DMR contains 9 target sites for this gRNA. By combining these components, we demethylated the C19MC DMR. (b) Schematic representation of the derivation of hTSL^{C19MC} and hTSL^{primed-C19MC} cells. We delivered the gRNA and the dCas-TET expression vectors into a primed hES cells and derived hTSL^{C19MC} cells. We further isolated two C19MC-positive clones from these bulk hTSL^{C19MC} cells by single-cell cloning. As a negative control, we used an empty gRNA vector and derived hTSL^{empty} cells. hTSL^{primed-C19MC} cells were generated from hTSL^{primed} cells that were originated from genetically unmodified primed hES cells. (c) Phase-contrast images of C19MC-reactivated hES cells. Similar results were obtained with two independent experiments. The scale bar indicates 300 μ m. (d) PCA of hTS and hTSL cells. hTSL^{C19MC} cells had transcriptome profiles similar to hTS and hTSL^{naïve} cells both before and after differentiation. Only 12,853 differentially expressed genes (TPM > 4 in at least one cell type, adjusted p-value < 0.05, fold change > 2) were analyzed. (e) Volcano plot comparing ST-hTSL^{C19MC} cells with ST-hTS and -hTSL^{naïve} cells. Gene expression levels were compared and expressed as log2 fold change. The down- and up-regulated genes are shown in blue and red dots, respectively (TPM > 4 in at least one cell type, adjusted p-value < 0.05, fold change > 2). (f) Pathways enriched among the genes down- and up-regulated in ST-hTSL^{C19MC} cells (blue and red dots in Supplementary Fig. 5e). The top five pathways of the down- and up-regulated genes are represented with q-values on the left and right charts, respectively. (g) Phase-contrast images of hTSL^{primed-empty} and hTSL^{primed-C19MC} cells. Similar results were obtained with two independent experiments. The scale bar indicates 300 μ m. (h) RNA-FISH of C19MC in hTSL^{primed-empty} and hTSL^{primed-C19MC} cells. C19MC expression is shown in yellow, and nuclei are in gray. The scale bar indicates 100 μ m. (i) Percentage of the number of C19MC FISH signals per nucleus in hTSL^{primed-empty} and hTSL^{primed-C19MC} cells. 500-1000 nuclei were examined for each cell type. (j) Phase-contrast images of EVT- and ST-like cells derived from hTSL^{primed-empty} and hTSL^{primed-C19MC} cells. Similar results were obtained with two independent experiments. The scale bar indicates 100 μ m.

Supplementary Fig. 6



Supplementary Figure 6. Prediction of genes targeted by C19MC miRNAs.

(a) Potential target genes of Class I miRNAs located between the C19MC DMR and the putative second promoter. Target genes were predicted using mirDIP⁸ by comparing the gene expression levels in C19MC active (hTS, hTSL^{naïve}, hTS^{Δm/+}, and hTSL^{C19MC}) cells with those in C19MC inactive (hTSL^{primed}, hTS^{+/ Δ p}, hTS^{Δm/ Δ p}, and hTSL^{empty}) cells. The data are expressed as the log2 fold change. Up- and down-regulated genes have positive and negative values, respectively. We considered only genes with Transcripts Per Million (TPM) values greater than zero in at least one cell type. Statistical analysis was performed by two-sided Kolmogorov-Smirnov test. (b) Venn diagram showing the most likely target genes of Class I miRNAs. The gene expression levels in C19MC inactive (hTSL^{primed}, hTS^{+/ Δ p}, hTS^{Δm/ Δ p}, and hTSL^{empty}) cells were compared with those in C19MC active (hTS, hTSL^{naïve}, hTS^{Δm/+} and hTSL^{C19MC}) cells, and 1,030 up-regulated genes were defined as follows: TPM > 4 in at least one cell type, adjusted p-value < 0.05, and fold change > 1.5. By merging these up-regulated genes with the computationally predicted target genes in (a), we identified 310 genes as the most likely candidates. (c) Overrepresented pathways among the most likely target genes of Class I miRNAs. The top five pathways are represented with q-values. (d) Circos plot of the miRNA-mRNA combinations associated with “p53 transcriptional gene network” and “p53 signaling pathway”. The expression levels in C19MC active cells were compared to those in C19MC inactive cells and expressed as the log2 fold change.

Supplementary Table 1. List of the antibodies used in this study.

Description	Company	Catalog number	Dilution rate (Usage)
Primary antibody			
Anti-TP63	Cell Signaling	#13109	1:100 (Immunostaining)
Anti-TFAP2C	Santa Cruz Biotechnology	#sc-12762	1:200 (Immunostaining)
Ant-KRT7	Abcam	#ab119697	1:100 (Flow cytometry)
Anti-hCG	DAKO	#IR508	1:10 (Immunostaining)
Anti-KLF17	Sigma	#HPA024629	1:200 (Immunostaining)
Anti-H3K4me3	MBL	#MAB10304	0.8 ng/ul (ChIP)
Anti-ELF5	Sigma	#HPA062706	1:100 (Immunostaining)
Anti-TEAD4	Abcam	#ab58310	1:100 (Immunostaining)
Anti-CDX2	Abcam	#ab157524	1:100 (Immunostaining)
PE-conjugated anti-OCT-4A	Cell Signaling	#56159	1:800 (Immunostaining)
PE-conjugated anti-HLA-G	Abcam	#ab24384	1:50 (Flow cytometry)
PE-conjugated anti-SDC1	Miltenyi Biotec	#130-119-928	1:500 (Immunostaining)
Alexa Fluor 488-conjugated anti-HLA-ABC	Biolegend	#311415	1:50 (Flow cytometry)
Secondary antibody			
PE-conjugated mouse IgG1	R&D	#IC006P	
Alexa Fluor 488-conjugated anti-rabbit IgG	Cell Signaling	#4412	
Alexa Fluor 488-conjugated anti-mouse IgG2a	Biolegend	#400233	
Alexa Fluor 555-conjugated anti-mouse IgG	Cell Signaling	#4409	
Alexa Fluor 647-conjugated anti-rabbit IgG	Cell Signaling	#4414	
HRP-conjugated anti-rabbit IgG	Cell Signaling	#7074	

Supplementary References

1. Bernardo, A. S., et al. BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. *Cell Stem Cell* **9**, 144-155 (2011).
2. Cinkornpumin, J. K., et al. Naive Human Embryonic Stem Cells Can Give Rise to Cells with a Trophoblast-like Transcriptome and Methylome. *Stem Cell Rep.* **15**, 198-213 (2020).
3. Guo, G., et al. Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* **28**, 1040-1056 e1046 (2021).
4. Bai, Q., et al. Dissecting the first transcriptional divergence during human embryonic development. *Stem Cell Rev. Rep.* **8**, 150-162 (2012).
5. Okae, H., et al. Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* **22**, 50-63 e56 (2018).
6. Pastor, W. A., et al. Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell* **18**, 323-329 (2016).
7. Bernstein, B. E., et al. The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.* **28**, 1045-1048 (2010).
8. Tokar, T., et al. mirDIP 4.1-integrative database of human microRNA target predictions. *Nucleic Acids Res.* **46**, D360-D370 (2018).