The American Journal of Human Genetics, Volume *99*

Supplemental Data

Allele-Specific Methylome and Transcriptome Analysis Reveals Widespread Imprinting in the Human Placenta

Hirotaka Hamada, Hiroaki Okae, Hidehiro Toh, Hatsune Chiba, Hitoshi Hiura, Kenjiro Shirane, Tetsuya Sato, Mikita Suyama, Nobuo Yaegashi, Hiroyuki Sasaki, and Takahiro Arima

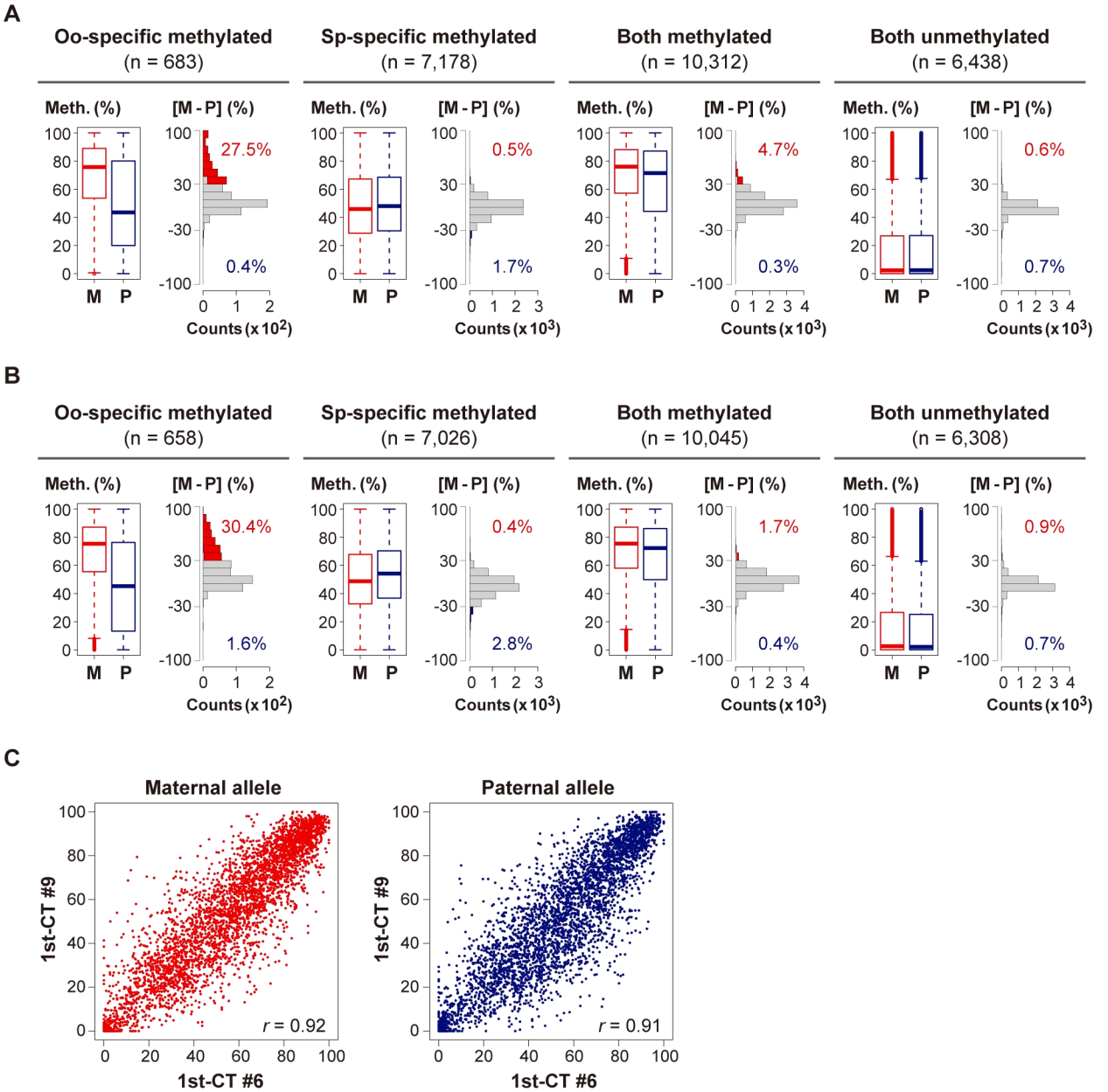

A
Oo-specific methylated
(n = 683)
Sp-specific methylated
(n = 7,178)
Both methylated
(n = 10,312)
Both unmethylated
(n = 6,438)
Meth. (%)
[M - P] (%)
27.5%
0.4%
0.5%
1.7%
4.7%
0.3%
0.6%
0.7%
M
P
Counts (x 10^2)
Counts (x 10^3)
B
Oo-specific methylated
(n = 658)
Sp-specific methylated
(n = 7,026)
Both methylated
(n = 10,045)
Both unmethylated
(n = 6,308)
30.4%
1.6%
0.4%
2.8%
1.7%
0.4%
0.9%
0.7%
C
Maternal allele
Paternal allele
1st-CT #9
1st-CT #6
r = 0.92
r = 0.91

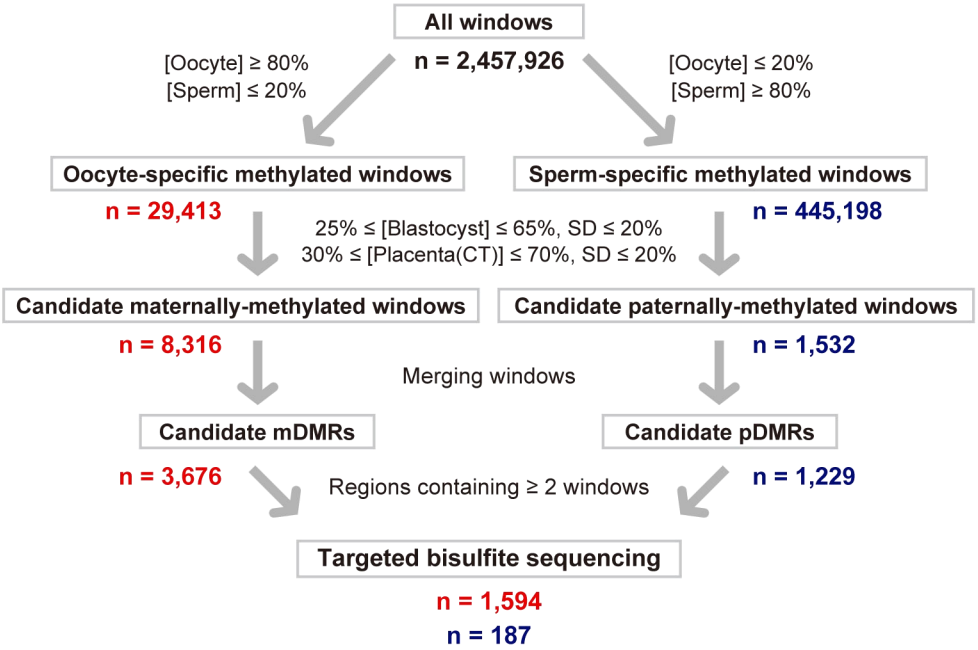
Figure S1. Maintenance of allelic DNA methylation of oocyte-specific methylated windows

(A) Box plots of methylation levels of the maternal (M) and paternal (P) alleles of windows (1st-CT #9). Boxes represent lower and upper quartiles and horizontal lines indicate the median. Whiskers extend to the most extreme data points within 1.5 times the interquartile range from the boxes. The open circles indicate the data points outside the whiskers. Histograms of the distribution of the [M-P] levels are also shown. 27.5% of oocyte-specific methylated windows were maternally methylated.

(B) Box plots of methylation levels of the maternal (M) and paternal (P) alleles of windows (1st-CT #6: the sample analyzed in Figure 1). For comparison, an equal number of reads of 1st-CT#9 were randomly extracted from the reads of 1st-CT #6. Histograms of the distribution of the [M-P] levels are also shown.

(C) Comparison of allelic methylation levels between 1st-CT #6 and 1st-CT #9. As described above, we used randomly extracted reads for 1st-CT #6. 4,307 windows were covered by both 1st-CT #6 and 1st-CT #9, and were analyzed. The allelic methylation levels were strongly correlated between these two samples (Pearson's $r > 0.9$).

A



B

Present study

[Oocyte]	[Sperm]	[Blastocyst]	[Placenta]	[M-P]
≥ 80%	≤ 20%	25-65%	30-70%	≥ 30%

Hanna et al.

[Oocyte - Sperm]	[Blastocyst]	[M-P]
> 50%	15-60%	> 15%*

Figure S2. Selection of candidate gDMRs for targeted bisulfite sequencing

(A) We focused on oocyte- and sperm-specific methylated windows maintaining 25-65% methylation levels with ≤ 20% standard deviations (SD) in blastocysts and 30-70% methylation levels with ≤ 20% SD in CT cells (the WGBS data of 1st-CT #6 was used). After merging windows within 500 bp of each other, we obtained 3,676 candidate mDMRs and 1,229 pDMRs except for known gDMRs. 1,594 candidate mDMRs and 187 candidate pDMRs containing ≥ 2 windows were selected for targeted bisulfite sequencing analyses.

(B) Comparison of the cutoff values for mDMRs. Our cutoff values were more stringent than those of Hanna et al. [1]. *: the original criterion is "> 5% difference in methylation between diandric and digynic triploids", which corresponds to > 15% [M-P].

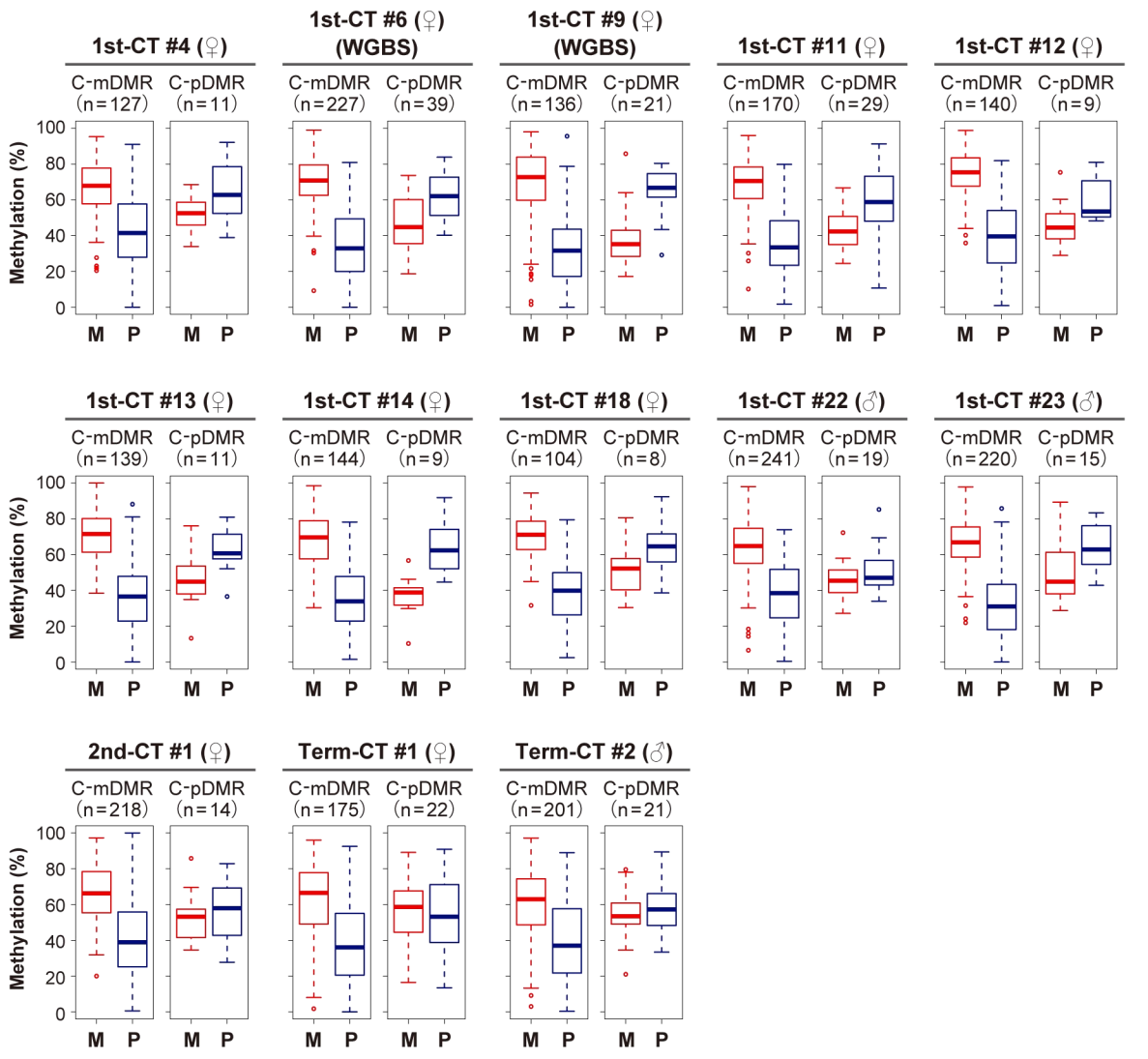

1st-CT #4 (♀)
C-mDMR (n=127)
C-pDMR (n=11)
1st-CT #6 (♀) (WGBS)
C-mDMR (n=227)
C-pDMR (n=39)
1st-CT #9 (♀) (WGBS)
C-mDMR (n=136)
C-pDMR (n=21)
1st-CT #11 (♀)
C-mDMR (n=170)
C-pDMR (n=29)
1st-CT #12 (♀)
C-mDMR (n=140)
C-pDMR (n=9)
1st-CT #13 (♀)
C-mDMR (n=139)
C-pDMR (n=11)
1st-CT #14 (♀)
C-mDMR (n=144)
C-pDMR (n=9)
1st-CT #18 (♀)
C-mDMR (n=104)
C-pDMR (n=8)
1st-CT #22 (♂)
C-mDMR (n=241)
C-pDMR (n=19)
1st-CT #23 (♂)
C-mDMR (n=220)
C-pDMR (n=15)
2nd-CT #1 (♀)
C-mDMR (n=218)
C-pDMR (n=14)
Term-CT #1 (♀)
C-mDMR (n=175)
C-pDMR (n=22)
Term-CT #2 (♂)
C-mDMR (n=201)
C-pDMR (n=21)
Methylation (%)
0
20
40
60
80
100
M
P

Figure S3. Allelic DNA methylation levels in each CT sample

Box plots of methylation levels of the maternal (M) and paternal (P) alleles of candidate gDMRs. Boxes represent lower and upper quartiles and horizontal lines indicate the median. Whiskers extend to the most extreme data points within 1.5 times the interquartile range from the boxes. The open circles indicate the data points outside the whiskers. Ten female 1st-CT samples and three 2nd/term-CT samples were analyzed. 1st-CT #6 and #9 were analyzed using WGBS and the other samples were analyzed using targeted bisulfite sequencing.

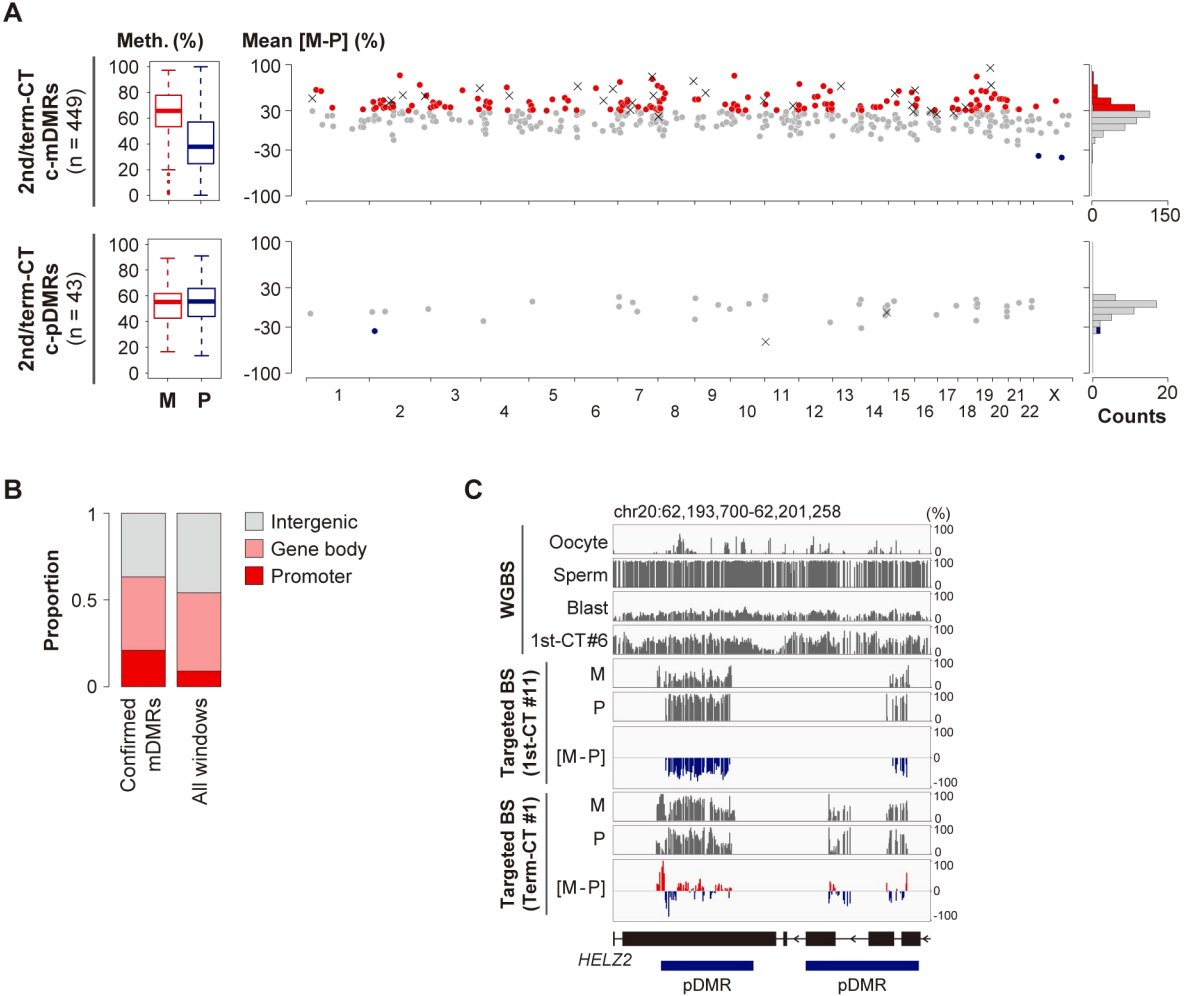

A
Meth. (%)
Mean [M-P] (%)
2nd/term-CT c-mDMRs (n = 449)
2nd/term-CT c-pDMRs (n = 43)
M
P
Counts
B
Proportion
Intergenic
Gene body
Promoter
Confirmed mDMRs
All windows
C
chr20:62,193,700-62,201,258
(%)
WGBS
Oocyte
Sperm
Blast
1st-CT#6
Targeted BS (1st-CT #11)
Targeted BS (Term-CT #1)
[M-P]
HELZ2
pDMR
pDMR

Figure S4. Distribution of gDMRs in 2nd/term CT samples

(A) Allelic DNA methylation patterns of candidate DMRs (c-DMRs) in 2nd/term-CT samples. Data from three placental samples were combined. Box plots of methylation levels of the maternal (M) and paternal (P) alleles, chromosomal distribution of [M-P] levels and histograms of the distribution of the [M-P] levels are shown. In the chromosome maps, red circles indicate c-DMRs showing ≥ 30% [M-P] levels and statistically significant allelic methylation differences (BH-corrected $P < 0.05$). Similarly, blue circles indicate those with ≤ -30% [M-P] levels. The other c-DMRs are shown as gray circles. Cross marks (×) represent known gDMRs.

(B) Distribution of the confirmed mDMRs and all 20 CpG windows. Compared to the 20 CpG windows, the confirmed mDMRs were significantly enriched at promoter regions ($P = 1.0 \times 10^{-15}$: Chi-square test).

(C) DNA methylation patterns of pDMRs located in the gene body region of *HELZ2* [MIM: 611265]. The pDMRs were paternally methylated in 1st-CT cells, but the paternally biased methylation was lost in term-CT cells.

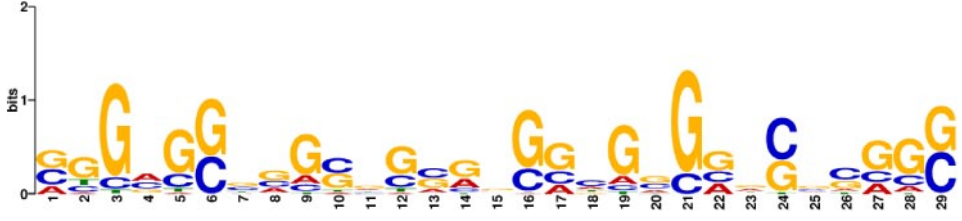
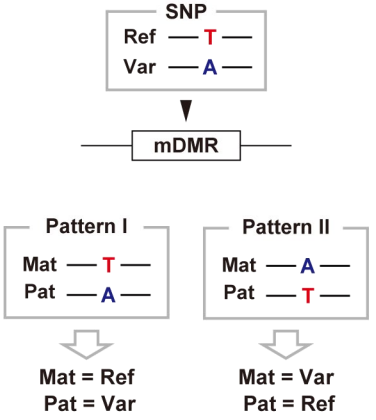
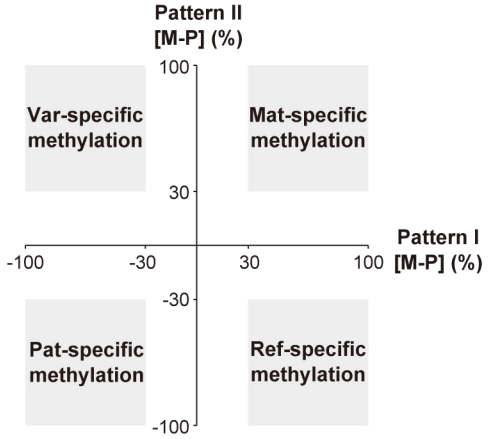
Motif	*E*-value	% of targets	Similar motifs
	3.4e-88	36.8%	EGR1 SP2 SP1

Figure S5. Search for sequence motifs in the confirmed mDMRs

We used MEME-ChIP[2] for the motif search. The confirmed mDMRs (n = 440) were analyzed. We considered only statistically significant (*E*-value < 0.05) motifs that were found in > 20% of the confirmed mDMRs. The three most similar motifs are also indicated.

A


SNP
Ref T
Var A
mDMR
Pattern I
Mat T
Pat A
Mat = Ref
Pat = Var
Pattern II
Mat A
Pat T
Mat = Var
Pat = Ref


Pattern II
[M-P] (%)
100
30
-30
-100
-100
-30
30
100
Pattern I
[M-P] (%)
Var-specific
methylation
Mat-specific
methylation
Pat-specific
methylation
Ref-specific
methylation

B

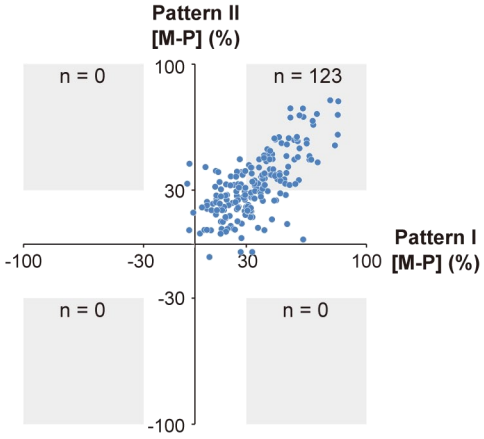

Pattern II
[M-P] (%)
n = 0
n = 123
n = 0
n = 0
100
30
-30
-100
-100
-30
30
100
Pattern I
[M-P] (%)

Figure S6. Haplotype-dependent allele-specific methylation (hap-ASM) analysis

(A) Schematic representation of the hap-ASM analysis. For each mDMR, CT samples were classified into two groups according to the SNP sequences (exemplified by T/A). Samples whose maternal alleles are the reference alleles are classified as Pattern I, and those whose maternal alleles are the variant alleles are classified as Pattern II. Parent-of-origin-dependent allele-specific methylation (Mat- or Pat-specific methylation) and hap-ASM (Ref- or Var-specific methylation) can be distinguished by comparing the mean [M-P] levels of Pattern I and II samples. Ref: reference allele; Var: variant allele; Mat: maternal allele; Pat: paternal allele.

(B) Hap-ASM analysis of candidate mDMRs. The data of the 1st-CT samples were used. In our data set, 306 SNPs overlapping 149 candidate mDMRs (including those that were not confirmed to be imprinted) were available for the hap-ASM analysis. The numbers of SNPs associated with parent-of-origin-dependent allele-specific methylation and hap-ASM are indicated. 123 SNPs overlapping 65 candidate mDMRs were confirmed to be associated with parent-of-origin-dependent allele-specific methylation. No region was found to show hap-ASM.

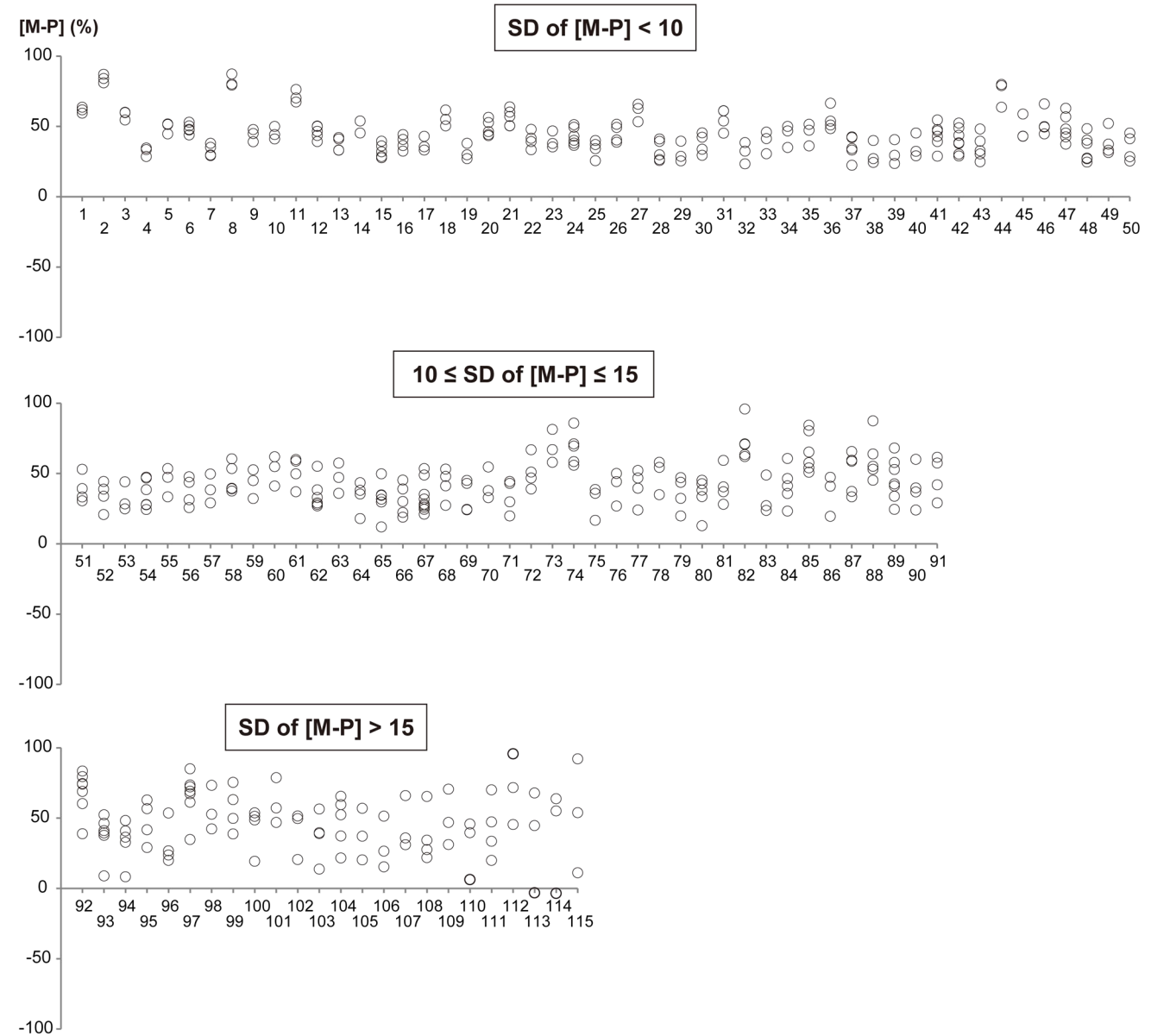
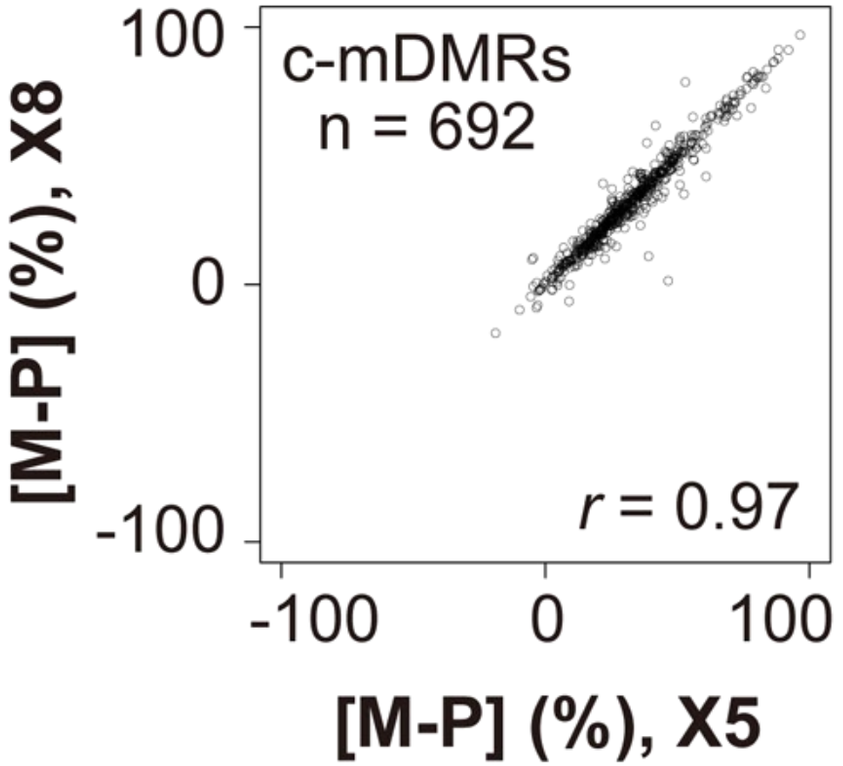
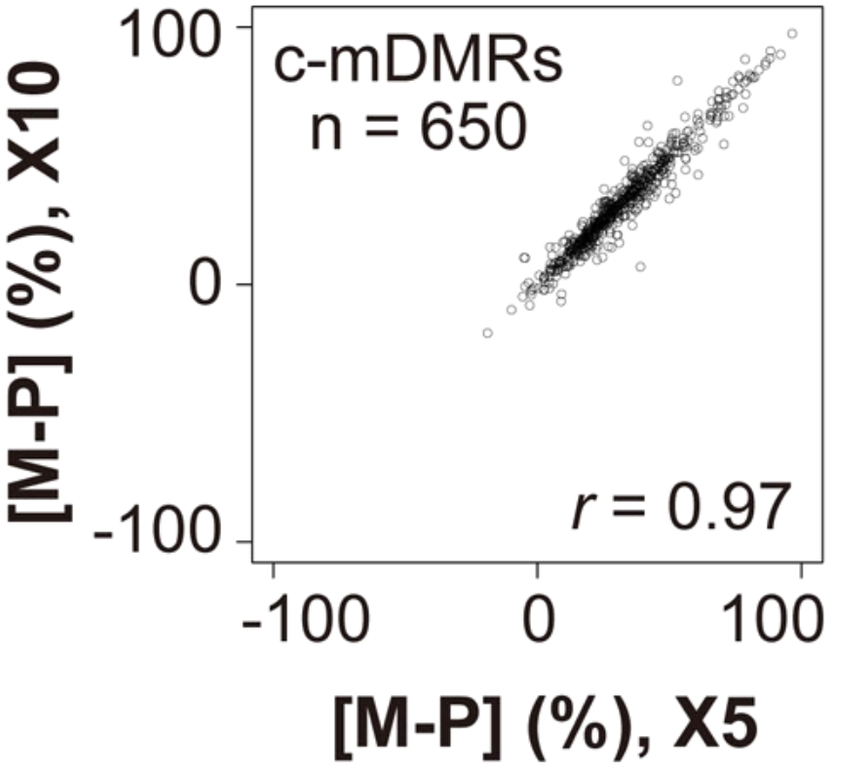
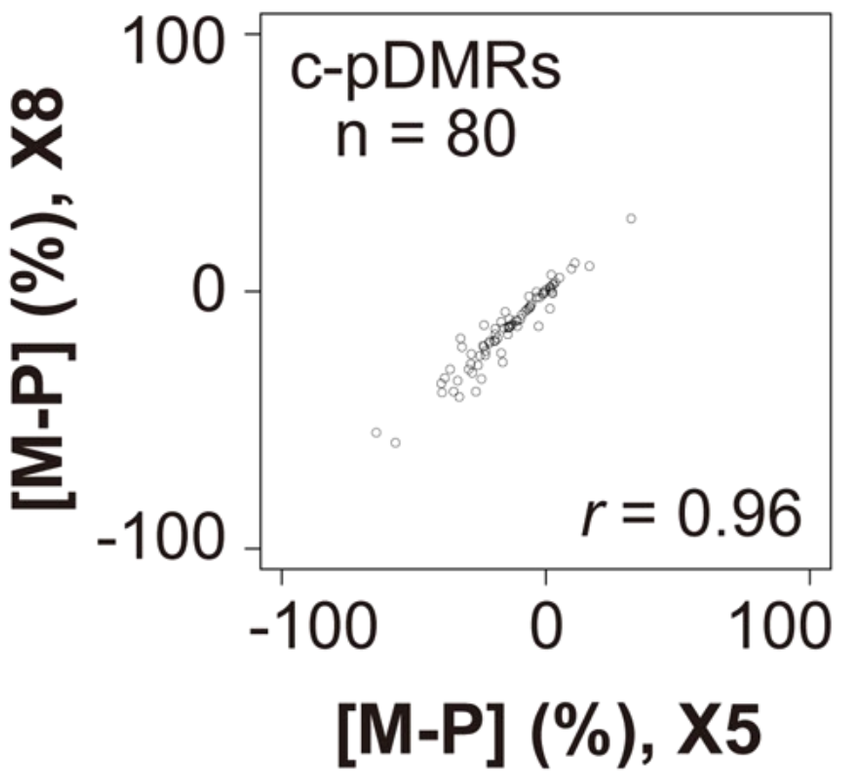

[M-P] (%)
SD of [M-P] < 10
100
50
0
-50
-100
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
10 ≤ SD of [M-P] ≤ 15
100
50
0
-50
-100
51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91
SD of [M-P] > 15
100
50
0
-50
-100
92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115

Figure S7. Variations in allelic methylation levels between samples

Confirmed mDMRs for which there were data from three or more samples are shown. The SD value for each confirmed mDMR was calculated as shown in Figure 2D, and the confirmed mDMRs were numbered according to the SD values. Some confirmed mDMRs with > 15% SD values showed polymorphic imprinting (e.g. No. 113-115). In contrast, the allelic methylation levels of mDMRs with < 10% SD values (No. 1-50) were relatively consistent across samples.


c-mDMRs
n = 692
r = 0.97
100
0
-100
-100
0
100
[M-P] (%), X8
[M-P] (%), X5


c-mDMRs
n = 650
r = 0.97
100
0
-100
-100
0
100
[M-P] (%), X10
[M-P] (%), X5


c-pDMRs
n = 80
r = 0.96
100
0
-100
-100
0
100
[M-P] (%), X8
[M-P] (%), X5

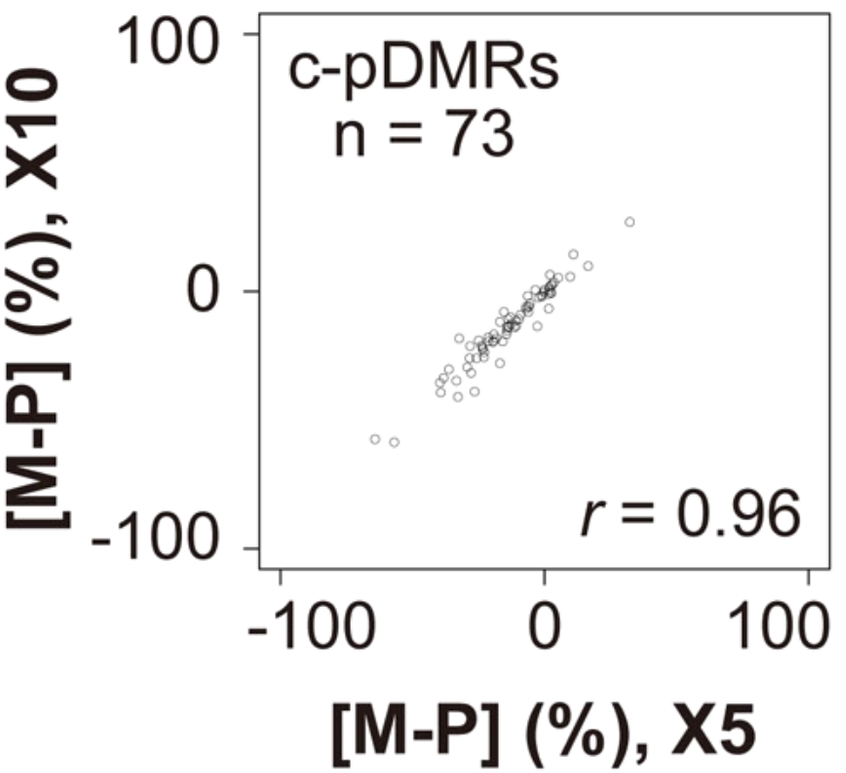
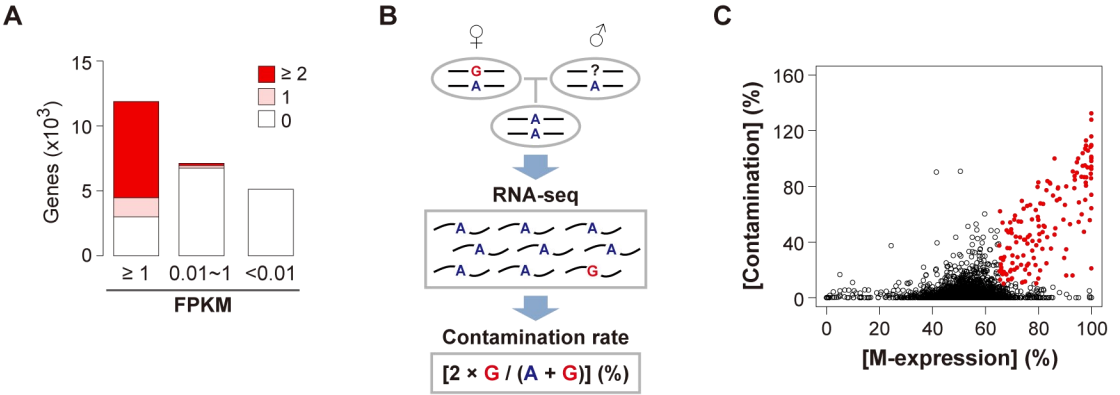

c-pDMRs
n = 73
r = 0.96
100
0
-100
-100
0
100
[M-P] (%), X10
[M-P] (%), X5

Figure S8. Comparison of allelic methylation levels obtained at 5X, 8X and 10X coverage

Allelic methylation levels of candidate mDMRs were calculated using CpGs covered with ≥ 5 reads (5X), ≥ 8 reads (8X) and ≥ 10 reads (10X). The data of the 1st-CT samples were used. The 8X and 10 data highly correlated with the 5X data ($r > 0.95$).



D

GO term	Count	Corrected *P*-value
inflammatory response	21	6.6×10^{-9}
response to wounding	26	9.1×10^{-9}
immune response	28	2.8×10^{-8}

Figure S9. Allelic expression analysis and identification of genes highly expressed in contaminating maternal cells

(A) Counts of autosomal genes covered by our allelic expression analyses. We analyzed 21 1st-CT samples (Table S1). Genes covered by ≥ 2 samples are shown in red, genes covered by one sample are in thin red and uncovered genes are in white. ~75% of genes with ≥ 1 FPKM were successfully analyzed in at least one sample but most genes with < 1 FPKM were not.

(B) Estimation of maternal cell contamination. We focused on SNPs that were homozygous in CT samples (exemplified by A/A) and heterozygous in mothers (exemplified by A/G). By doubling the rate of reads containing the non-embryonic SNP ("G" in this figure), we estimated the expression rate from the contaminating maternal cells ([Contamination] rate) for each gene.

(C) The relationship between the [M-expression] ratio and [Contamination] rate. Genes with > 10% mean [Contamination] rates and > 65% [M-expression] ratios are shown in red. Genes with > 10% [Contamination] rates tended to show maternally biased expression.

(D) GO analysis of genes with > 10% mean [Contamination] rates and > 65% [M-expression] ratios (n = 154). The top three GO terms are indicated with gene counts and BH-corrected *P*-values. Immune-related GO terms were significantly enriched.

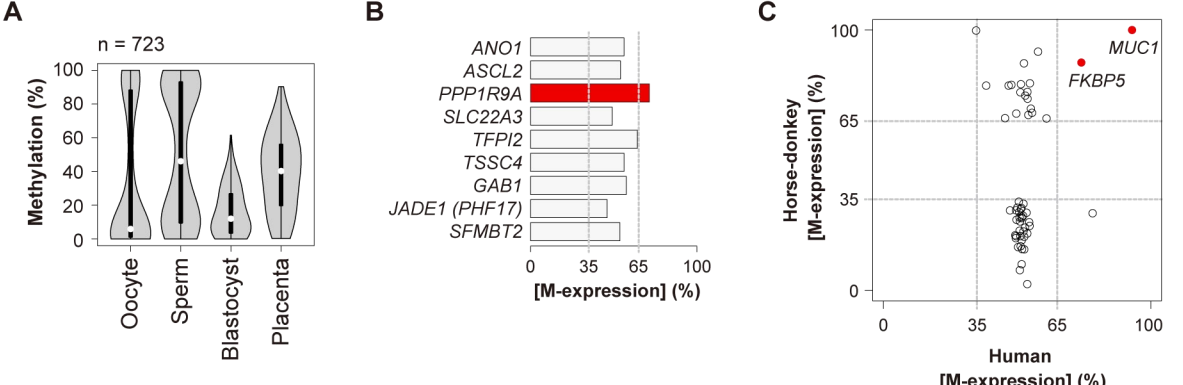

A
n = 723
Methylation (%)
100
80
60
40
20
0
Oocyte
Sperm
Blastocyst
Placenta
B
ANO1
ASCL2
PPP1R9A
SLC22A3
TFPI2
TSSC4
GAB1
JADE1 (PHF17)
SFMBT2
0
35
65
100
[M-expression] (%)
C
100
65
35
0
MUC1
FKBP5
Horse-donkey
[M-expression] (%)
0
35
65
100
Human
[M-expression] (%)

Figure S10. Low conservation of gDMRs and imprinted genes in mammalian placentas

(A) DNA methylation patterns of regions homologous to human candidate mDMRs in mouse oocytes, sperm, blastocysts and placenta. We used available WGBS data of mouse oocytes (DRA000570),[3] sperm, blastocysts (DRA000484)[4] and placenta (GSM1051161).[5] We selected regions having ≥ 80% methylation in oocytes, ≤ 20% methylation in sperm, 25-65% methylation with ≤ 20% SD in blastocysts and 30-70% methylation with ≤ 20% SD in the mouse placenta as candidate mDMRs. Only 5 regions were found to have the imprinted methylation patterns, suggesting that most human placental mDMRs may not be mDMRs in the mouse placenta.

(B) Allelic expression of human homologs of mouse placenta-specific imprinted genes. Mouse placenta-specific imprinted genes are from the Catalogue of Parental Origin Effects database. Only *PPP1R9A* showed maternal expression (> 65% [M-expression] ratio) in human 1st-CT cells.

(C) Allelic expression of candidate imprinted genes (n = 78) identified in the placenta of hybrids of the horse and donkey. The x-axis and y-axis show mean [M-expression] ratios in human 1st-CT cells and the placenta of the hybrids, respectively. These candidate imprinted genes did not have conserved imprinted expression patterns in the human placenta except for two genes, *MUC1* and *FKBP5*.

Table S1. CT samples used for bisulfite sequencing and RNA sequencing

The number of uniquely mapped reads and mean coverage per CpG site are shown.

Table S2. Chromosomal location of candidate and known gDMRs

Regions selected for the targeted bisulfite sequencing are indicated. Names of the known gDMRs are also indicated.

Table S3. Allelic DNA methylation levels of candidate gDMRs in 1st-CT samples

Ten 1st-CT samples were analyzed. For each gDMR, the methylation levels of the maternal (M) and paternal (P) alleles and the [M-P] level are shown. The candidate mDMRs and pDMRs were respectively ranked according to their [M-P] levels. The statistical significance (BH-corrected *P*-value) of the methylation differences between parental alleles is shown. Associated genes are also indicated. All informative candidate gDMRs including those that were not confirmed to be imprinted are indicated. NA: Not available.

Table S4. Allelic DNA methylation levels of candidate gDMRs in 2nd/term-CT samples

Three 2nd/term-CT samples were analyzed. For each gDMR, the methylation levels of the maternal (M) and paternal (P) alleles and the [M-P] level are shown. The candidate mDMRs and pDMRs were respectively ranked according to their [M-P] levels. The statistical significance (BH-corrected *P*-value) of the methylation differences between parental alleles is shown. Associated genes are also indicated. All informative candidate gDMRs including those that were not confirmed to be imprinted are indicated. NA: Not available.

Table S5. Allelic DNA methylation levels of known gDMRs in 1st-CT samples

Known gDMRs are classified as described in Materials and Methods. For each gDMR, the methylation levels of the maternal (M) and paternal (P) alleles and the [M-P] level are shown. The statistical significance (BH-corrected *P*-value) of the methylation differences between parental alleles is also indicated. NA: Not available.

Table S6. Allelic DNA methylation levels of known gDMRs in 2nd/term-CT samples

Known gDMRs are classified as described in Materials and Methods. For each gDMR, the methylation levels of the maternal (M) and paternal (P) alleles and the [M-P] level are shown. The statistical significance (BH-corrected *P*-value) of the methylation differences between parental alleles is also indicated. NA: Not available.

Table S7. Allelic expression of autosomal genes

For each gene, the proportion of reads derived from the maternal allele to total reads ([M-expression] ratio), [Contamination] rate and FPKM are shown. The statistical significance (BH-corrected *P*-value) of the allelic expression differences is also shown. NA: Not available.

Table S8. Strong candidate imprinted genes

For each gene, the mean [M-expression] ratio, BH-corrected *P*-value and expression allele are shown. The numbers of samples with maternal (< 35% [M-expression] ratios),

paternal (> 65% [M-expression] ratios) and biallelic (35-65% [M-expression] ratios) expression are also indicated. Allelic DNA methylation patterns of associated gDMRs, which were obtained using the 1st-CT samples, are also included. Maternal methylation of the *CYP2J2* DMR was previously reported.[6] NA: Not available; M: maternal; P: paternal.

Table S9. Allelic expression of X-linked genes

For each gene, the [M-expression] ratio, [Contamination] rate and FPKM are shown. Genes in the pseudoautosomal regions are also indicated. NA: Not available.

References

1. Hanna, C.W., Penaherrera, M.S., Saadeh, H., Andrews, S., McFadden, D.E., Kelsey, G., and Robinson, W.P. (2016). Pervasive polymorphic imprinted methylation in the human placenta. Genome Res.

2. Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics *27*, 1696-1697.

3. Shirane, K., Toh, H., Kobayashi, H., Miura, F., Chiba, H., Ito, T., Kono, T., and Sasaki, H. (2013). Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet *9*, e1003439.

4. Kobayashi, H., Sakurai, T., Imai, M., Takahashi, N., Fukuda, A., Yayoi, O., Sato, S., Nakabayashi, K., Hata, K., Sotomaru, Y., et al. (2012). Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genet *8*, e1002440.

5. Hon, G.C., Rajagopal, N., Shen, Y., McCleary, D.F., Yue, F., Dang, M.D., and Ren, B. (2013). Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nat Genet *45*, 1198-1206.

6. Sanchez-Delgado, M., Martin-Trujillo, A., Tayama, C., Vidal, E., Esteller, M., Iglesias-Platas, I., Deo, N., Barney, O., Maclean, K., Hata, K., et al. (2015). Absence of Maternal Methylation in Biparental Hydatidiform Moles from Women with NLRP7 Maternal-Effect Mutations Reveals Widespread Placenta-Specific Imprinting. PLoS Genet *11*, e1005644.