Disruption of folate metabolism causes germline epigenetic instability and distinguishes HIRA as a biomarker of maternal transgenerational epigenetic inheritance

Georgina E.T. Blake[1,2,†], Xiaohui Zhao[1,2], Hong wa Yung[1,2], Graham J. Burton[1,2], Anne C. Ferguson-Smith[2,3], Russell S. Hamilton[2,3], Erica D. Watson[1,2,*]

[1]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge UK

[2]Centre for Trophoblast Research, University of Cambridge, Cambridge UK

[3]Department of Genetics, University of Cambridge, Cambridge, UK

[†]Current address: College of Medicine and Health, University of Exeter Medical School, Exeter, UK

*Correspondence: Dr Erica Watson (edw23@cam.ac.uk)

ABSTRACT

The mechanism behind transgenerational epigenetic inheritance (TEI) is unclear, particularly through the maternal line. We previously showed that disruption of folate metabolism in mice by the *Mtrr*^{gt} hypomorphic mutation results in TEI of congenital malformations. Either maternal grandparent can initiate this phenomenon, which persists for at least four wildtype generations. Using genome-wide approaches, we reveal genetic stability in the *Mtrr*^{gt} model and epigenome-wide differential DNA methylation in the germline of *Mtrr*$^{+/gt}$ maternal grandfathers. While epigenetic reprogramming occurs, wildtype grandprogeny and great grandprogeny exhibit memory of germline methylation defects. One such region is associated with misexpression of the *Hira* gene at least until the F3 generation in a manner that distinguishes the HIRA histone chaperone as a biomarker of maternal epigenetic inheritance.

MAIN TEXT

Environmental stressors can impact an individual's health and that of their progeny[1-5]. Phenotypic risk that persists for several generations once the stressor is removed is termed transgenerational epigenetic inheritance (TEI)[6]. While the mechanism is unclear, this non-conventional inheritance likely occurs independent of DNA base-sequence and involves the inheritance of an epigenetic factor(s) via the germline[6,7]. Candidates in mammals include: DNA methylation, histone modifications, and/or non-coding RNA[1,2,4,8-10]. How an epigenetic message resists reprogramming and is transmitted between generations remains elusive.

Few mammalian models of TEI exist and the majority focus on paternal inheritance[1,2,8-11]. We reported the *Mtrr*^{gt} mouse line, a rare model of maternal TEI in which congenital malformations are transgenerationally inherited for at least four wildtype generations[3]. MTRR (methionine synthase reductase) is a key enzyme required for folate and methionine metabolism (**Supplementary Fig. 1**)[12-14]. Folate is best known for its role in neural tube closure, yet its function in development is complex and remains poorly understood. Folate metabolism is required for thymidine synthesis[15] and cellular methylation. It provides one-carbon methyl groups for the methylation of homocysteine by methionine synthase (MTR) to form methionine and tetrahydrofolate[16]. Methionine acts as a precursor for S-adenosylmethionine (SAM), which serves as the sole methyl-donor for substrates involved in epigenetic regulation including DNA, RNA, and proteins[17-19]. MTRR activates MTR through the reductive methylation of its vitamin B_{12} cofactor[14]. Consequently, MTRR helps to maintain genetic and epigenetic stability during pregnancy.

The *Mtrr*^{gt} mutation in mice knocks down *Mtrr* expression, reduces MTR enzymatic activity, and consequently disrupts folate metabolism[3,12]. Similar to humans with an *MTRR* mutation[13,20-22] or dietary folate deficiency[23], *Mtrr*$^{gt/gt}$ mice display hyperhomocysteinemia[3,12], macrocytic anaemia[24], and a wide spectrum of developmental phenotypes at midgestation (e.g., growth defects and congenital malformations including neural tube, heart and placenta defects)[3]. Therefore, *Mtrr*^{gt} mice are relevant for studying the effects of abnormal folate metabolism.

Furthermore, the *Mtrr*^{gt} mouse line is a model of maternal TEI. Through highly-controlled genetic pedigrees (**Supplementary Fig. 2a-b**), we determined that a male or female carrier of the *Mtrr*^{gt} allele (i.e., *Mtrr*$^{+/gt}$) is sufficient to cause TEI of developmental phenotypes in wildtype (*Mtrr*$^{+/+}$)

descendants[3]. Phenotypic inheritance occurs only via wildtype daughters for at least four generations[3] in conditions of normal folate metabolism[24]. The spectrum and frequency of developmental phenotypes is largely comparable in each generation, regardless of whether an *Mtrr*$^{+/gt}$ maternal grandmother or grandfather initiated the effect[3]. An exception is the F1 generation where phenotypic risk occurs only when individuals are derived from an F0 *Mtrr*$^{+/gt}$ female (**Supplementary Fig. 2a**)[3,25]. Though the mechanism is unclear, this phenomenon suggests a maternally inherited factor. Embryo transfer of F2 blastocysts demonstrated that congenital malformations (but not the growth phenotypes) occurred independent of a defective uterine environment[3] (**Supplementary Fig. 2c**) and emphasizes the mechanistic importance of germline epigenetic inheritance in the *Mtrr*^{gt} model. Epigenetic instability occurs in the *Mtrr*^{gt} model, particularly in placentas of wildtype F1 and F2 generations as indicated locus-specific dysregulation of DNA methylation associated with gene misexpression[3]. A genome-wide analysis has not yet been performed.

Here, we investigate potential mechanism(s) of TEI in the *Mtrr*^{gt} model. First, we demonstrate that *Mtrr*$^{gt/gt}$ mice are genetically stable and hence reassert focus on an epigenetic mechanism. Second, we show that the *Mtrr*^{gt} mutation alters differential DNA methylation in the germline. Sperm were chosen for analysis because of their experimental tractability and because an F0 *Mtrr*$^{+/gt}$ male can initiate TEI through his wildtype F1 daughters. We observe that differentially methylated regions (DMRs) in sperm of F0 *Mtrr*$^{+/gt}$ males are reprogrammed in somatic tissue of wildtype progeny and grandprogeny. Yet, memory of germline epigenetic disruption persists at least until the F3 generation. This includes misexpression of *Hira*, a gene important for chromatin stability. We propose *Hira* as a biomarker of maternal TEI in the *Mtrr*^{gt} model.

RESULTS

Genetic stability in *Mtrr*^{gt} mice

Since folate metabolism is directly linked to DNA synthesis, we first addressed whether the *Mtrr*^{gt} allele influences genetic stability. Whole genome sequencing (WGS) was performed on phenotypically normal C57Bl/6J control embryos (n=2) and *Mtrr*$^{gt/gt}$ embryos with congenital malformations (n=6) (**Supplementary fig. 2b,e**). DNA libraries were sequenced resulting in ~30x coverage per sample (~3.5 x 10^8 paired-end reads/genome). The sequenced genomes were compared to the C57Bl/6J reference genome to identify structural variants [SVs] and single nucleotide polymorphisms [SNPs].

The *Mtrr*^{gt} mutation was generated by a genetrap (gt) insertion in the 129P2Ola/Hsd background before backcrossing for eight generations into the C57Bl/6J strain[3]. As expected, the majority of variants identified in *Mtrr*$^{gt/gt}$ embryos were located on Chr13 in the genomic region surrounding the *Mtrr* gene (**Supplementary Fig. 3a-b**). Many SNPs in this region showed sequence similarity to the 129P2Ola/Hsd genome and likely persisted due to *Mtrr*^{gt} genotype selection and regional crossover frequency. Using these SNPs, we defined a 20 Mb region of 129P2Ola/Hsd sequence surrounding the *Mtrr*^{gt} allele (**Fig. 1a**). When this region was masked, C57Bl/6J and *Mtrr*$^{gt/gt}$ embryos contained a similar frequency of SNPs (C57Bl/6J: 4,871 ± 791 SNPs/embryo; *Mtrr*$^{gt/gt}$: 5,138 ± 398 SNPs/embryo; p=0.781) and SVs (C57Bl/6J: 342 SVs/embryo; *Mtrr*$^{gt/gt}$: 301 SVs/embryo; p=0.6886; **Fig. 1b-c**) implying that the de novo mutation rate was unchanged by the mutation. The majority of SNPs and SVs identified in C57Bl/6J and *Mtrr*$^{gt/gt}$ embryos represented non-coding variants or were located in non-coding regions (**Supplementary Fig. 3c-d**). Moreover, genetic variation within the masked region had minimal functional effect (beyond that of the gene-trap insertion) since expression of individual genes from this region was similar among C57Bl/6J, 129P2Ola/Hsd and *Mtrr*$^{gt/gt}$ mice (**Fig. 1a,d**). As further indication of genomic stability, we showed that expression of several transposable elements[26] was similar in C57Bl/6J and *Mtrr*$^{gt/gt}$ tissue (**Fig. 1e**) indicating that their repressive state was preserved. Overall, these data suggested that the genome of the *Mtrr*^{gt} model was stable, and that phenotypic inheritance was unlikely caused by an increased frequency of de novo mutation. Therefore, focus shifted to an epigenetic mechanism.

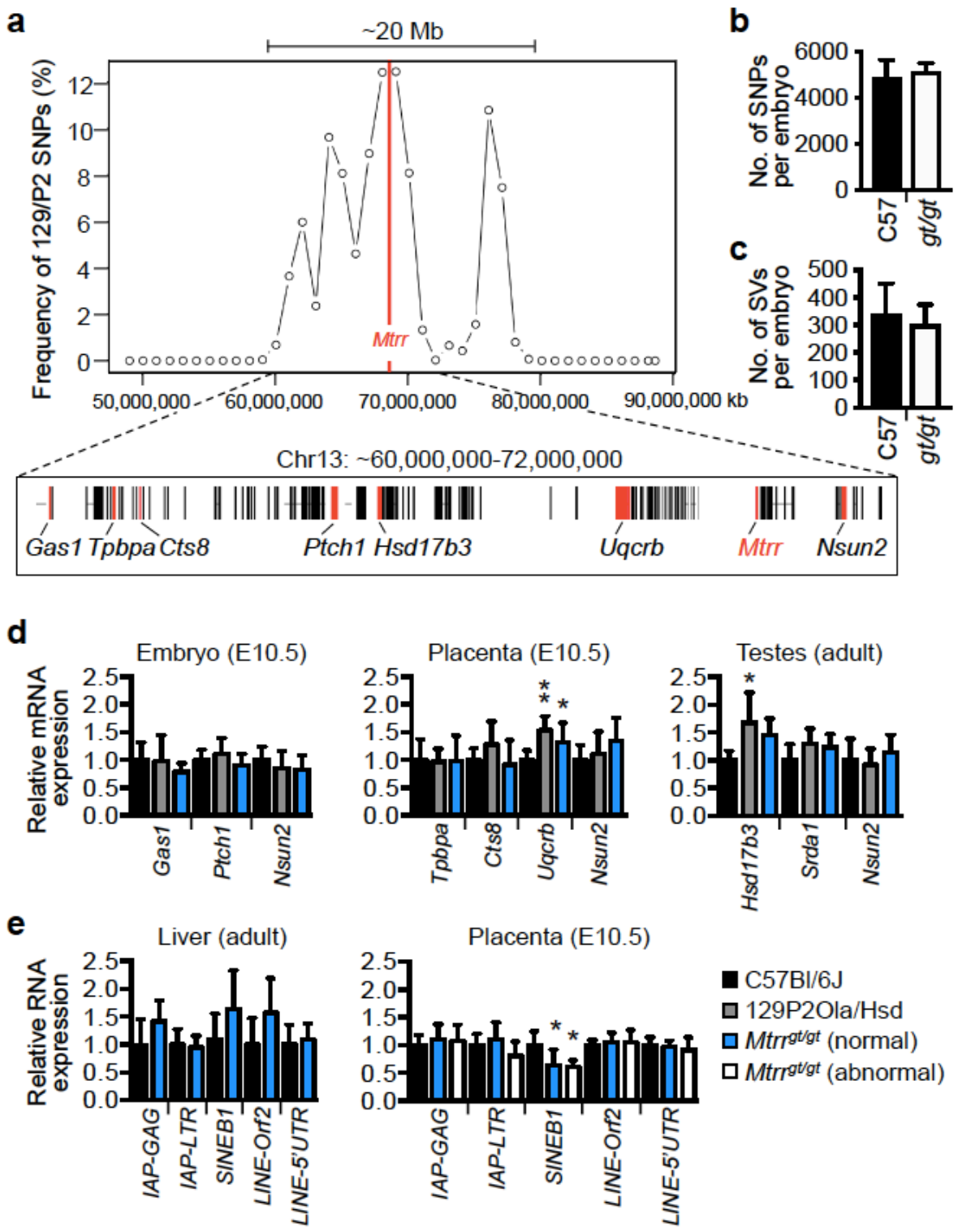


Fig. 1. The *Mtrr*^{gt} mouse line is genetically stable.
a-c, Whole genome sequencing of normal C57Bl/6 embryos (N=2) and severely affected *Mtrr*$^{gt/gt}$ embryos (N=6) at E10.5 to determine the frequency of genetic variants compared to the C57Bl/6J reference genome. **a,** The frequency of 129P2Ola/Hsd single nucleotide polymorphisms (SNPs) in the region surrounding the gene-trap insertion site in the *Mtrr* gene (red line). The majority of genes within the 20 Mb region surrounding the *Mtrr* gene are shown below the graph. **b, c,** The average number of (**b**) SNPs and (**c**) structural variants (SVs) per embryo in C57Bl/6 embryos (C57, black bars) and *Mtrr*$^{gt/gt}$ embryos (*gt/gt*, white bars). The 20 Mb region shown in **a** was masked when calculating the average number of genetic variants in (**b, c**). Data is plotted as mean

[Fig. 1 continued] ± standard deviation (sd). Independent t test. **d,** Graphs showing RT-qPCR analysis of selected genes (highlighted red in **a**) in embryos (E10.5), placentas (E10.5) and/or adult testes from C57Bl/6J (black bars), 129P2Ola/Hsd (grey bars) and phenotypically-normal *Mtrr*$^{gt/gt}$ (blue bars) mice. **e,** Data indicating RNA expression of specific groups of transposable elements as determined by RT-qPCR in C57Bl/6J tissue (black bars) and *Mtrr*$^{gt/gt}$ tissue (blue bars: phenotypically normal; white bars: severely affected). Adult liver and placenta at E10.5 were assessed. Data from RT-qPCR analyses in (**d**, **e**) are shown as mean ± sd and relative to C57Bl/6J tissue levels (normalized to 1). N=5-6 samples per group. One-way ANOVA with Dunnett's multiple comparison test, *p<0.05; **p<0.01.

Germline DNA methylation is altered in the *Mtrr*^{gt} model

MTRR plays a direct role in the transmission of one-carbon methyl groups for DNA methylation[3,12,14]. Therefore, germline DNA methylation was considered as a potential mediator of phenotype inheritance. Since an *Mtrr*$^{+/gt}$ maternal grandmother (MGM) or maternal grandfather (MGF) initiates TEI (**Supplementary Fig. 2a**)[3,6] and due to the experimental tractability of male gametes, we chose to focus our analysis on sperm. Mature spermatozoa were collected from C57Bl/6J, *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ mice (**Supplementary Fig. 2b,d-e**) and the purity was confirmed by assessing imprinted regions of known methylation status via bisulfite pyrosequencing (**Supplementary Fig. 4a**). Global 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) levels were normal across all *Mtrr* genotypes relative to C57Bl/6J controls as determined by mass spectrometry (**Fig. 2a**).

To analyse genome-wide distribution of sperm DNA methylation, methylated DNA immunoprecipitation followed by sequencing (MeDIP-seq) was performed. This approach allowed the unbiased detection of robust locus-specific changes in DNA methylation by identifying clusters of differentially methylated cytosines, thus reducing the potential impact of single-nucleotide variants[4,27]. MeDIP libraries of sperm DNA were prepared using eight males each from C57Bl/6J, *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ genotypes (**Supplementary Fig. 4b**). Sequencing generated 179 million paired-end mappable reads on average per group (C57Bl/6J: 164 million reads; *Mtrr*$^{+/+}$: 172 million reads; *Mtrr*$^{+/gt}$: 203 million reads; *Mtrr*$^{gt/gt}$: 179 million reads). Using MEDIPS package[28], each *Mtrr* genotype was independently compared to C57Bl/6J controls. Loci of >500 bp with a methylation change of >1.5-fold and p value of <0.01 were defined as DMRs. The number of DMRs identified increased with the severity of *Mtrr* genotype: 91 DMRs in *Mtrr*$^{+/+}$ males, 203 DMRs in *Mtrr*$^{+/gt}$ males and 599 DMRs in *Mtrr*$^{gt/gt}$ males (**Fig. 2b**). The presence of DMRs in sperm from *Mtrr*$^{+/+}$ males

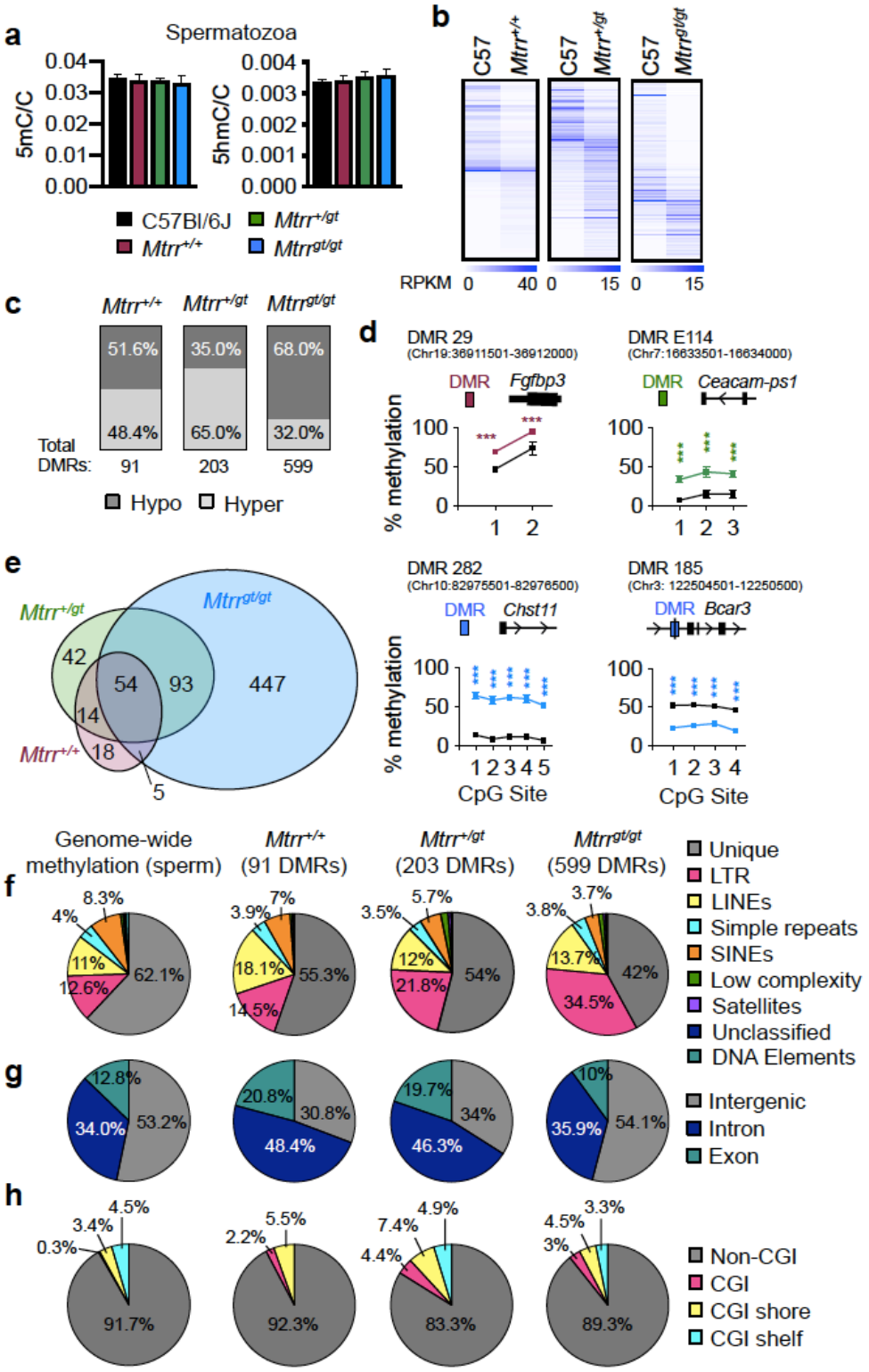


Fig. 2. Characterization of differential DNA methylation in spermatozoa from *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$, and *Mtrr*$^{gt/gt}$ mice.

a, Global 5mC and 5hmC in spermatozoa from C57Bl/6J (black bars), *Mtrr*$^{+/+}$ (purple bars), *Mtrr*$^{+/gt}$ (green bars) and *Mtrr*$^{gt/gt}$ (blue bars) adult males (N=9 males/group) as assessed by mass spectrometry. Data is presented as ratio of methylated cytosines per genomic cytosines (mean ± sd). One-way ANOVA. **b,** Heat maps plotting reads per kilobase million (RPKM, window size=500 bp) from MeDIP-seq analysis of spermatozoa from *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males compared to C57Bl/6J males ($p<0.05$). N=8 males/group. **c**, Percentage of total DMRs identified in each genotype that were hypomethylated (dark grey) or hypermethylated (light grey). **d,** Examples of

[Fig. 2 continued] sperm DMRs identified via MeDIP-seq and validated by bisulfite pyrosequencing from *Mtrr*$^{+/+}$ (purple line), *Mtrr*$^{+/gt}$ (green line), or *Mtrr*$^{gt/gt}$ (blue lines) males compared to C57Bl/6J sperm (N=8 males/group including four MeDIP-seq samples and four independent samples). Data is shown as percentage methylation at each CpG site assessed (mean ± sd). Schematic of each DMR is indicated in relation to the closest gene. Two-way ANOVA with Sidak's multiple comparisons test performed. **p<0.01, ***p<0.0001. **e,** An intersection analysis of sperm DMRs from *Mtrr*$^{+/+}$ (purple), *Mtrr*$^{+/gt}$ (green) and *Mtrr*$^{gt/gt}$ (blue) males compared to C57Bl/6J sperm. **f-h,** Relative distribution of methylated regions in C57Bl/6J sperm (background methylome) and sperm DMRs from *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males among (**f**) unique sequences and repetitive elements, (**g**) coding and non-coding regions, and (**h**) CpG islands (CGIs), shores and shelves. **f** and **h**, Chi-squared test; **g**, Two-way ANOVA with Dunnett's multiple comparison test.

derived from *Mtrr*$^{+/gt}$ intercrosses (**Supplementary Fig. 2d**) indicated a parental effect of the *Mtrr*^{gt} allele. Hypomethylated and hypermethylated regions were identified in each *Mtrr* genotype when compared to C57Bl/6J controls (**Fig. 2c**), consistent with earlier findings in placentas[3]. These data suggested that the *Mtrr*^{gt} allele was sufficient to dysregulate sperm DNA methylation.

To ensure the robustness and reliability of the MeDIP-seq data, we randomly selected hyper- and hypomethylated DMRs to validate using bisulfite pyrosequencing. Sperm DNA from C57Bl/6J, *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males was assessed (N=8 males/group: four sperm samples from MeDIP-seq experiment plus four independent samples). DMRs were validated in the *Mtrr* genotype in which they were identified (**Figs. 2d, 3, Supplementary Fig. 5**). The overall validation rate was 94.1% in hypomethylated DMRs and 58.3% in hypermethylated DMRs (**Supplementary Table 1**) and indicated a high degree of corroboration between techniques. The majority of DMRs that did not validate showed extensive methylation (>80% CpG methylation) in C57Bl/6J sperm and were identified as hypermethylated in the MeDIP-seq experiment (**Supplementary Fig. 5**). This might reflect some false positives in line with another study[4].

For most DMRs assessed, methylation change was consistent across all CpG sites and the absolute change in CpG methylation ranged from 10 to 80% of control levels (**Figs. 2d, 3, Supplementary Fig. 5**). Within each genotypic group, a high degree of inter-individual consistency of methylation change was also observed. Therefore, we conclude that the *Mtrr*^{gt} mutation, or parental exposure to it as in *Mtrr*$^{+/+}$ males, is sufficient to lead to distinct DNA methylation changes in sperm.

Most DMRs associated with metabolic dysregulation rather than genetic effects

A proportion of the DMRs was located within the region around the gene-trap insertion site in *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males (**Fig. 1a**, **Supplementary Fig. 6**), consistent with *Mtrr*$^{gt/gt}$ liver[29] and suggesting that the gene-trap or underlying 129P2Ola/Hsd sequence might epigenetically dysregulate the surrounding region. Comparison to the whole genome sequencing data revealed that genetic variation did not influence DMR calling to a great extent since only a small proportion (2.8-5.5%) of these DMRs contained one or more SNP. Outside of the *Mtrr* genomic region, 54 DMRs were common to *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males (**Fig. 2e**) and were primarily located in distinct chromosomal clusters (**Supplementary Fig. 6**). These data implicate epigenetic hotspots or underlying genetic effects. However, beyond a polymorphic duplication on Chr19 in the C57Bl/6J strain[30] that accounted for a minor number of DMRs (2.5-15.8% of DMRs), no DMRs were directly associated with an SV or located <1 kb. Once accounting for potential genetic effects, the majority of sperm DMRs in *Mtrr*$^{+/+}$ and *Mtrr*$^{+/gt}$ males (76/91 DMRs and 142/203 DMRs, respectively), and a proportion of sperm DMRs in *Mtrr*$^{gt/gt}$ males (174/599 DMRs) were attributed to the metabolic consequences of the *Mtrr*^{gt} mutation.

Sperm DMR genomic distribution and potential regulatory function

DMR distribution was determined to explore regional susceptibility of the sperm methylome to the *Mtrr*^{gt} allele. First, the sperm 'background methylome' was established to resolve the expected genome-wide distribution of CpG methylation (see Methods). By comparing the regional distribution of sperm DMRs to the background methylome, DMRs in all *Mtrr* genotypes were not significantly enriched in repetitive regions (**Fig. 2f**). However, sperm of *Mtrr*$^{+/+}$ and *Mtrr*$^{+/gt}$ males had an over-representation of DMRs in introns and exons, and under-representation in intergenic regions ($p<0.0003$, Chi-squared test; **Fig. 2g**). This was not the case for *Mtrr*$^{gt/gt}$ males since DMRs became more proportionately distributed among most genomic regions (**Fig. 2g**). While the majority of sperm DMRs were located within CpG deserts, DMRs from *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males were over-represented in CpG islands ($p<0.0014$, Chi-squared; **Fig. 2h**). Altogether, these results might have implications for gene regulation.

During sperm maturation, histones are replaced by protamines[31]. However, nucleosome

retention occurs primarily at promoters of developmentally-regulated genes and gene-poor repeat regions providing scope for epigenetic inheritance[32-34]. Using a published data set[34], the genome-wide nucleosome retention rate was estimated at 1.94% as determined by assessing 10,000 randomly-selected 500 bp windows. Some DMRs were significantly enriched at nucleosome retention regions (14.5-34.1% of DMRs, $p<0.0001$, binomial test; **Table 1**) and represented key candidate regions for epigenetic inheritance.

Table 1. Number of DMRs identified in regions of nucleosome retention in sperm or regions resistant to reprogramming in the pre-implantation embryo and/or germline.

Male genotype	***Mtrr*$^{+/+}$**	***Mtrr*$^{+/gt}$**	***Mtrr*$^{gt/gt}$**
Total No. of sperm DMRs	91	203	599
Nucleosome retention[34]	31 (34.1%)	57 (28.1%)	87 (14.5%)
Resistant to reprogramming in pre-implantation embryo[35]	37 (40.7%)	96 (47.3%)	325 (54.3%)
Resistant to reprogramming in germline[36]	2 (2.2%)	5 (2.5%)	23 (3.8%)
Resistant to reprogramming in pre-implantation embryo[35] and germline[36]	2 (2.2%)	4 (2.0%)	16 (2.7%)
Nucleosome retention[34] and resistant to reprogramming in pre-implantation embryo[35]	19 (20.9%)	36 (17.7%)	49 (8.9%)
Nucleosome retention[34] and resistant to reprogramming in germline[36]	1 (1.1%)	2 (1.0%)	4 (0.67%)

Therefore, to explore the epigenetic signature of DMRs, mean enrichment for histone modifications and/or Tn5 transposase sensitive sites (THSS) in mouse spermatozoa[37], epiblast[38] and extraembryonic ectoderm (ExE)[38] at E6.5 was determined using published ChIP-seq and ATAC-seq data sets. All DMRs, except those surrounding the *Mtrr* gene-trapped site, were analysed (N=379 DMRs, all *Mtrr* genotypes combined) alongside randomly selected regions (N=379; see Methods) representing the 'baseline genome'. In sperm[37], DMRs were collectively enriched for the repressive histone mark H3K9me3 and not active histone marks (e.g., H3K4me1, H3K27ac) compared to the background methylome (**Supplementary Fig. 7a**), and were more likely to associate with a closed chromatin state due to enrichment for protamine 1 (PRM1) and not THSS (**Supplementary Fig. 7a-b**). In contrast, the same regions in epiblast and ExE tissues[38]

were more likely in an open chromatin conformation when compared to tissue-specific baseline methylomes (**Supplementary Fig. 7c**). Therefore, a regulatory role for the DMRs during development is possible and required further evaluation (see below).

DMRs were located in regions of reprogramming resistance

DNA methylation is largely reprogrammed during pre-implantation development and in the developing germline[39,40]. Recently, several loci were identified as 'reprogramming resistant'[35,36,41]. Using published data sets[35,36], we determined that 40.7-54.3% of sperm DMRs across all *Mtrr* genotypes fell within loci resistant to pre-implantation reprogramming (**Table 1**). Sixteen of these DMRs were common among *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males. Fewer DMRs correlated with regions resistant to germline reprogramming or both pre-implantation and germline reprogramming (2.2-3.8% and 2.0-2.7% of DMRs/*Mtrr* genotype, respectively; **Table 1**). Only one DMR located in a region resistant to germline reprogramming was common to all genotypes. Interestingly, several DMRs in reprogramming resistant regions[35,36] also overlapped with regions of nucleosome retention[34] (**Table 1**). Overall, DMRs in these key regions might be important for epigenetic inheritance.

Sperm DMRs are reprogrammed in wildtype F1 and F2 generations

TEI in the *Mtrr*^{gt} model occurs via the maternal lineage. However, an F0 *Mtrr*$^{+/gt}$ male can initiate the effect[3] (**Supplementary Fig. 2a**). To determine the heritability of sperm DMRs, bisulfite pyrosequencing was used to validate ten DMRs from F0 *Mtrr*$^{+/gt}$ males in tissue of wildtype F1 and F2 progeny. Candidate DMRs displayed features including reprogramming resistance, localization to intragenic or intergenic regions, and/or hyper- or hypomethylation (**Supplementary Table 2**). In general, all DMRs tested lost their differential methylation in wildtype F1 and F2 embryos and placentas at embryonic day (E) 10.5, and showed DNA methylation patterns similar to C57Bl/6J tissue (**Fig. 3**). This result occurred even when wildtype F2 conceptuses displayed congenital malformations (**Fig. 3**). DMRs were also assessed in *Mtrr*$^{gt/gt}$ conceptuses at E10.5 to determine whether these regions were capable of differential methylation outside of the germline. Seven out of 10 sperm DMRs were hypermethylated in *Mtrr*$^{gt/gt}$ embryos and/or placentas (**Supplementary**

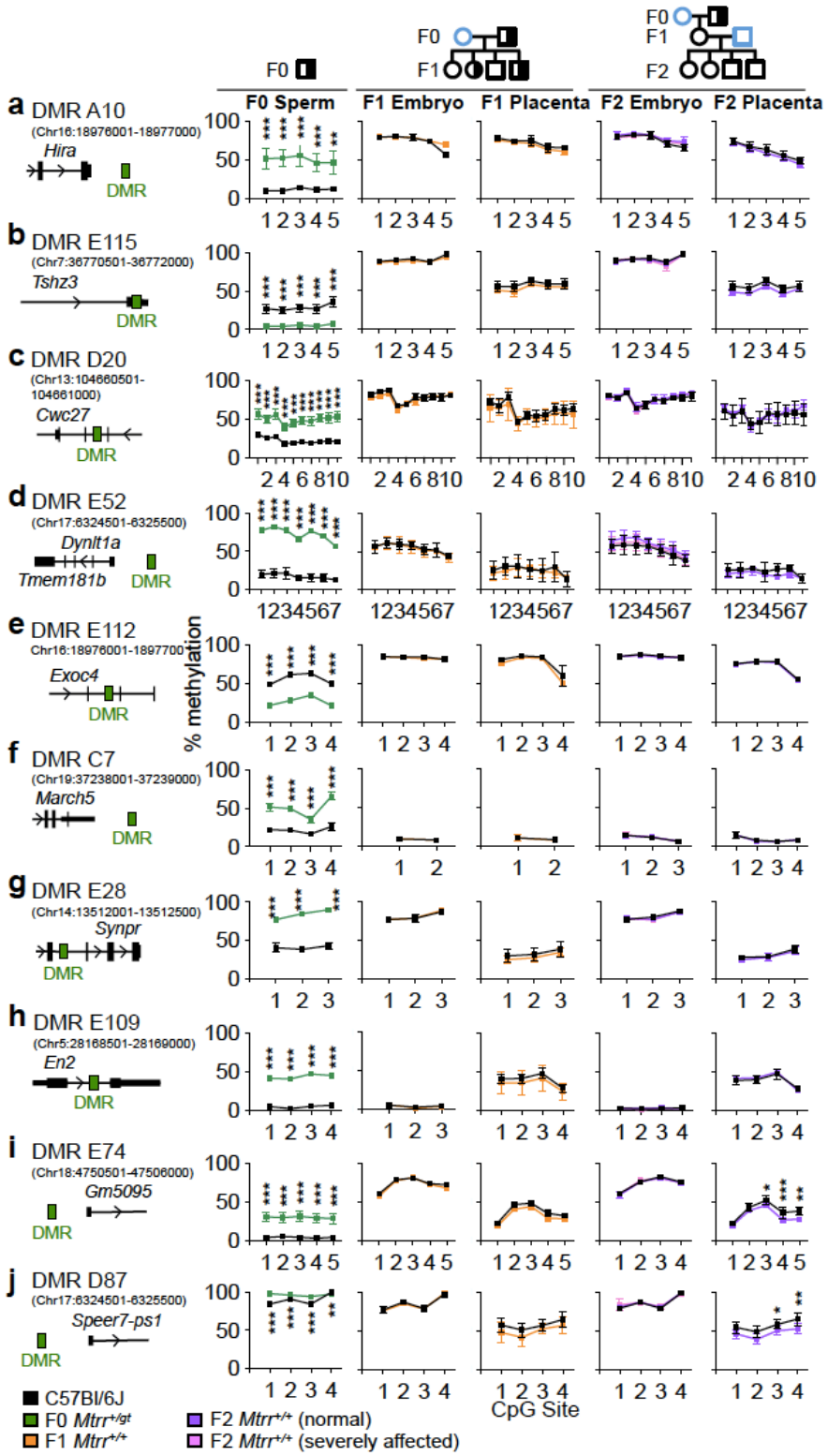


Fig. 3. Sperm DMRs are reprogrammed in somatic tissue of F1-F2 wildtype generations. CpG methylation at specific sperm DMRs identified in F0 *Mtrr*$^{+/gt}$ males was assessed in wildtype embryos and placentas at E10.5 from F1 and F2 progeny. Pedigrees indicate specific mating scheme. Pedigree legend: circle, female; square, male; blue outline, C57Bl/6J control; black outline, *Mtrr*^{gt} mouse line; white filled, *Mtrr*$^{+/+}$; half-white/half-black filled, *Mtrr*$^{+/gt}$. **a-j,** Schematic drawings of each DMR assessed indicate its relationship to the closest gene and are followed by graphs showing the average percentage of methylation at individual CpGs for the corresponding DMR as determined by bisulfite pyrosequencing. In each case, methylation was assessed in sperm from F0 *Mtrr*$^{+/gt}$ males (green line), phenotypically normal F1 wildtype (*Mtrr*$^{+/+}$) embryos and

[Fig. 3 continued] placentas at E10.5 (orange lines), and phenotypically normal (purple line) or severely affected (pink line) F2 wildtype (*Mtrr*$^{+/+}$) embryos and placentas at E10.5. C57Bl/6J (black lines) are shown as controls. N=4-8 individuals assessed per group. Data is shown as mean ± sd for each CpG site. Two-way ANOVA, with Sidak's multiple comparisons test, performed on mean methylation per CpG site per genotype group. *p<0.05, **p<0.01, ***p<0.001.

Fig. 8a-j), a similar pattern to sperm from *Mtrr*$^{+/gt}$ males (**Fig. 3**). It was unclear whether these regions resisted reprogramming, or were erased and re-established or maintained due to intrinsic *Mtrr*$^{gt/gt}$ homozygosity. Overall, the altered DNA methylation in sperm of *Mtrr*$^{+/gt}$ males was not evident in somatic tissue of wildtype progeny and grandprogeny.

Epigenetic memory of sperm DMRs

Previous studies in an intergenerational model suggested that sperm DMRs might be associated with perturbed transcription in offspring even when DNA methylation was re-established to normal levels[4]. Therefore, expression of six genes located in or near sperm DMRs identified in F0 *Mtrr*$^{+/gt}$ males was assessed in F1 and F2 wildtype individuals. While these genes displayed normal expression in F1 tissues (**Fig. 4a-c**), *Hira* (histone chaperone), *Cwc27* (spliceosome-associated protein), and *Tshz3* (transcription factor) (**Supplementary Table 2**) were misexpressed in F2 wildtype embryos or adult livers compared to C57Bl/6J controls (**Fig. 4d-f**). This result might reflect an epigenetic memory of the associated sperm DMR or wider epigenetic dysregulation in sperm of the F0 *Mtrr*$^{+/gt}$ males.

Next, the genetic and epigenetic characteristics of *Hira*, *Cwc27*, and *Tshz3* DMRs were considered to predict their regulatory potential. Genomically, *Cwc27* and *Tshz3* DMRs were intragenic (**Supplementary Figs. 9-10**) while the *Hira* DMR was located ~6 kb downstream of the gene (**Fig. 5a-b**). To understand the DMR epigenetic signature in a developmental context, published ChIP-seq data sets in spermatozoa[37], and wildtype embryonic stem cell (ESC)[42,43] and trophoblast stem cell (TSC)[44-46] lines were assessed. In TSCs, histone marks at all three DMRs were largely absent (**Fig. 5a, Supplementary Figs. 9-10**). However, *Tshz3* and *Hira* DMRs, which overlapped with CpG islands[47], demonstrated CTCF binding in ESCs and TSCs (**Fig. 5a, Supplementary Fig. 10**). In contrast, both ESCs and sperm displayed enrichment for active (e.g., H3K4me1, H3K27ac) and repressive (e.g., H3K27me3, H3K9me3) marks at *Hira*, *Cwc27*, and *Tshz3* DMRs (**Fig. 5a, Supplementary Fig. 9-10**). Of note, all three DMRs displayed ESC-specific

TET1 binding (**Fig. 5a, Supplementary Fig. 9-10**), which further implicates a tissue-specific regulatory role, potentially involving 5hmC. Whether histone marks at these DMRs and others are dysregulated by an ancestral or intrinsic *Mtrr*^{gt} allele is yet-to-be determined.

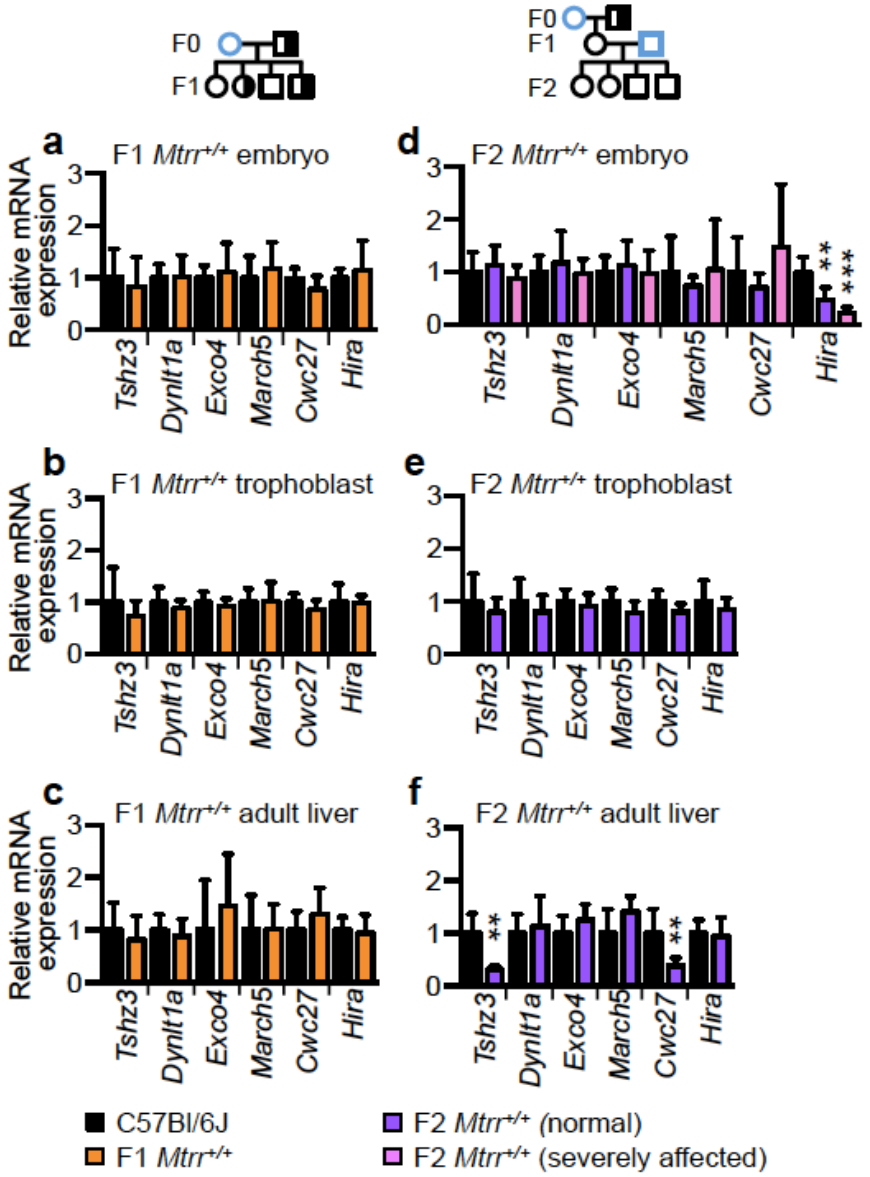


Fig. 4. Transcriptional disruption of some DMR-associated genes in F1 and F2 wildtype somatic tissue indicates the potential for epigenetic memory.
a-f, RT-qPCR analysis of mRNA expression of genes located in close proximity to sperm DMRs in (**a, d**) embryos and (**b, e**) placentas at E10.5, and (**c, f**) adult livers of wildtype F1 and F2 progeny. Tissue from (**a-c**) phenotypically normal F1 *Mtrr*$^{+/+}$ individuals (orange bars), and (**d-f**) phenotypically normal F2 *Mtrr*$^{+/+}$ individuals (purple bars) or severely affected (pink bars) F2 *Mtrr*$^{+/+}$ individuals derived from an F0 *Mtrr*$^{+/gt}$ male. Pedigrees indicate specific mating scheme assessed. C57Bl/6J control tissues (black bars) were assessed as controls. Data is plotted as mean ± sd, and is presented as relative expression to C57Bl/6J levels (normalized to 1). N=4-7 individuals per group. Independent t tests or one-way ANOVA with Dunnett's multiple comparisons tests were performed. **p<0.01, ***p<0.001. Each gene was associated with the following sperm DMR: *Tshz3*, DMR E115; *Dynlt1a*, DMR E52; *Exoc4* DMR E112; *March5*, DMR C7; *Cwc27*, DMR D20; *Hira*, DMR A10. Pedigree legend: circle, female; square, male; blue outline, C57Bl/6J control; black outline, *Mtrr*^{gt} mouse line; white filled, *Mtrr*$^{+/+}$; half-white/half-black filled, *Mtrr*$^{+/gt}$.

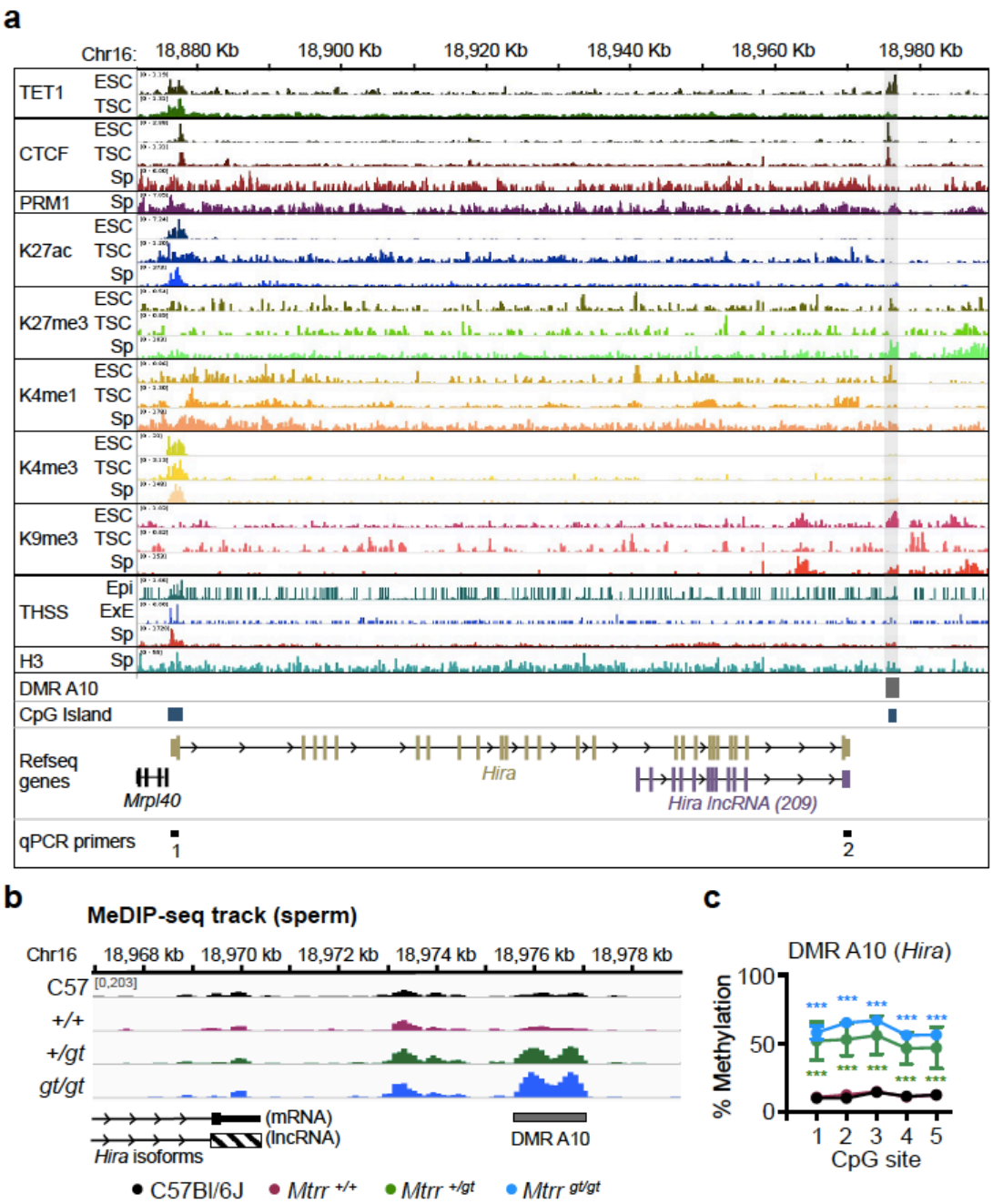


Fig. 5. Epigenetic characteristics of *Hira* DMR in normal cells and sperm from *Mtrr*^{gt} males. **a**, Enrichment of DNA binding proteins (TET1, CTCF, PRM1, H3) and histone modifications (H3K27ac, H3K27me3, H3K4me1, H3K4me3, H3K9me3) in the *Hira* locus on Chr16 using published ChIP-seq data sets in cultured wildtype embryonic stem cells (ESCs), trophoblast stem cells (TSCs)[42], and wildtype spermatozoa (Sp)[37]. Tn5 transposase sensitive sites (THSS) were also determined using published ATAC-seq data sets of normal B6D2F1 mouse epiblast and extraembryonic ectoderm (ExE) at embryonic day 6.5[38] or CD1 wildtype spermatozoa[37]. Grey box and shading indicated the *Hira* DMR identified in sperm of *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males. Blue boxes indicate CpG islands. Schematics of protein encoding (brown) and lncRNA encoding (purple) *Hira* isoforms are shown at the bottom. Region of RT-qPCR primer sets 1 and 2 are also indicated. **b**, Schematic drawing of *Hira* transcripts and *Hira* DMR (DMR A10; grey rectangle) in relation to MeDIP-seq reads in sperm from C57Bl/6J (black), *Mtrr*$^{+/+}$ (purple), *Mtrr*$^{+/gt}$ (green) and *Mtrr*$^{gt/gt}$ (blue) males. N=8 males per group. **c**, The average percentage methylation at individual CpGs (mean ± sd) in the *Hira* DMR in sperm from C57Bl/6J (black line), *Mtrr*$^{+/+}$ (purple line), *Mtrr*$^{+/gt}$ (green line) and *Mtrr*$^{gt/gt}$ (blue line) males. N=3 males per group. Two-way ANOVA with Sidak's multiple comparisons test performed on mean methylation per CpG site. ***p<0.001.

HIRA as a potential mediator of maternal inheritance in the *Mtrr^{gt}* model

The *Hira* DMR was further considered based on its resistance to germline reprogramming[36] and potential for epigenetic memory (**Fig. 4d**). HIRA is a histone H3.3 chaperone central to transcriptional regulation[48], and to maintenance of chromatin structure during oogenesis[49] and in the male pronucleus after fertilization[50]. *Hira$^{-/-}$* mice[51] and the *Mtrr^{gt}* mouse line[3] display similar phenotypes including growth defects, congenital malformations and embryonic lethality by E10.5. Furthermore, *Mtrr^{gt}* genotypic severity correlated with the degree of hypermethylation in the *Hira* DMR in sperm (**Fig. 5b-c**) suggesting that the *Hira* DMR is responsive to alterations in folate metabolism.

We sought to elucidate how *Hira* expression was affected by *Mtrr$^{gt/gt}$* homozygosity during development. *Mtrr$^{gt/gt}$* conceptuses at E10.5 derived from *Mtrr$^{gt/gt}$* intercrosses display similar phenotypes to the MGF pedigrees but at a higher frequency[3]. While methylated at control levels in phenotypically-normal *Mtrr$^{gt/gt}$* embryos, five CpGs (out of 128 CpGs) in the *Hira* DMR were 9.1 ± 0.5% (mean ± sd, $p<0.05$) more methylated than controls in the associated *Mtrr$^{gt/gt}$* placentas and 48.6 ± 4.9% ($p<0.001$) in spermatozoa from *Mtrr$^{gt/gt}$* males (**Fig. 5c**, **Supplementary Fig. 8b**). Yet, the effect of *Mtrr$^{gt/gt}$* homozygosity on *Hira* RNA and protein expression was embryo-specific. For instance, *Mtrr$^{gt/gt}$* embryos exhibited a significant up-regulation of *Hira* mRNA (2.18-fold) and protein (2.90-fold; **Fig. 6a-b, Supplementary fig. 11e**) compared to controls and a coincidental down-regulation of *Hira* lncRNA (*Hira* lncRNA 209; 2.25-fold, $p<0.01$; **Figs. 5a, 6a**). Contrastingly, both *Hira* isoforms and HIRA protein were normally expressed in the associated *Mtrr$^{gt/gt}$* placentas (**Fig. 6a-b, Supplementary fig. 11e**).

Next, the potential regulatory legacy of the *Hira* DMR in sperm was explored in the *Mtrr$^{+/gt}$* MGF pedigree. We previously reported that nearly all MGF F1 conceptuses from this pedigree were phenotypically normal at E10.5 whereas F2-F3 wildtype conceptuses displayed a wide spectrum and frequency of developmental phenotypes[3] (**Supplementary Fig. 2a**). Here, we showed substantial hypermethylation of the *Hira* DMR (39.0 ± 4.1% average increase per CpG) was observed in sperm of F0 *Mtrr$^{+/gt}$* males compared to controls (**Fig. 5c**), yet the *Hira* DMR was reprogrammed in MGF F1-F3 wildtype embryos and placentas (**Fig. 3a, Supplementary fig. 11a**). Correlated with this finding was embryo-specific *Hira* mRNA and lncRNA misexpression in a

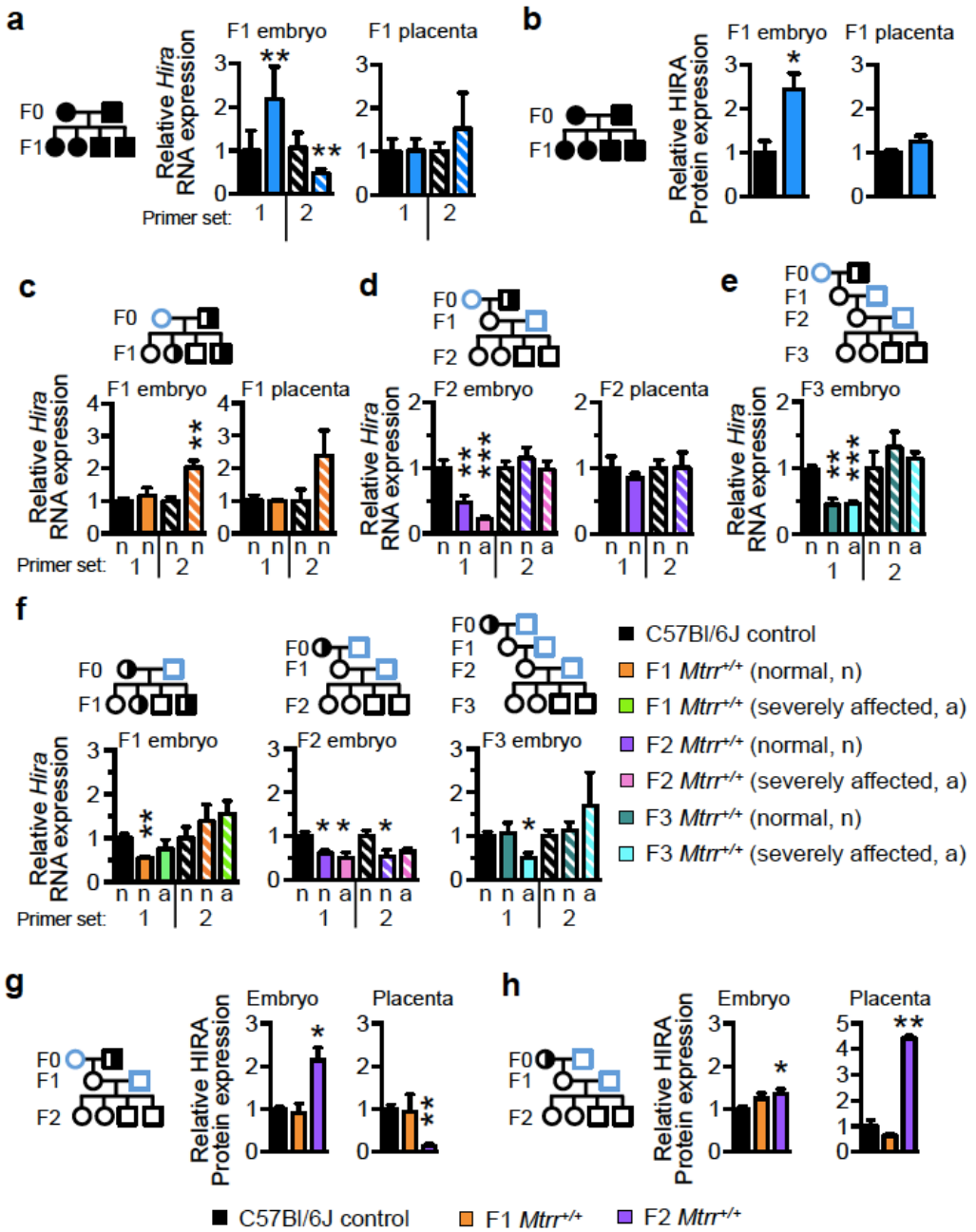


Fig. 6. Dysregulation of *Hira* RNA and protein in F1-F3 conceptuses coincides with a pattern of maternal inheritance.

a, RT-qPCR data showing *Hira* mRNA (solid bars, primer set 1 (1)) and *Hira* lncRNA (striped bars, primer set 2 (2)) in embryos and placentas from C57Bl/6J (black bars) and F1 $Mtrr^{gt/gt}$ conceptuses (blue bars). **b**, Western blot analysis of HIRA protein in C57Bl/6 (black bars) and $Mtrr^{gt/gt}$ (blue bars) embryos and placentas. **c-f**, RT-qPCR data showing *Hira* mRNA (solid bars, primer set 1) and *Hira* lncRNA (striped bars, primer set 2) in embryos and placentas from (**c**) F1 $Mtrr^{+/+}$ (orange bars), (**d**) F2 $Mtrr^{+/+}$, and (**e**) F3 $Mtrr^{+/+}$ conceptuses derived from (**c-e**) F0 $Mtrr^{+/gt}$ males or (**f**) F0 $Mtrr^{+/gt}$ females. C57Bl/6J conceptuses (black bars) were controls. Phenotypically-normal (n) and severely affected (a) conceptuses were assessed. **g,h,** Western blot analysis showing HIRA protein expression in embryos and placentas of C57Bl/6 conceptuses (black bars) and F1 $Mtrr^{+/+}$ (orange bars) and F2 $Mtrr^{+/+}$ (purple) conceptuses derived from either (**g**) F0 $Mtrr^{+/gt}$ males or (**h**) F0 $Mtrr^{+/gt}$ females. For RNA and protein data, N=3-10 individuals were assessed. HIRA protein levels were first normalised to β-actin (see **Supplementary Fig. 11c-e**). All data was plotted as mean ± sd and relative to C57Bl/6J (normalised to 1). Independent t tests or Kruskal-Wallis test with Dunn's multiple comparison test. *p<0.05, **p<0.01, ***p<0.001. Pedigree legend, circle, female; square, male; blue outline, C57Bl/6J line; black outline, $Mtrr^{gt}$ mouse line; white fill, $Mtrr^{+/+}$; half black-half white fill, $Mtrr^{+/gt}$; black fill, $Mtrr^{gt/gt}$.

pattern indicative of phenotypic inheritance; *Hira* mRNA was down-regulated in MGF F2-F3 wildtype embryos and not F1 wildtype embryos when compared to controls (**Fig. 6c-e**). Conversely, significant dysregulation of *Hira* lncRNA expression was apparent only in MGF F1 wildtype embryos (**Fig. 6c-e**). As phenotypic severity increased, so too did the degree of *Hira* mRNA dysregulation in the MGF F2 and F3 embryos (**Fig. 6d-e**). Similar to *Mtrr*$^{gt/gt}$ conceptuses, placental *Hira* RNA expression was unaffected (**Fig. 6a,c-d**).

To investigate a link between *Hira* expression and phenotypic inheritance, we analysed wildtype F1-F3 conceptuses at E10.5 derived from F0 *Mtrr*$^{+/gt}$ maternal grand*mothers* (**Supplementary Fig. 2a**), of which all generations display a wide spectrum of developmental phenotypes[3]. Supporting our hypothesis, *Hira* mRNA expression was down-regulated in MGM F1-F3 wildtype embryos compared to C57Bl/6J controls (**Fig. 6f**). As expected, *Hira* lncRNA transcripts were unchanged in MGM F1 and F3 wildtype embryos, yet were down-regulated in MGM F2 embryos (**Fig. 6f**), which display the highest frequency of phenotypes among the three generations[3]. Overall, these data indicated that *Hira* RNA is a potential marker of maternal phenotypic inheritance.

Disruption of post-transcriptional regulation of HIRA was also apparent in F2 embryos and placentas of both *Mtrr*$^{+/gt}$ maternal grandparental pedigrees, yet in a manner not predicted by *Hira* mRNA expression. For example, down-regulated *Hira* mRNA (**Fig. 6c-d,f**) was associated with normalized or up-regulated HIRA protein (**Fig. 6g-h, Supplementary fig. 11c-d**), and normal *Hira* mRNA was associated a down-regulation of HIRA protein (**Fig. 6d,g, Supplementary fig. 11c-d**). Importantly, dysregulation of *Hira* RNA and HIRA protein correlated with a pattern of maternal phenotype inheritance.

DISCUSSION

We investigated potential mechanisms contributing to epigenetic inheritance in *Mtrr*^{gt} mice, a unique model of mammalian TEI[3]. In the *Mtrr*^{gt} model, inheritance of developmental phenotypes and epigenetic instability occurs via the maternal lineage, though the mechanism is complex because an *Mtrr*$^{+/gt}$ female or male can initiate the effect through their daughters and granddaughters[3]. Due to its experimental tractability, we assessed DNA methylation in spermatozoa to understand how the germline epigenome was affected by the *Mtrr*^{gt} allele. We identified several distinct DMRs in regions of predicted transcriptional regulation, nucleosome retention and reprogramming resistance. This result illustrates widespread epigenetic instability in the male germline of the *Mtrr*^{gt} model, particularly in the F0 *Mtrr*$^{+/gt}$ males of the MGF pedigree. While largely resolved in somatic tissue of subsequent wildtype generations, some germline DMRs were associated with transcriptional changes in the F1-F3 progeny. This result is compatible with another model of epigenetic inheritance, which displayed association between the germline DMRs and local gene expression in embryos despite resolution of the DMR[4]. We suggest that an epigenetic memory associated with a germline DMR persists for at least three generations. Furthermore, the histone chaperone HIRA emerged as a biomarker and potential mediator of maternal epigenetic inheritance.

Separating genetic from epigenetic factors in this and other TEI studies is challenging. In our model, TEI is initiated in the F0 generation by the *Mtrr*^{gt} mutation that generates a metabolic defect linked to cellular methylation[3,12]. Though subsequent generations were genetically wildtype and metabolically normal[24], widespread dysregulation of DNA methylation was evident alongside several developmental phenotypes[3]. Whole genome sequencing of *Mtrr*$^{gt/gt}$ embryos discounted an alternate single mutation in the *Mtrr*^{gt} mouse line that might contribute to phenotypes. Moreover, folate metabolism is important for thymidine synthesis[52] and folate deficiency might trigger increased DNA breaks due to uracil misincorporation as demonstrated in erythrocytes of splenectomised patients[53] and prostate adenoma cells[54]. We excluded genetic instability in *Mtrr*^{gt} mice because de novo mutations occurred at a comparable frequency in control C57Bl/6J and *Mtrr*$^{gt/gt}$ embryos relative to the reference genome. Moreover, comparison of our whole genome sequencing and MeDIP-seq data sets determined that most sperm DMRs identified were true

methylation changes instead of an outcome of underlying genetic variation. Therefore, the epigenetic consequences of the *Mtrr*^{gt} allele rather than genetic instability are the likely instigator of TEI in this model.

Our data show that inheritance of sperm DMRs into offspring somatic tissue is unlikely, as have other studies[4]. Instead, somatic cell lineages might inherit germline epigenetic instability in a broader sense. For instance, genomic regions identified as sperm DMRs in F0 *Mtrr*$^{+/gt}$ males showed normal methylation patterns in embryos and placentas of F1-F3 wildtype offspring at midgestation. However, widespread epigenetic instability is still evident in F1 and F2 wildtype placentas[3]. It is possible that abnormalities in the sperm epigenome of F0 *Mtrr*$^{+/gt}$ males might be reprogrammed and then stochastically abnormally re-established/maintained in other genomic regions in a cell-type specific manner in wildtype offspring. This hypothesis could explain why many distinct phenotypes result in the F2-F3 generations. Apart from DNA methylation, this phenomenon might be linked to dysregulation of epigenetic regulators, such as HIRA[49] (see below), and/or other epigenetic mechanisms, such as histone modifications[10] or ncRNA expression[1,9].

Dietary folate deficiency causes differential methylation in sperm[55], though whether it causes TEI is unknown. Similar phenotypes appear in *Mtrr*^{gt} mutation in mice and in folate-deficient humans[13,20-23], but have not been reported in genetically wildtype mice fed folate-deficient diets[56,57]. There was no overlap between sperm DMRs in the diet model versus *Mtrr*$^{+/gt}$ males, which disputes the existence of folate-specific epigenomic hotspots in sperm. Severity of insult or technical differences (e.g., MeDIP-array[55] versus MeDIP-seq) might explain this discrepancy.

Whether DNA methylation patterns observed in F0 sperm are reconstructed[58,59] in the F1 germline is yet-to-be determined in the MGF pedigree. Exploring oocyte methylation has significant challenges and comparing DNA methylation patterns in F0 sperm to F1 oocytes might be incongruous. However, we showed that sperm of *Mtrr*$^{+/+}$ males derived from *Mtrr*$^{+/gt}$ intercrosses exhibited several DMRs that were independent of genetic variation yet overlapped with sperm DMRs in *Mtrr*$^{+/gt}$ males (representing their fathers). Therefore, reconstruction of specific atypical F0 germline methylation patterns likely occurs in the F1 germline in the *Mtrr*^{gt} model. In contrast, vinclozolin toxicant exposure of rats results in TEI and dissimilar DMRs in spermatozoa of F1 and

F3 offspring[8]. However, epigenetic patterns might shift as generational distance from the F0 individual increases. The effects of vinclozolin on testis formation[60], not observed in the *Mtrr*^{gt} model[61], might explain the discrepancies between TEI models.

Hira transcription emerged as a biomarker of maternal phenotypic inheritance even though *Hira* was identified through an associated sperm DMR. Specifically, embryos derived from wildtype oocytes of *Mtrr*$^{+/gt}$ females (or of wildtype females with *Mtrr*^{gt} ancestry) displayed a broad spectrum and frequency of developmental phenotypes alongside abnormal *Hira* mRNA expression at E10.5. In contrast, embryos derived from wildtype sperm of *Mtrr*$^{+/gt}$ males displayed normal phenotypes and *Hira* mRNA levels, yet *Hira* lncRNA was misexpressed at E10.5. The HIRA histone chaperone complex is important for histone deposition/recycling to maintain chromatin integrity during transcription[62]. The function of *Hira* lncRNA is unknown, though lncRNA-based mechanisms often control cell fates during development by influencing nuclear organization and transcriptional regulation[63]. Regardless, *Hira*$^{-/-}$ embryos[51] phenocopy embryos in the *Mtrr*^{gt} model[3] suggesting general epigenetic instability is detrimental to embryo development, and that *Hira* transcription might be responsive to epigenetic instability caused by the *Mtrr*^{gt} allele.

It is of interest that F1 progeny are phenotypically different when derived from an F0 *Mtrr*$^{+/gt}$ male versus F0 *Mtrr*$^{+/gt}$ female[3] (**Supplementary fig. 2a**). The difference might relate to a maternally-supplied factor present in ooplasm that is cytoplasmically-inherited by the F1 wildtype zygote (**Fig. 7**). Maternal *Hira* mRNA and protein are present in the oocyte and zygote[49,64], where it is involved in protamine replacement by histones in the paternal pronucleus[50]. Therefore, HIRA is a maternal factor suitably placed to perpetuate epigenetic instability between generations in the *Mtrr*^{gt} model. Altogether, dysregulation of *Hira* and/or other maternal factors in oocytes might have a greater impact on chromatin integrity in the early embryo than when dysregulated in sperm.

Overall, our study shows the long-term impact of abnormal folate metabolism on DNA methylation in the germline and emphasizes the importance of the *Mtrr*^{gt} mouse model in understanding the complex molecular and epigenetic mechanisms involved in TEI. Our data further suggest the importance of normal folate intake in both women and men of reproductive age for healthy pregnancies in their daughters and granddaughters.

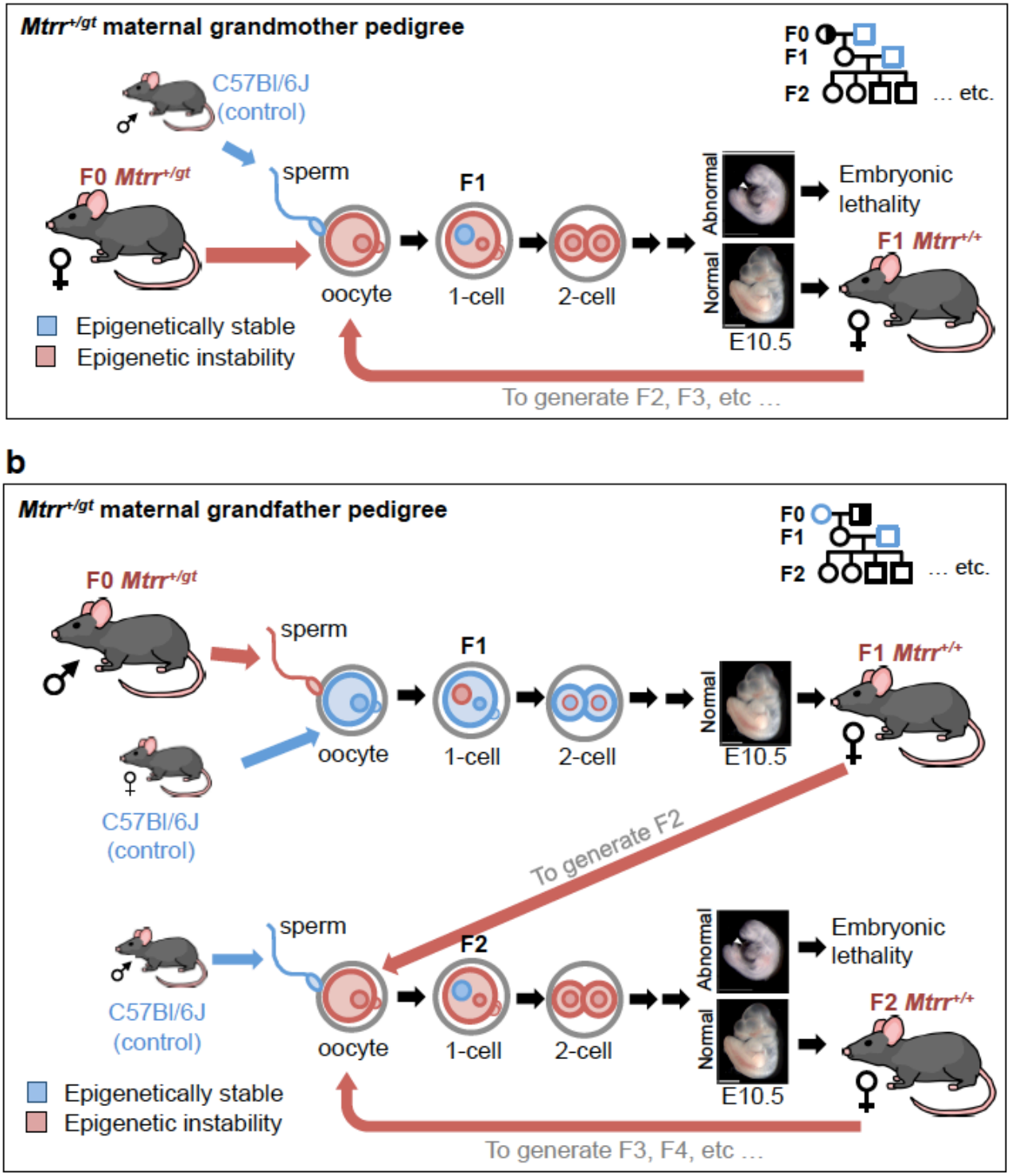


Fig. 7. Proposed model of maternal inheritance of epigenetic instability in *Mtrr*^{gt} model. Model suggesting how epigenetic instability generated by the *Mtrr*^{gt} mutation might be differently inherited over multiple generations depending upon whether the (**a**) maternal grandmother or (**b**) maternal grandfather is a carrier for the *Mtrr*^{gt} allele. Pedigree legend: circle, female; square, male; blue outline, C57Bl/6J control; black outline, *Mtrr*^{gt} mouse line; white filled, *Mtrr*$^{+/+}$; half-white/half-black filled, *Mtrr*$^{+/gt}$.

METHODS

Ethics statement. This research was regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body.

Animal model. *Mtrr*$^{Gt(XG334)Byg}$ (MGI:3526159), referred to as *Mtrr*^{gt} mice, were generated as previously described[3]. Briefly, a *β-geo* gene-trap (gt) vector was inserted into intron 9 of the *Mtrr* gene in 129P2Ola/Hsd embryonic stem cells (ESCs). *Mtrr*^{gt} ESCs were injected into C57Bl/6J blastocysts. Upon germline transmission, the *Mtrr*^{gt} allele was backcrossed into the C57Bl/6J genetic background for at least eight generations[3]. *Mtrr*$^{+/+}$ and *Mtrr*$^{+/gt}$ mice were produced from *Mtrr*$^{+/gt}$ intercrosses. *Mtrr*$^{gt/gt}$ mice were produced from *Mtrr*$^{gt/gt}$ intercrosses. C57Bl/6J mice from The Jackson Laboratories (www.jaxmice.jax.org) and 129P2Ola/Hsd from Envigo (previously Harlan Laboratories [www.envigo.com]) were used as controls and were bred in house and separately from the *Mtrr*^{gt} mouse line. All mice were fed a normal chow diet (Rodent No. 3 breeding chow, Special Diet Services) *ad libitum* from weaning. A detailed breakdown of the diet was reported previously[3]. Mice were euthanized via cervical dislocation. Genotyping for the *Mtrr*^{gt} allele and/or sex was performed using PCR on DNA extracted from ear tissue or yolk sac as previously described[3,65,66].

To determine the multigenerational effects of the *Mtrr*^{gt} allele in the maternal grandfather (MGF), the following mouse pedigree was established (**Supplementary Fig. 2a**). For the F1 generation, F0 *Mtrr*$^{+/gt}$ males were mated with C57Bl/6J females and the resulting *Mtrr*$^{+/+}$ progeny were analysed. For the F2 generation, F1 *Mtrr*$^{+/+}$ females were mated with C57Bl/6J males and the resulting *Mtrr*$^{+/+}$ progeny were analysed. For the F3 generation, F2 *Mtrr*$^{+/+}$ females were mated with C57Bl/6J males and the resulting *Mtrr*$^{+/+}$ progeny was analysed. A similar pedigree was established to assess the effects of the *Mtrr*^{gt} allele in the maternal grandmother (MGM) with the exception of the F0 generation, which involved the mating of an *Mtrr*$^{+/gt}$ female with a C57Bl/6J male.

Tissue dissection and phenotyping. For embryo and placenta collection, timed matings were

established and noon on the day that the vaginal plug was detected was considered embryonic day (E) 0.5. Embryos and placentas were dissected at E10.5 in cold 1x phosphate buffered saline and were scored for phenotypes (see below), photographed, weighed, and snap frozen in liquid nitrogen for storage at -80°C. Livers were collected from pregnant female mice (gestational day 10.5), weighed and snap frozen in liquid nitrogen for storage at -80°C.

A rigorous phenotyping regime was performed at E10.5 as previously described[3]. Briefly, all conceptuses were scored for one or more congenital malformation including failure of the neural tube to close in the cranial or spinal cord region, malformed branchial arches, pericardial edema, reversed heart looping, enlarged heart, and/or off-centered chorioallantoic attachment. Twinning or haemorrhaging was also scored as a severe abnormality. Embryos with <30 somite pairs were considered developmentally delayed. Embryos with 30-39 somite pairs but a crown-rump length more than two standard deviations (SD) from the mean crown-rump length of C57Bl/6J control embryos were considered growth restriction or growth enhanced. Conceptuses were considered phenotypically-normal if they were absent of congenital malformations, had 30-39 somite pairs and had crown-rump lengths within two SD of controls.

Spermatozoa collection. Spermatozoa from cauda epididymides and vas deferens were collected from 16-20 week-old fertile mice as previously described[67] with the following amendments. Spermatozoa were released for 20 minutes at 37°C in Donners Medium (25 mM NaHCO3, 20 mg/ml bovine serum albumin (BSA), 1 mM sodium pyruvate and 0.53% (vol/vol) sodium dl-lactate in Donners stock (135 mM NaCl, 5 mM KCl, 1 mM MgSO4, 2 mM $CaCl_2$ and 30 mM HEPES)). Samples were centrifuged at 500 x g (21°C) for 10 minutes. The supernatant was transferred and centrifuged at 1,300 x *g* (4°C) for 15 minutes. After the majority of supernatant was discarded, the samples were centrifuged at 1,300 x *g* (4°C) for 5 minutes. Further supernatant was discarded and the remaining spermatozoa were centrifuged at 12,000 x *g* for 1 minute and stored at -80°C.

Nucleic Acid Extraction. For embryo, trophoblast and liver tissue, genomic DNA (gDNA) was extracted using DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions. RNA was extracted from tissues using the AllPrep DNA/RNA Mini Kit (Qiagen). For

sperm, Solution A (75 mM NaCl pH 8; 25 mM EDTA) and Solution B (10 mM Tris-HCl pH 8; 10 mM EDTA; 1% SDS; 80 mM DTT) were added to the samples followed by RNAse A incubation (37°C, 1 hour) and Proteinase K incubation (55°C, overnight) as was previously described (Radford *et al.* 2014). DNA was extracted using phenol/chloroform/isoamyl alcohol mix (25:24:1) (Sigma-Aldrich) as per the manufacturer's instructions. DNA was precipitated using 10 M ammonium acetate, 0.1 mg/ml glycogen, and 100% ethanol and incubated at -80°C for at least 30 minutes. DNA was collected by centrifugation (13,000 rpm, 30 minutes). The pellet was washed twice in 70% ethanol, air-dried, and resuspended in TE buffer. DNA quality and quantity was confirmed using gel electrophoresis and QuantiFluor dsDNA Sample kit (Promega) as per the manufacturer's instructions.

Whole Genome Sequencing (WGS). DNA was extracted from two whole C57Bl/6J embryos at E10.5 (one male, one female) and six whole *Mtrr*$^{gt/gt}$ embryos with congenital malformations at E10.5 (four males, two females). Non-degraded gDNA was sent to BGI (Hong Kong) for library preparation and sequencing. Sequencing was performed with 150 bp paired-end reads on an Illumina HiSeq X machine. Quality control of reads assessed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters and low quality bases removed using Trim Galore (https://www.bioinformatics.babraham.ac.uk/projects/trim galore/). Summary metrics were created across all samples using the MultiQC package (http://multiqc.info)[68]. Sequencing reads were aligned to the C57Bl/6J reference genome (GRCm38, mm10) using BowTie2 with default parameters (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml)[69]. Duplicates were marked using Picard (http://broadinstitute.github.io/picard).

Structural variant analysis was performed using Manta[70]. Structural variants (SVs) were filtered using vcftools (version 0.1.15)[71]. In order to identify single nucleotide polymorphisms (SNPs), the data was remapped to the *mm10* reference mouse genome using BWA (version 0.7.15-r1144- dirty)[72]. Reads were locally realigned and SNPs and short indels identified using GenomeAnalysisTK (GATK, version 3.7)[73]. Homozygous variants were called when more than 90% of reads at the locus supported the variant call, whereas variants with at least 30% of reads

supporting the variant call were classified as heterozygous. Two rounds of filtering of variants were performed as follows. Firstly, low quality and biased variant calls were removed. Secondly, variants with: i) simple repeats with a periodicity <9 bp, ii) homopolymer repeats >8 bp, iii) dinucleotide repeats >14 bp, iv) low mapping quality (<40), v) overlapping annotated repeats or segmental duplications, and vi) >3 heterozygous variants fell within a 10 kb region were removed using vcftools (version 0.1.15) as was previously described[74]. The 129P2/OlaHsd mouse genome variation data was downloaded from Mouse Genomes Project[75].

Methylated DNA immunoprecipitation and sequencing (MeDIP-Seq). MeDIP-seq was carried out as described previously[4,76]. Briefly, 3 µg of sperm gDNA was sonicated using a Diagenode Bioruptor UCD-200 to yield 200-700 bp fragments that were end-repaired and dA-tailed. Illumina adaptors for paired-end sequencing were ligated using the NEB Next DNA Library Prep Master Mix for Illumina kit (New England Biolabs). After each step, the DNA was washed using Agencourt AMPure XP SPRI beads (Beckman Coulter). Immunoprecipitations (IPs) were performed in triplicate using 500 ng of DNA per sample, 1.25 µl of mouse anti-human 5mC antibody (clone 33D3; Eurogentec Ltd., Cat No. BI-MECY, RRID:AB_2616058), and 10 µl of Dynabeads coupled with M-280 sheep anti-mouse IgG bead (Invitrogen). The three IPs were pooled and purified using MinElute PCR Purification columns (Qiagen). Libraries were amplified by PCR (12 cycles) using Phusion High-Fidelity PCR Master Mix and adaptor specific iPCR tag primers (**Supplementary File 1**), and purified using Agencourt AMPure XP SPRI beads. The efficiency of the IP was verified using qPCR to compare the enrichment for DNA regions of known methylation status (e.g., methylated in sperm: *H19* and *Peg3* ICR, **Supplementary Fig. 4a**)[55] in the pre-amplification input and the IP fractions. MeDIP library DNA concentrations were estimated using the Kapa Library Quantification kit (Kapa Biosystems) and were further verified by running on an Agilent High Sensitivity DNA chip on an Agilent 2100 BioAnalyzer. Sequencing of MeDIP libraries was performed using 100 bp paired-end reads on an Illumina HiSeq platform at the Babraham Institute Next Generation Sequencing Facility (Cambridge, UK).

Quality assessment of the sequencing reads was performed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adaptor trimming was performed

using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim galore/). Reads were mapped to the GRCm38 (mm10) reference genome using Bowtie2 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml)[69]. All programmes were run with default settings unless otherwise stated. Sample clustering was assessed using principle component analysis (PCA), using the 500 most variable windows with respect to read coverage (as a proxy for methylation) for 5 kb window across all samples. Further data quality checks and differential methylation analysis was performed using the MEDIPS package in R[28]. The following key parameters were defined: BSgenome = BSgenome.Mmusculus.UCSC.mm10, uniq = 1e-3, extend = 300, ws = 500, shift=0. Differentially methylated regions (DMRs) were defined as windows (500 bp) in which there was at least 1.5-fold difference in methylation (reads per kilobase million mapped reads (RPKM)) between C57Bl/6J and *Mtrr* sperm methylation level with a p-value <0.01. Adjacent windows were merged using BEDTools (version 2.27.0)[77]. The background methylome was defined as all 500 bp windows across the genome at which the sum of the average RPKM per genotype group was >1.0. The genomic localisations of DMRs including association with coding/non-coding regions and CpG islands was determined using annotation downloaded from University of California, Santa Cruz (UCSC)[78]. The percentage of DMRs associated with repetitive regions of the genome was calculated using RepeatMasker software (http://www.repeatmasker.org).

Enrichment analysis using published ChIP-seq and ATAC-seq data sets. To identify enrichment of specific histone modifications and THSS, CTCF, TET1 and protamine1 (PRM1) binding in the DMR regions, published data sets were analysed included ChIP-seq data in CD1 spermatozoa collected from cauda epididymis[37], ESCs[42] and TSCs[42], and ATAC-seq data in CD1 spermatozoa collected from cauda epididymis[37] and B6D2F1 epiblast and extraembryonic ectoderm (ExE)[38] at E6.5. The source and accession numbers of processed ChIP-seq and ATAC-seq wig/bigwig files are shown in **Supplementary File 2** and accessible on GitHub (https://github.com/CTR-BFX/Blake_Watson). To ensure that the analysis was consistent across public data sets, all wig files were converted to bigwig using UCSC tools “wigToBigWig -clip” (http://hgdownload.soe.ucsc.edu/admin/exe/).

DMRs identified in sperm from all three *Mtrr* genotypes (MeDIP-seq analysis) were

combined to generate a list of 893 DMRs. To prevent the inclusion of DMRs associated with genomic variation, the 20 Mb region surrounding the *Mtrr* gene (Chr13:58060780-80060780) that was identified as 129P2Ola/Hsd genomic sequence was masked. This resulted in 459 DMRs for subsequent analysis. The ChIP-seq and ATAC-seq files in sperm[37] were originally aligned to mouse reference genome mm9. The files were converted to mouse reference genome mm10 to match the other published data sets in this analysis using hgLiftOver with the mm9ToMm10.over.chain (http://genome.ucsc.edu/cgi-bin/hgLiftOver). To identify the baseline enrichment profiles around the DMRs for specific histone modifications, THSS, CTCF, TET1, or PRM1, a similar number of genomic regions were randomly selected using bedtools (v2.26.0)[77] with the following command: "bedtools2/2.26.0, bedtools shuffle -i DMRs.bed -g Mus_GRCm38_chrall.fa.fai -chrom -seed 27442958 -noOverlapping -excl Mtrr_mask20Mb.bed". The DMR profiles were created using deeptools (version 2.3.1)[79], via computeMatix 3 kb scaled windows, flanking regions of 6 kb, and a bin size of 200 bp, and plotted with plotProfile.

Quantitative reverse transcription PCR (RT-qPCR). For RNA expression analysis, cDNA was synthesised using RevertAid H Minus reverse transcriptase (Thermo Scientific) and random hexamer primers (Thermo Scientific) using 1-2 µg of RNA in a 20-µl reaction according to manufacturer's instructions. PCR amplification was conducted using MESA Green qPCR MasterMix Plus for SYBR Assay (Eurogentec Ltd.) on a DNA Engine Opticon2 thermocycler (BioRad). The following cycling conditions were used: 95°C for 10 minutes, 40 cycles: 95°C for 30 seconds, 60°C for 1 minute, followed by melt curve analysis. Transcript levels were normalised to *Hprt* and/or *Gapdh* RNA levels. Relative cDNA expression levels were analysed as previously described[80]. Transcript levels in C57Bl/6J tissue were normalized to 1. For primer sequences and concentrations, refer to **Supplementary File 1**.

Bisulfite mutagenesis and pyrosequencing. Between 250 ng and 2 µg of gDNA extracted from each tissue sample was bisulfite treated with an Imprint DNA Modification Kit (Sigma). Control samples lacking DNA template were run to ensure that there was no contamination during the bisulfite conversion of DNA. To quantify DMR CpG methylation, pyrosequencing was performed.

50 ng of bisulfite-converted DNA was used as template for PCR, together with 0.2 uM of each biotinylated primer and 0.25 units of HotStarTaq PlusDNA Polymerase (Qiagen). Refer to **Supplementary File 1** for primer sequences, which were designed using PyroMark Assay Design Software 2.0 (Qiagen). PCR was performed in triplicate using the following conditions: 95°C for 5 minutes, 40 cycles of 94°C for 30 seconds, 56°C for 30 seconds, 72°C for 55 seconds, and then 72°C for 5 minutes. PCR products were purified using Strepdavidin Sepharose High Performance beads (GE healthcare). The beads bound to DNA were washed in 70% ethanol, 0.4 M NaOH and 10 mM Tris-acetated (pH 7.6) and then hybridized to the sequencing primer in PyroMark annealing buffer (Qiagen) according to the manufacturer's instructions. Pyrosequencing was conducted using PyroMark Gold reagents kit (Qiagen) on a PyroMark MD pyrosequencer (Biotage). The mean CpG methylation was calculated using three to eight biological replicates and at least two technical replicates. Analysis of methylation status was performed using Pyro Q-CpG software (Biotage).

Mass spectrometry. Sperm gDNA was digested into individual nucleoside components using the DNA Degradase Plus kit (Zymo Research) according to the manufacturer's instructions. The heat inactivation step was omitted. 100 ng of degraded DNA per individual was sent to the Babraham Institute Mass Spectrometry Facility (Cambridge, UK), where global cytosine, 5mC and 5- 5hmC was determined by liquid chromatography-tandem mass spectrometry (LC-MS/MS) as previously described[81]. Sperm of nine males were assessed per genotype and analysed in pools each containing three unique individuals. All pooled samples were analysed in triplicate. Global 5mC and 5hmC levels are reported as percentages relative to C.

Western blotting. Embryos and placentas at E10.5 were homogenised in lysis buffer (20 mM Tris [pH 7.5], 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 2.5 mM sodium pyrophosphate, 1 mM β-glycerolphosphate, 1 mM Na_3VO_4 and complete mini EDTA-free proteases inhibitor cocktail [Roche Diagnostics]) with Lysing Matrix D ceramic beads (MP Biomedical) using a MagNA Lyser (Roche Diagnostics) at 5,500 rpm for 20 seconds. Samples were incubated on ice for 5 minutes and then homogenized again at 5,500 rpm for 20 seconds. Homogenates were then incubated on ice for 20 minutes with brief intervening vortexing steps

occurring every 5 minutes. Samples were then centrifuged at 10,000 x *g* for 5 minutes. Supernatant from each sample was transferred to a new tube and centrifuged again at 10,000 x *g* for 5 minutes to ensure that all residual tissue was removed. Protein concentration of tissue lysates was determined using bicinchoninic acid (Sigma-Aldrich). Proteins were denatured with gel loading buffer (50 mM Tris [pH 6.8], 100 mN DTT, 2% SDS, 10% glycerol and trace amount of bromophenol blue) at 70°C for 10 minutes. Equivalent amounts of protein were resolved by 8-10% SDS-PAGE and blotted onto nitrocellulose (0.2 µm, Amersham Protran) with a semi-dry blotter (GE Healthcare). The membrane was stained with Ponceau S solution (Sigma-Aldrich) and the resulting scanned image was used as a loading control. After washing, the membrane was blotted with 5% skimmed milk in Tris buffered saline containing 0.1% Tween-20 (TBS-T) before incubation with 1:1,000 dilution of monoclonal rabbit anti-human HIRA (clone D2A5E, Cell Signalling Technology, Cat. No. 13307, RRID:AB_2798177) overnight at 4°C or 1:10,000 dilution of monoclonal mouse anti-human β-actin (clone AC-74, Sigma-Aldrich, Cat. No. A2228, RRID:AB_476997) for 1 hour at room temperature. Primary antibodies were diluted in TBS-T. The membrane was incubated with 1:10,000 dilution of anti-rabbit IgG conjugated to horse radish peroxidase (HRP; GE Healthcare) diluted in 2.5% skimmed milk in TBS-T or anti-mouse IgG conjugated to HRP (GE Healthcare) diluted in TBS-T. The signal of resolved protein was visualized by Amersham enhanced chemiluminescence (ECL) Western Blotting Analysis System (GE Healthcare) using Amersham Hyperfilm ECL (GE Healthcare). A flat-bed scanner (HP Scanjet G4050) was used to scan films. Band intensities were determined with background subtraction using ImageJ (64-bit) software (NIH, USA).

Statistical analysis. Statistical analysis was performed using GraphPad Prism software (version 7). RT-qPCR data were analysed by independent unpaired t tests or ordinary one-way ANOVA with Dunnett's or Sidak's multiple comparison testing. SV and SNP data were analysed by independent unpaired t tests with Welch's correction. Bisulfite pyrosequencing data, SV chromosome frequency and DMR distribution at repetitive elements were analysed by two-way ANOVAs with Dunnett's, Sidak's or Tukey's multiple comparisons tests. Western blot data were analysed using a non-parametric Kruskal-Wallis test with Dunn's multiple comparisons test. $p<0.05$

was considered significant unless otherwise stated.

Data availability. All relevant data are available from the corresponding author upon reasonable request. Raw sequencing data have been deposited in ArrayExpress database at EMBL-EBI. The WGS data is under accession number E-MTAB-8513 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8513) and the raw MeDIP-Seq data is under accession number E-MTAB-8533 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8533).

Code availability. The in-house scripts used for the analysis can be found in the following online repository: https://github.com/CTR-BFX/Blake_Watson.

ACKNOWLEDGEMENTS

We thank Drs Nozomi Takahashi and Tessa Bertozzi for their technical support, Dr Claire Senner for critical discussion, and Drs Jamie Hackett and Miguel Branco for .bed files. The MeDIP-seq and LC-MS/MS was performed at the Babraham Institute (UK). WGS was performed by BGI (Hong Kong). This work was funded by grants from Lister Institute for Preventative Medicine and the Centre for Trophoblast Research (CTR) to E.D.W and from the MRC MR/J001597 and Wellcome Trust WT095606 (to A.C.F.-S). The following support was provided: Wellcome Trust 4-year DTP in Developmental Mechanisms (to G.E.T.B.) and CTR funding (to G.E.T.B., H.W.Y., X.Z., R.S.H.).

AUTHOR CONTRIBUTIONS

G.E.T.B. collected sperm, performed DNA/RNA extractions, generated WGS and MeDIP libraries, and performed RT-qPCR and bisulfite pyrosequencing analyses. G.E.T.B. and E.D.W. dissected tissue and phenotyped conceptuses. G.E.T.B. and E.D.W. collected and analysed the data. X.Z., R.S.H., and G.E.T.B. designed and performed bioinformatics analyses. H.W.Y. performed the western blotting analysis. E.D.W. conceived the project. G.E.T.B., A.C.F.-S., G.J.B., and E.D.W. designed the experiments and interpreted the results. E.D.W. and G.E.T.B. wrote the manuscript. All authors read and revised the manuscript.

COMPETING INTERESTS

None to declare.

REFERENCES

1. Chen, Q. *et al.* Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science* **351**, 397-400 (2016).
2. Gapp, K. *et al.* Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat Neurosci* **17**, 667-9 (2014).
3. Padmanabhan, N. *et al.* Mutation in folate metabolism causes epigenetic instability and transgenerational effects on development. *Cell* **155**, 81-93 (2013).
4. Radford, E.J. *et al.* In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. *Science* **345**, 1255903 (2014).
5. Anway, M.D., Cupp, A.S., Uzumcu, M. & Skinner, M.K. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* **308**, 1466-9 (2005).
6. Blake, G.E. & Watson, E.D. Unravelling the complex mechanisms of transgenerational epigenetic inheritance. *Curr Opin Chem Biol* **33**, 101-7 (2016).
7. Heard, E. & Martienssen, R.A. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* **157**, 95-109 (2014).
8. Beck, D., Sadler-Riggleman, I. & Skinner, M.K. Generational comparisons (F1 versus F3) of vinclozolin induced epigenetic transgenerational inheritance of sperm differential DNA methylation regions (epimutations) using MeDIP-Seq. *Environ Epigenet* **3**(2017).
9. Sharma, U. *et al.* Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science* **351**, 391-396 (2016).
10. Siklenka, K. *et al.* Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science* **350**, aab2006 (2015).
11. Jimenez-Chillaron, J.C. *et al.* Intergenerational transmission of glucose intolerance and obesity by in utero undernutrition in mice. *Diabetes* **58**, 460-8 (2009).
12. Elmore, C.L. *et al.* Metabolic derangement of methionine and folate metabolism in mice deficient in methionine synthase reductase. *Mol Genet Metab* **91**, 85-97 (2007).
13. Leclerc, D. *et al.* Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria. *Proc Natl Acad Sci U S A* **95**, 3059-64 (1998).
14. Yamada, K., Gravel, R.A., Toraya, T. & Matthews, R.G. Human methionine synthase reductase is a molecular chaperone for human methionine synthase. *Proc Natl Acad Sci U S A* **103**, 9476-81 (2006).
15. Field, M.S., Kamynina, E., Chon, J. & Stover, P.J. Nuclear Folate Metabolism. *Annu Rev Nutr* **38**, 219-243 (2018).
16. Shane, B. & Stokstad, E.L. Vitamin B12-folate interrelationships. *Annu Rev Nutr* **5**, 115-41 (1985).
17. Ghandour, H., Lin, B.F., Choi, S.W., Mason, J.B. & Selhub, J. Folate status and age affect the accumulation of L-isoaspartyl residues in rat liver proteins. *J Nutr* **132**, 1357-60 (2002).
18. Jacob, R.A. *et al.* Moderate folate depletion increases plasma homocysteine and decreases lymphocyte DNA methylation in postmenopausal women. *J Nutr* **128**, 1204-12 (1998).
19. Wainfan, E., Moller, M.L., Maschio, F.A. & Balis, M.E. Ethionine-induced changes in rat liver transfer RNA methylation. *Cancer Res* **35**, 2830-5 (1975).
20. Wilson, A. *et al.* A common variant in methionine synthase reductase combined with low cobalamin (vitamin B12) increases risk for spina bifida. *Mol Genet Metab* **67**, 317-23 (1999).
21. Rosenblatt, D.S., Cooper, B.A., Schmutz, S.M., Zaleski, W.A. & Casey, R.E. Prenatal vitamin B12 therapy of a fetus with methylcobalamin deficiency (cobalamin E disease). *Lancet* **1**, 1127-9 (1985).

22. Schuh, S. *et al.* Homocystinuria and megaloblastic anemia responsive to vitamin B12 therapy. An inborn error of metabolism due to a defect in cobalamin metabolism. *N Engl J Med* **310**, 686-90 (1984).
23. Koury, M.J. & Ponka, P. New insights into erythropoiesis: the roles of folate, vitamin B12, and iron. *Annu Rev Nutr* **24**, 105-31 (2004).
24. Padmanabhan, N. *et al.* Abnormal folate metabolism causes age-, sex- and parent-of-origin-specific haematological defects in mice. *J Physiol* **596**, 4341-4360 (2018).
25. Padmanabhan, N. *et al.* Multigenerational analysis of sex-specific phenotypic differences at midgestation caused by abnormal folate metabolism. *Environ Epigenet* **3**, dvx014 (2017).
26. Kim, S. *et al.* PRMT5 protects genomic integrity during global DNA demethylation in primordial germ cells and preimplantation embryos. *Mol Cell* **56**, 564-79 (2014).
27. Taiwo, O. *et al.* Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat Protoc* **7**, 617-36 (2012).
28. Lienhard, M., Grimm, C., Morkel, M., Herwig, R. & Chavez, L. MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics* **30**, 284-6 (2014).
29. Orozco, L.D. *et al.* Epigenome-wide association of liver methylation patterns and complex metabolic traits in mice. *Cell Metab* **21**, 905-17 (2015).
30. Watkins-Chow, D.E. & Pavan, W.J. Genomic copy number and expression variation within the C57BL/6J inbred mouse strain. *Genome Res* **18**, 60-6 (2008).
31. Casas, E. & Vavouri, T. Sperm epigenomics: challenges and opportunities. *Front Genet* **5**, 330 (2014).
32. Brunner, A.M., Nanni, P. & Mansuy, I.M. Epigenetic marking of sperm by post-translational modification of histones and protamines. *Epigenetics Chromatin* **7**, 2 (2014).
33. Carone, B.R. *et al.* High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Dev Cell* **30**, 11-22 (2014).
34. Erkek, S. *et al.* Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat Struct Mol Biol* **20**, 868-75 (2013).
35. Kobayashi, H. *et al.* Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet* **8**, e1002440 (2012).
36. Hackett, J.A. *et al.* Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448-52 (2013).
37. Jung, Y.H. *et al.* Chromatin States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes. *Cell Rep* **18**, 1366-1382 (2017).
38. Smith, Z.D. *et al.* Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature* **549**, 543-547 (2017).
39. Seah, M.K.Y. & Messerschmidt, D.M. From Germline to Soma: Epigenetic Dynamics in the Mouse Preimplantation Embryo. *Curr Top Dev Biol* **128**, 203-235 (2018).
40. Tang, W.W., Kobayashi, T., Irie, N., Dietmann, S. & Surani, M.A. Specification and epigenetic programming of the human germ line. *Nat Rev Genet* **17**, 585-600 (2016).
41. Smallwood, S.A. *et al.* Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat Genet* **43**, 811-4 (2011).
42. Schoenfelder, S. *et al.* Divergent wiring of repressive and active chromatin interactions between mouse embryonic and trophoblast lineages. *Nat Commun* **9**, 4189 (2018).
43. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343-8 (2011).
44. Calabrese, J.M. *et al.* Site-specific silencing of regulatory elements as a mechanism of X inactivation. *Cell* **151**, 951-63 (2012).
45. Chrysanthou, S. *et al.* A Critical Role of TET1/2 Proteins in Cell-Cycle Progression of Trophoblast Stem Cells. *Stem Cell Reports* **10**, 1355-1368 (2018).
46. Chuong, E.B., Rumi, M.A., Soares, M.J. & Baker, J.C. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet* **45**, 325-9 (2013).
47. Gardiner-Garden, M. & Frommer, M. CpG islands in vertebrate genomes. *J Mol Biol* **196**, 261-82 (1987).
48. Pchelintsev, N.A. *et al.* Placing the HIRA histone chaperone complex in the chromatin landscape. *Cell Rep* **3**, 1012-9 (2013).

49. Nashun, B. *et al.* Continuous Histone Replacement by Hira Is Essential for Normal Transcriptional Regulation and De Novo DNA Methylation during Mouse Oogenesis. *Mol Cell* **60**, 611-25 (2015).
50. Lin, C.J., Koh, F.M., Wong, P., Conti, M. & Ramalho-Santos, M. Hira-mediated H3.3 incorporation is required for DNA replication and ribosomal RNA transcription in the mouse zygote. *Dev Cell* **30**, 268-79 (2014).
51. Roberts, C. *et al.* Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. *Mol Cell Biol* **22**, 2318-28 (2002).
52. Stover, P.J. One-carbon metabolism-genome interactions in folate-associated pathologies. *J Nutr* **139**, 2402-5 (2009).
53. Blount, B.C. *et al.* Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc Natl Acad Sci U S A* **94**, 3290-5 (1997).
54. Bistulfi, G., Vandette, E., Matsui, S. & Smiraglia, D.J. Mild folate deficiency induces genetic and epigenetic instability and phenotype changes in prostate cancer cells. *BMC Biol* **8**, 6 (2010).
55. Lambrot, R. *et al.* Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes. *Nat Commun* **4**, 2889 (2013).
56. Bills, N.D., Koury, M.J., Clifford, A.J. & Dessypris, E.N. Ineffective hematopoiesis in folate-deficient mice. *Blood* **79**, 2273-80 (1992).
57. Heid, M.K., Bills, N.D., Hinrichs, S.H. & Clifford, A.J. Folate deficiency alone does not produce neural tube defects in mice. *J Nutr* **122**, 888-94 (1992).
58. Jablonka, E. Epigenetic inheritance and plasticity: The responsive germline. *Prog Biophys Mol Biol* **111**, 99-107 (2013).
59. Miska, E.A. & Ferguson-Smith, A.C. Transgenerational inheritance: Models and mechanisms of non-DNA sequence-based inheritance. *Science* **354**, 59-63 (2016).
60. Uzumcu, M., Suzuki, H. & Skinner, M.K. Effect of the anti-androgenic endocrine disruptor vinclozolin on embryonic testis cord formation and postnatal testis development and function. *Reprod Toxicol* **18**, 765-74 (2004).
61. Blake, G.E.T., Hall, J., Petkovic, G.E. & Watson, E.D. Analysis of spermatogenesis and fertility in adult mice with a hypomorphic mutation in the Mtrr gene. *Reprod Fertil Dev* (2019).
62. Torne, J., Ray-Gallet, D., Boyarchuk, E., Garnier, M., Coulon, A., Orsi, G., Almouzni, G. Two distinct HIRA-dependent pathways handle H3.3 de novo deposition and recycling during transcription. (bioRxiv, 2019).
63. Batista, P.J. & Chang, H.Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**, 1298-307 (2013).
64. Park, S.J., Shirahige, K., Ohsugi, M. & Nakai, K. DBTMEE: a database of transcriptome in mouse early embryos. *Nucleic Acids Res* **43**, D771-6 (2015).
65. Jiang, Z. *et al.* Zic3 is required in the extra-cardiac perinodal region of the lateral plate mesoderm for left-right patterning and heart development. *Hum Mol Genet* **22**, 879-89 (2013).
66. Tunster, S.J. Genetic sex determination of mice by simplex PCR. *Biol Sex Differ* **8**, 31 (2017).
67. Hisano, M. *et al.* Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nat Protoc* **8**, 2449-70 (2013).
68. Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-8 (2016).
69. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-9 (2012).
70. Chen, X. *et al.* Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220-2 (2016).
71. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-8 (2011).
72. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-60 (2009).

73. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-303 (2010).
74. Oey, H., Isbel, L., Hickey, P., Ebaid, B. & Whitelaw, E. Genetic and epigenetic variation among inbred mouse littermates: identification of inter-individual differentially methylated regions. *Epigenetics Chromatin* **8**, 54 (2015).
75. Keane, T.M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-94 (2011).
76. Senner, C.E., Krueger, F., Oxley, D., Andrews, S. & Hemberger, M. DNA methylation profiles define stem cell identity and reveal a tight embryonic-extraembryonic lineage boundary. *Stem Cells* **30**, 2732-45 (2012).
77. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-2 (2010).
78. Meyer, L.R. *et al.* The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res* **41**, D64-9 (2013).
79. Ramirez, F., Dundar, F., Diehl, S., Gruning, B.A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**, W187-91 (2014).
80. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8 (2001).
81. Ficz, G. *et al.* FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**, 351-9 (2013).
82. Caubit, X. *et al.* Teashirt 3 is necessary for ureteral smooth muscle differentiation downstream of SHH and BMP4. *Development* **135**, 3301-10 (2008).
83. Friedrich, G.A., Hildebrand, J.D. & Soriano, P. The secretory protein Sec8 is required for paraxial mesoderm formation in the mouse. *Dev Biol* **192**, 364-74 (1997).
84. Millen, K.J., Wurst, W., Herrup, K. & Joyner, A.L. Abnormal embryonic cerebellar development and patterning of postnatal foliation in two mouse Engrailed-2 mutants. *Development* **120**, 695-706 (1994).
85. Rameix-Welti, M.A. *et al.* Visualizing the replication of respiratory syncytial virus in cells and in living mice. *Nat Commun* **5**, 5104 (2014).
86. Tuorto, F. *et al.* RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. *Nat Struct Mol Biol* **19**, 900-5 (2012).

SUPPLEMENTARY MATERIAL

Disruption of folate metabolism causes germline epigenetic instability and distinguishes HIRA as a biomarker of maternal transgenerational epigenetic inheritance

Georgina E.T. Blake[1,2,†], Xiaohui Zhao[1,2], Hong wa Yung[1,2], Graham J. Burton[1,2], Anne C. Ferguson-Smith[2,3], Russell S. Hamilton[2,3], Erica D. Watson[1,2,*]

[1]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge UK; [2]Centre for Trophoblast Research, University of Cambridge, Cambridge UK; [3]Department of Genetics, University of Cambridge, Cambridge, UK

[†]Current address: College of Medicine and Health, University of Exeter Medical School, Exeter, UK

*Correspondence: Dr Erica Watson (edw23@cam.ac.uk)

Supplementary table 1
Supplementary table 2
Supplementary figure 1
Supplementary figure 2
Supplementary figure 3
Supplementary figure 4
Supplementary figure 5
Supplementary figure 6
Supplementary figure 7
Supplementary figure 8
Supplementary figure 9
Supplementary figure 10
Supplementary figure 11
Supplementary file 1
Supplementary file 2

Supplementary Table 1. Summary of validation of sperm DMRs by bisulfite pyrosequencing

Male genotype*	**Number of DMRs that validated (%)**		**Total (%)**
	Hypomethylated	Hypermethylated	
Mtrr$^{+/+}$	0/0	3/3 (100%)	3/3 (100%)
Mtrr$^{+/gt}$	3/4 (75%)	10/20 (50%)	13/24 (54.2%)
Mtrr$^{gt/gt}$	13/13 (100%)	8/13 (61.5%)	21/26 (80.8%)
Total	16/17 (94.1%)	21/36 (58.3%)	37/53 (69.8%)

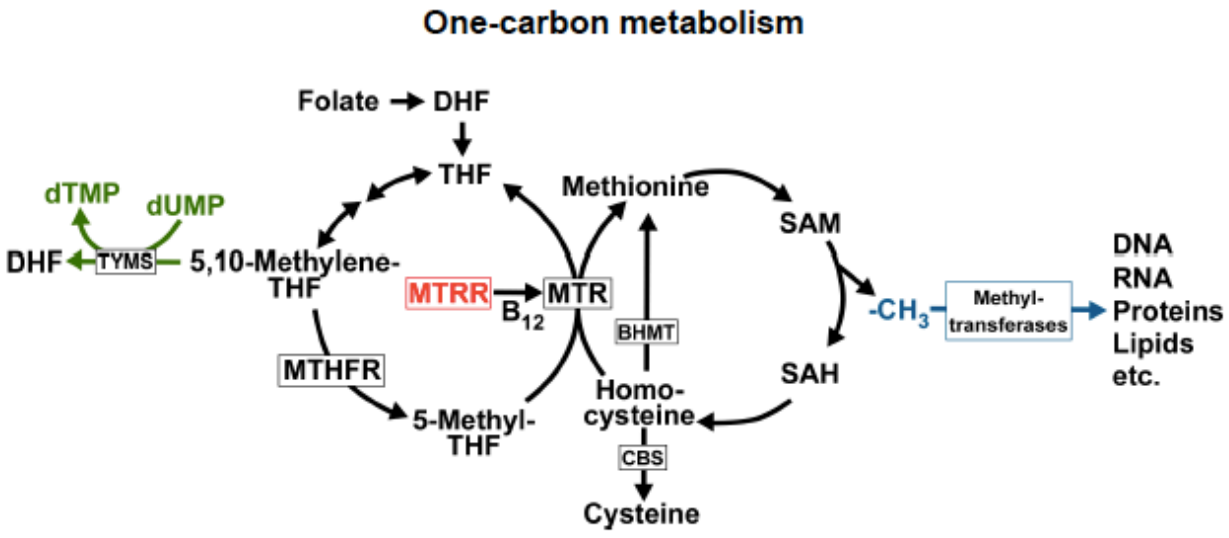
*N=8 males per genotypic group

Supplementary Table 2. Characteristics of sperm DMRs from *Mtrr*$^{+/gt}$ males that were assessed in F1-F2 wildtype somatic tissue.

DMR (Coordinates)	Genomic location	Reprogra-mming resistant region[35,36]	Methyl-ation status (sperm)	Closest gene(s)	Gene function	Gene knockout phenotype
DMR A10 (Chr16:18976001-18977000)	Intergenic	Yes	Hyper	*Hira*	Histone H3.3 chaperone	Failure of cardiac development, pericardial edema, abnormal neural tube morphology, embryo growth restriction, lethality by E10.5[51]
DMR D20 (Chr13:104660501-104661000)	Intragenic	No	Hypo	*Cwc27*	Spliceosome-associated protein; isomerase	Male infertility, pre-weaning lethality, abnormal blood homeostasis, postnatal growth restriction, retinal degeneration (IMPC database*)
DMR E115 (Chr7:36770501-36772000)	Intragenic	No	Hypo	*Tshz3*	Zinc finger transcription factor	Neonatal lethality likely due to respiratory distress, impaired uretic smooth muscle[82]
DMR E52 (Chr17:6324501-6325500)	Intergenic	Yes	Hyper	*Dynlt1a*	Dynein light chain	n/a
DMR D87 (Chr17:6324501-6325500)	Intergenic	Yes	Hyper	*4930572 O03Rik*	Unknown	n/a
				Speer7-cos	Spermatogenesis associated glutamate (E)-rich protein	n/a
DMR E112 (Chr16:33270501-33271000)	Intragenic	Yes	Hypo	*Exoc4*	Exocyst complex component	Embryonic abnormalities (e.g., heart defects and neural tube defects), gastrulation incomplete, embryonic lethality before E10.5[83]
DMR E109 (Chr5:28168501-28169000)	Intragenic	No	Hypo	*En2*	Homeobox transcription factor	Anterior-posterior cerebellar patterning defects, abnormal cerebellar foliation, motor defects[84]
DMR E28 (Chr14:13512001-13512500)	Intragenic	No	Hyper	*Synpr*	Synaptoporin	n/a

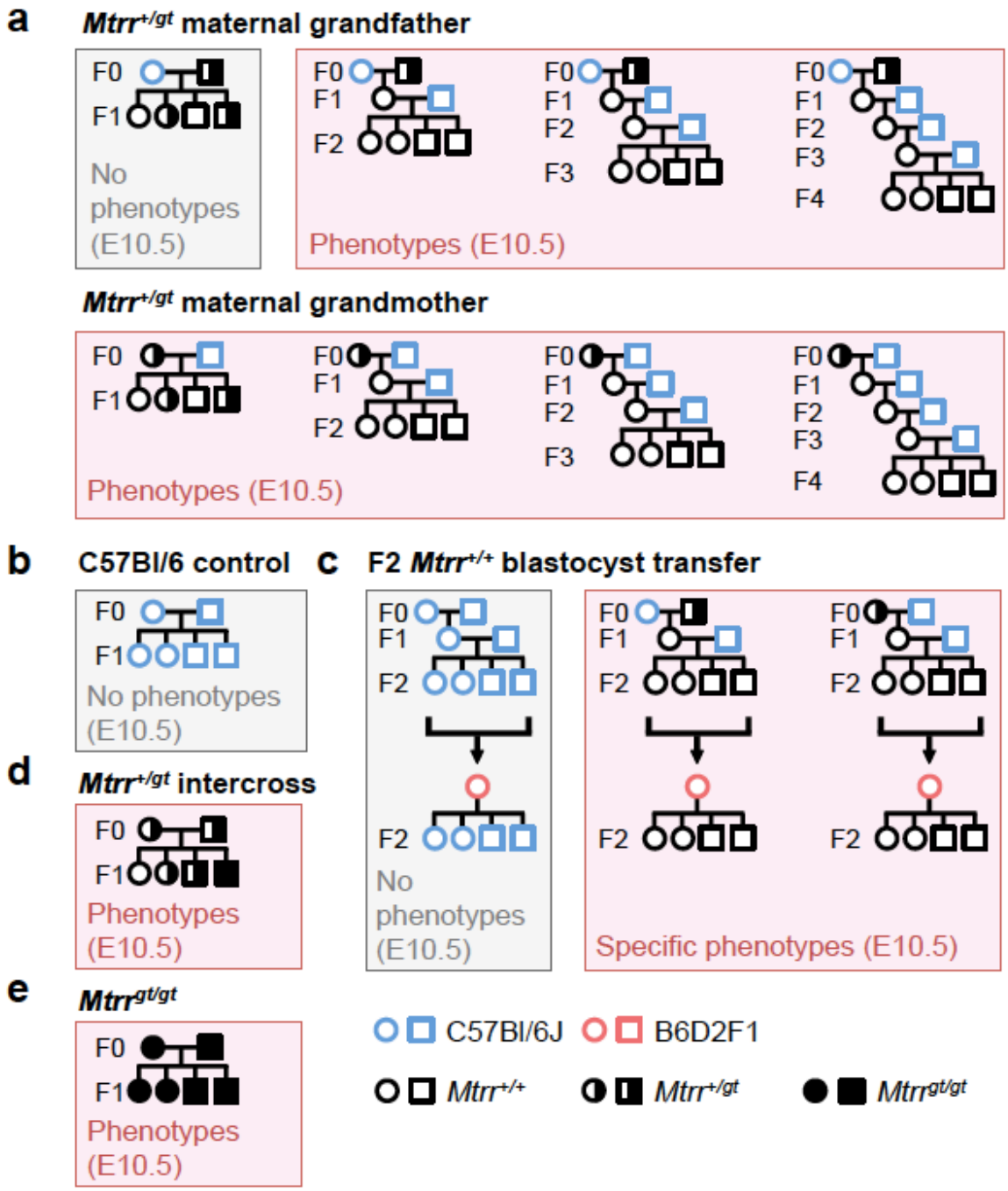
*International Mouse Phenotyping Consortium website: http://www.mousephenotype.org/

DMR, differentially methylated region; Hyper, hypermethylated; hypo, hypomethylated; n/a, not available.



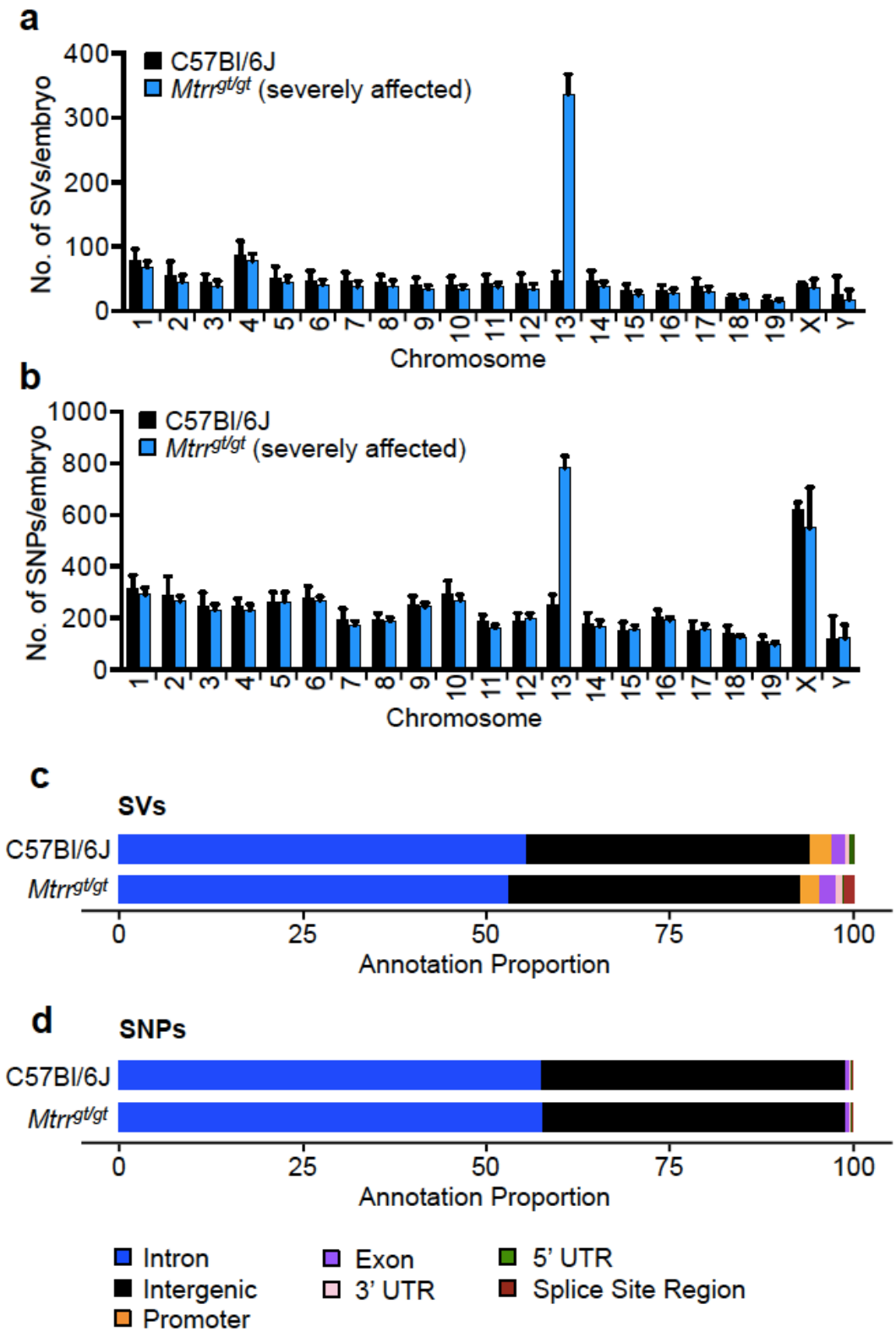
Supplementary Fig. 1. One-carbon metabolism.

Simplified schematic drawing of one-carbon metabolism[12]. MTRR (red) is a key enzyme required for progression of one-carbon metabolism, located at the intersection between folate and methionine pathways. Methionine synthase (MTR), which contains a cobalamin (B_{12}) cofactor, catalyses methyl (CH_3-) group transfer from methyl-tetrahydrofolate (5-methyl-THF) to homocysteine to form methionine. Every 1,000 turnovers, cob(I)alamin becomes oxidized to form inactive cob(II)alamin. MTRR provides an electron that, together with a CH_3- group from S-adenosyl-methionine (SAM), returns MTR to the active methylcobalamin form. BHMT, betaine-homocysteine S-methyltransferase; CBS, cystathionine beta-synthase; CH_3-, one-carbon methyl group; DHF, dihydrofolate; dTMP, deoxythymidine monophosphate; dUMP, deoyuridine monophosphate; MTHFR, methylenetetrahydrofolate reductase; MTR, methionine synthase; MTRR, methionine synthase reductase; SAH, S-adenosyl-homocysteine; SAM, S-adenosyl-methionine; THF, tetrahydrofolate; TYMS, thymidylate synthase.

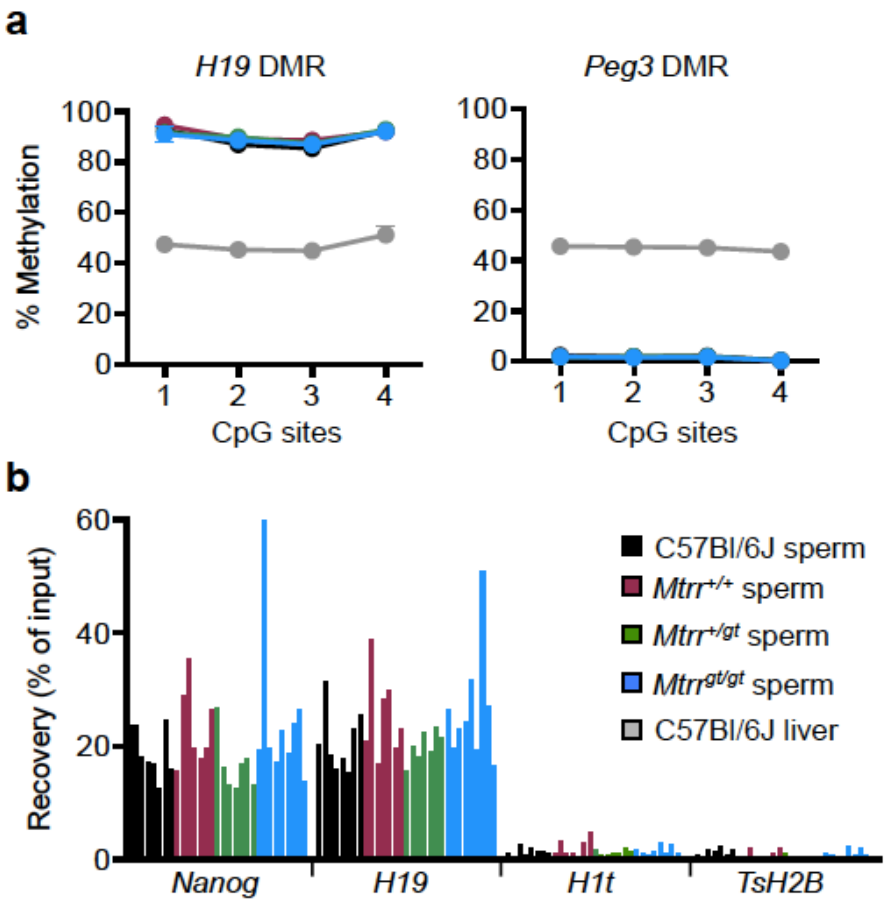


Supplementary Fig. 2. *Mtrr*^{gt} mouse line is a model of transgenerational epigenetic inheritance via the maternal lineage.

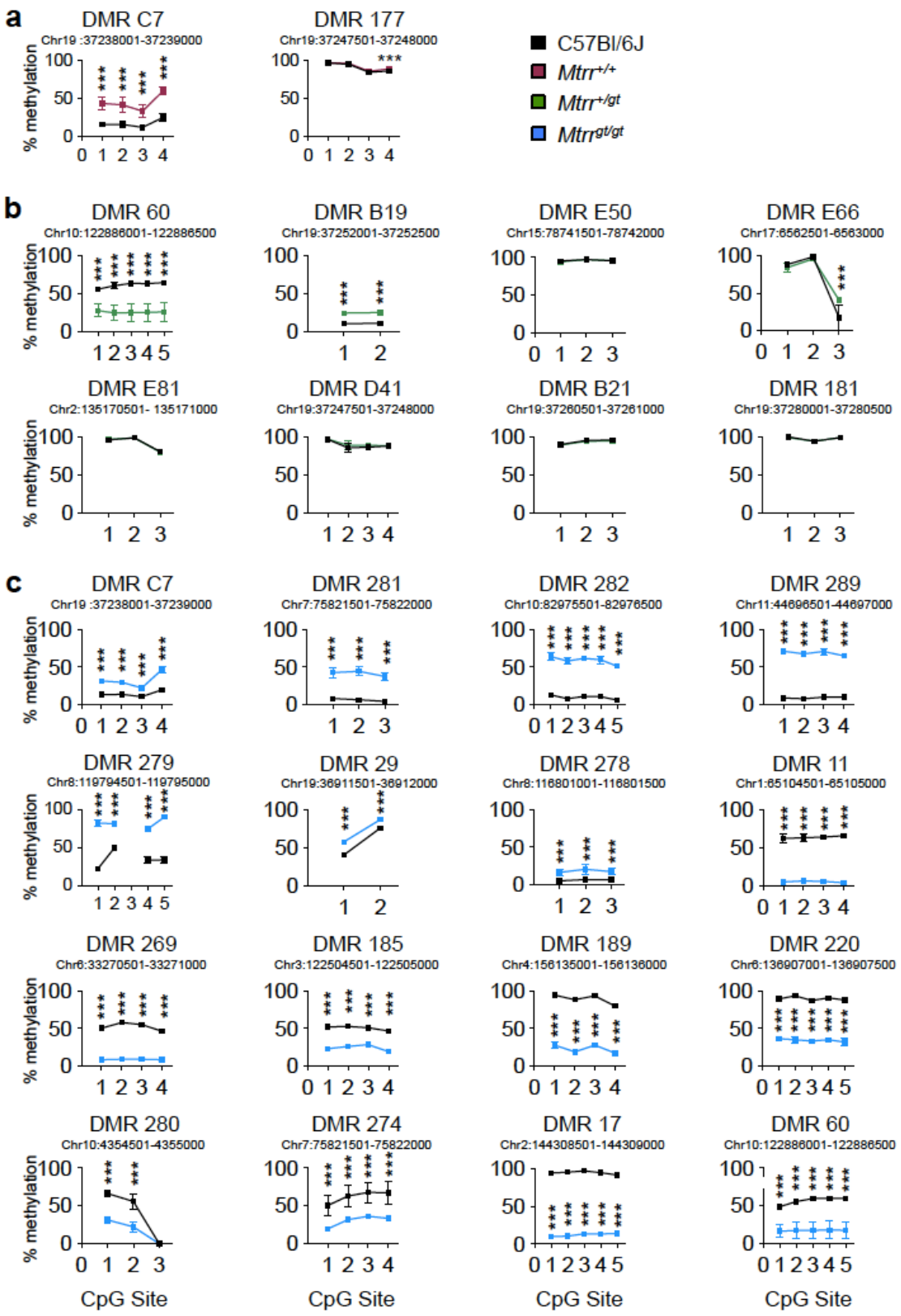
Highly controlled genetic pedigrees demonstrated that the *Mtrr*^{gt} mouse line is a model of transgenerational epigenetic inheritance[3], many of which are used in this study. Grey shaded boxes indicated pedigrees that do not display phenotypes at E10.5 in the final generation shown[3]. Pink shaded boxes indicate pedigrees that display a wide spectrum of phenotypes at E10.5 in the final generation shown[3]. **a** *Mtrr*$^{+/gt}$ maternal grandfather pedigree (top panel) and *Mtrr*$^{+/gt}$ maternal grandmother pedigree (bottom panel). **b** Control C57Bl/6J pedigree. **c** Embryo transfer of specified F2 wildtype blastocysts into control B6D2F1 pseudopregnant females demonstrated that some of the developmental phenotypes at E10.5 were independent of the F1 uterine environment[3]. **d** *Mtrr*$^{+/gt}$ intercross pedigree. **e** *Mtrr*$^{gt/gt}$ intercross pedigree. Pedigree legend: circle, female; square, male; blue outline, C57Bl/6J mouse line; black outline, *Mtrr*^{gt} mouse line; pink outline, B6D2F1 mouse line; white fill, *Mtrr*$^{+/+}$; half black-half white fill, *Mtrr*$^{+/gt}$; black fill, *Mtrr*$^{gt/gt}$.



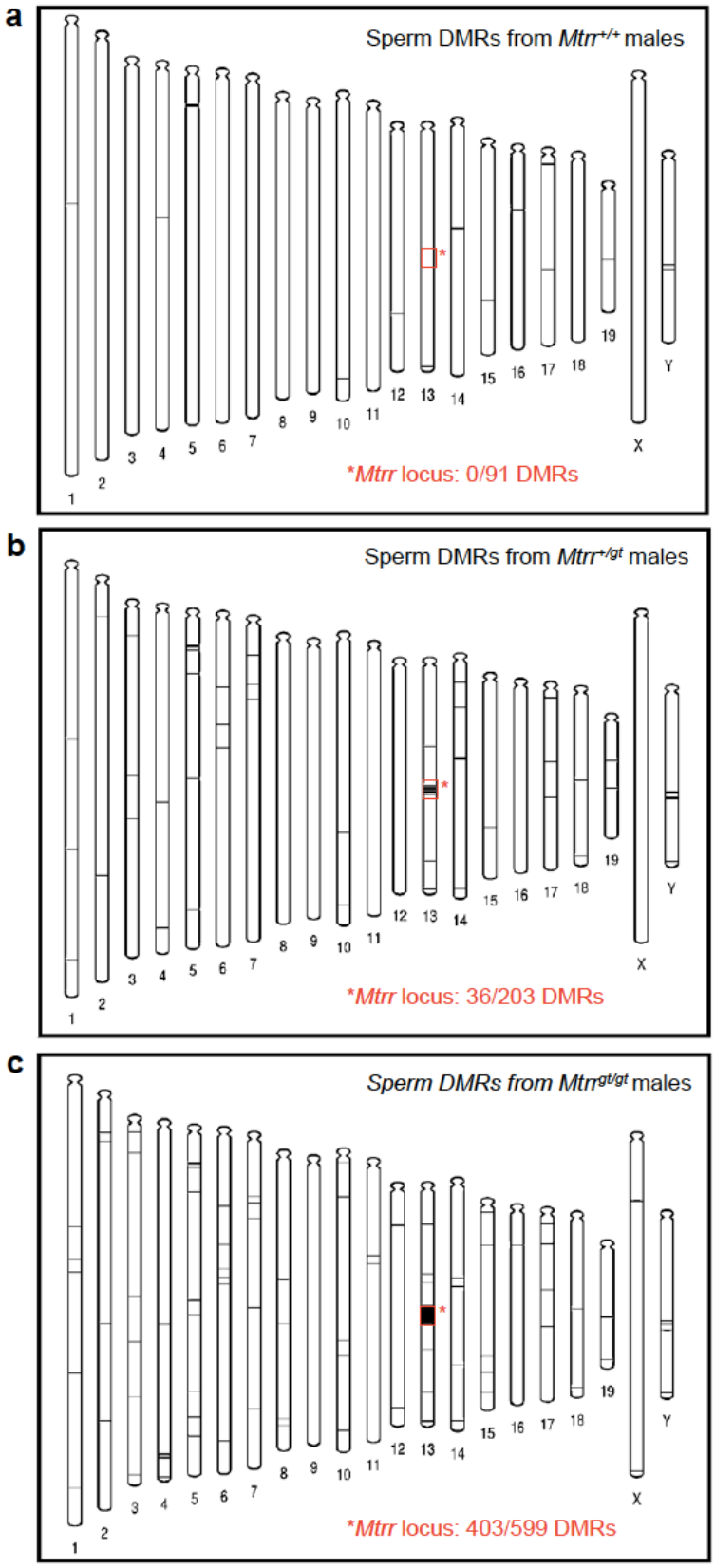
Supplementary Fig. 3. Frequency and location of genetic variants in C57Bl/6J and *Mtrr*$^{gt/gt}$ embryos.
a, b Whole genome sequencing data showing the average frequency of (**a**) structural variants (SVs) and (**b**) single nucleotide polymorphisms (SNPs) for each chromosome. Phenotypically normal C57Bl/6J embryos (N=2, black bars) and severely affected *Mtrr*$^{gt/gt}$ embryos (N=6, blue bars) were assessed. Data is presented as mean ± sd. Note that the gene-trap insertion in the *Mtrr* locus is on chromosome 13. **c, d** Genomic location of (**c**) SVs and (**d**) SNPs in C57Bl/6J and *Mtrr*$^{gt/gt}$ embryos after masking the region surrounding the gene-trap insertion site.



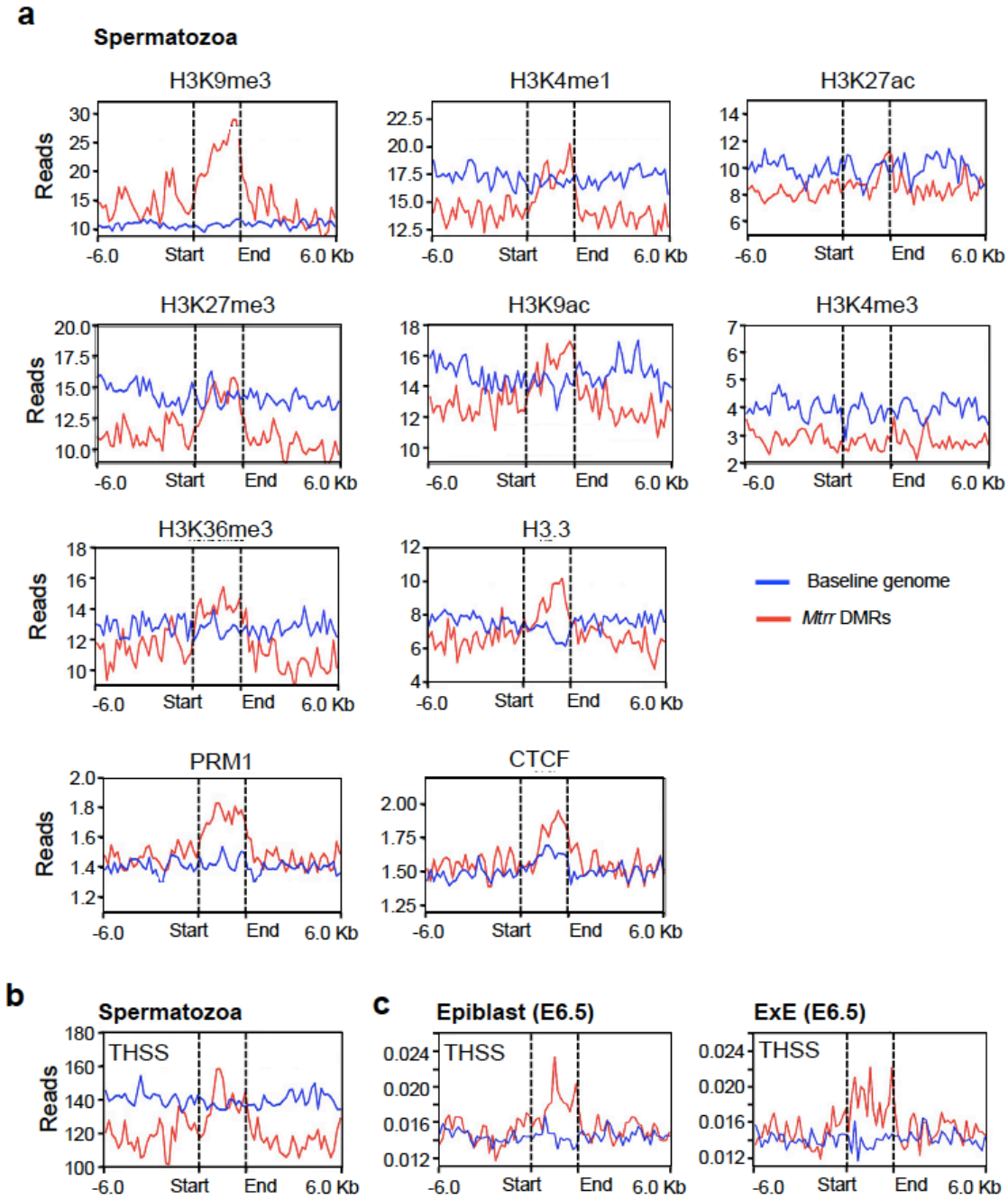
Supplementary Fig. 4. Confirmation of spermatozoa purity and validation of immunoprecipitation. a Bisulphite pyrosequencing of imprinting control regions in DNA from spermatozoa collected from cauda epididymides to determine sperm purity. Percentage methylation at specific CpG sites in the maternally imprinted *Peg3* DMR and paternally imprinted *H19* DMR were determined in sperm samples isolated from C57Bl/6J (black line), *Mtrr*$^{+/+}$ (purple line), *Mtrr*$^{+/gt}$ (green line) and *Mtrr*$^{gt/gt}$ (blue line) mice. C57Bl/6J liver (grey line) was assessed as a control. Data is represented as mean ± sd for each CpG site. N=8 samples per group. **b** Percentage recovery of DNA input after MeDIP experiment as determined using qPCR to amplify known methylated (*Nanog* and *H19*) and unmethylated (*H1t* and *TsH2B*) regions. MeDIP samples from sperm of C57Bl/6J (black), *Mtrr*$^{+/+}$ (purple), *Mtrr*$^{+/gt}$ (green), and *Mtrr*$^{gt/gt}$ (blue) males are shown. Each bar indicates one individual (N=8 males per genotype).



Supplementary Fig. 5. Validation of a panel of sperm DMRs identified in *Mtrr*$^{+/+}$, *Mtrr*$^{+/gt}$ and *Mtrr*$^{gt/gt}$ males. **a-c,** Validation of DMRs identified in a MeDIP-seq experiment of sperm from *Mtrr*$^{+/+}$ (purple line), *Mtrr*$^{+/gt}$ (green line) and *Mtrr*$^{gt/gt}$ (blue line) males relative to C57Bl/6J control sperm (black line). The average percentage methylation at individual CpG sites was determined by bisulfite pyrosequencing. Data is plotted as mean ± sd at each CpG site. N=8 males per group (four samples from MeDIP-seq analysis plus four unique samples). The coordinates for each DMR are given. Two-way ANOVA, with Sidak's multiple comparisons test, performed on mean methylation per CpG site per genotype group. ***p<0.001.

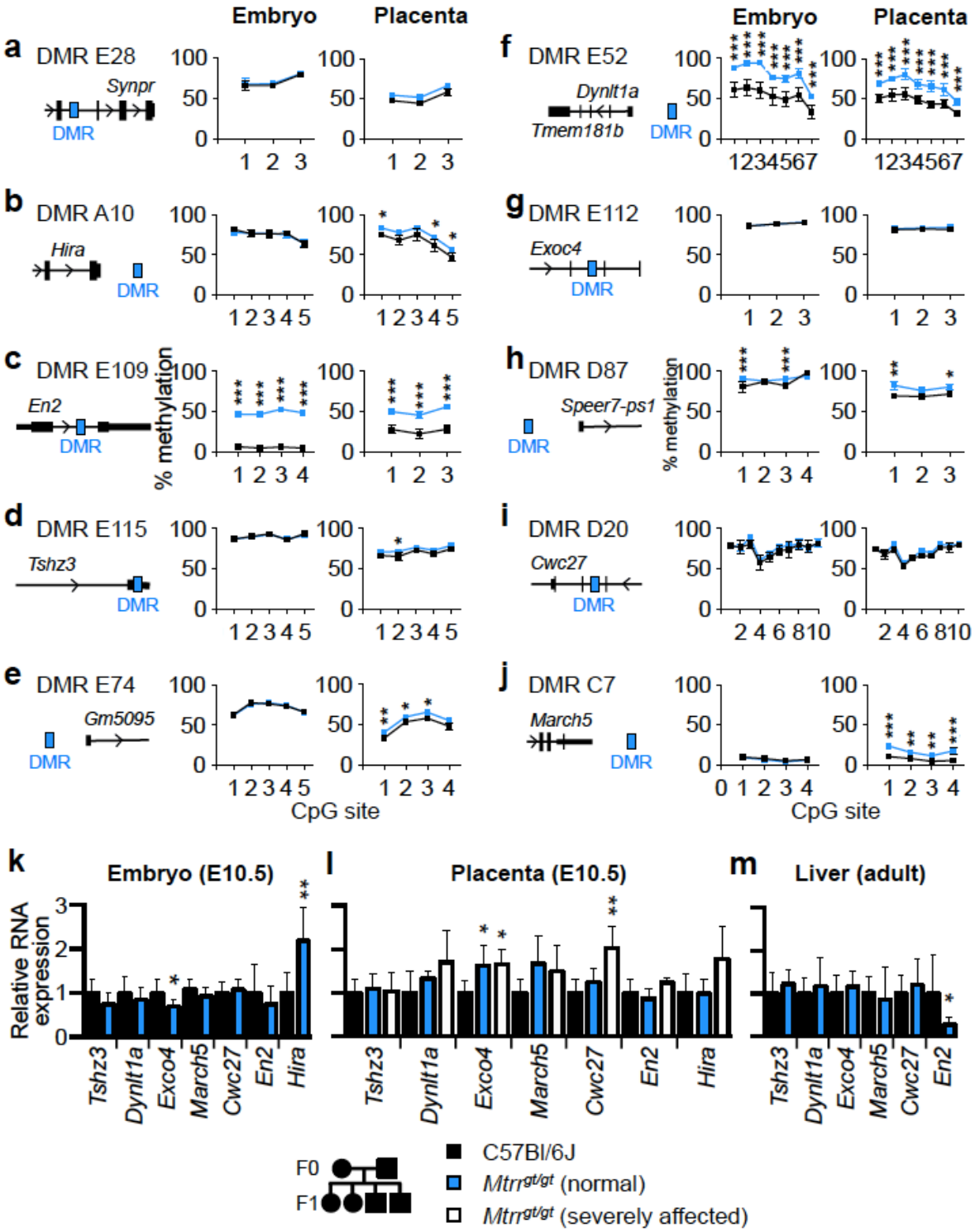


Supplementary Fig. 6. Chromosomal distribution of DMRs in sperm of *Mtrr* males. Phenograms showing chromosomal location of sperm DMRs (black lines) identified via MeDIP-seq analysis in (**a**) $Mtrr^{+/+}$, (**b**) $Mtrr^{+/gt}$, and (**c**) $Mtrr^{gt/gt}$ males. C57Bl/6J sperm were used as a control. Red box indicates region on Chromosome 13 surrounding *Mtrr* locus identified in **Fig. 1a**.



Supplementary Fig. 7. General epigenetic signature of genomic region identified as sperm DMRs in *Mtrr* males.

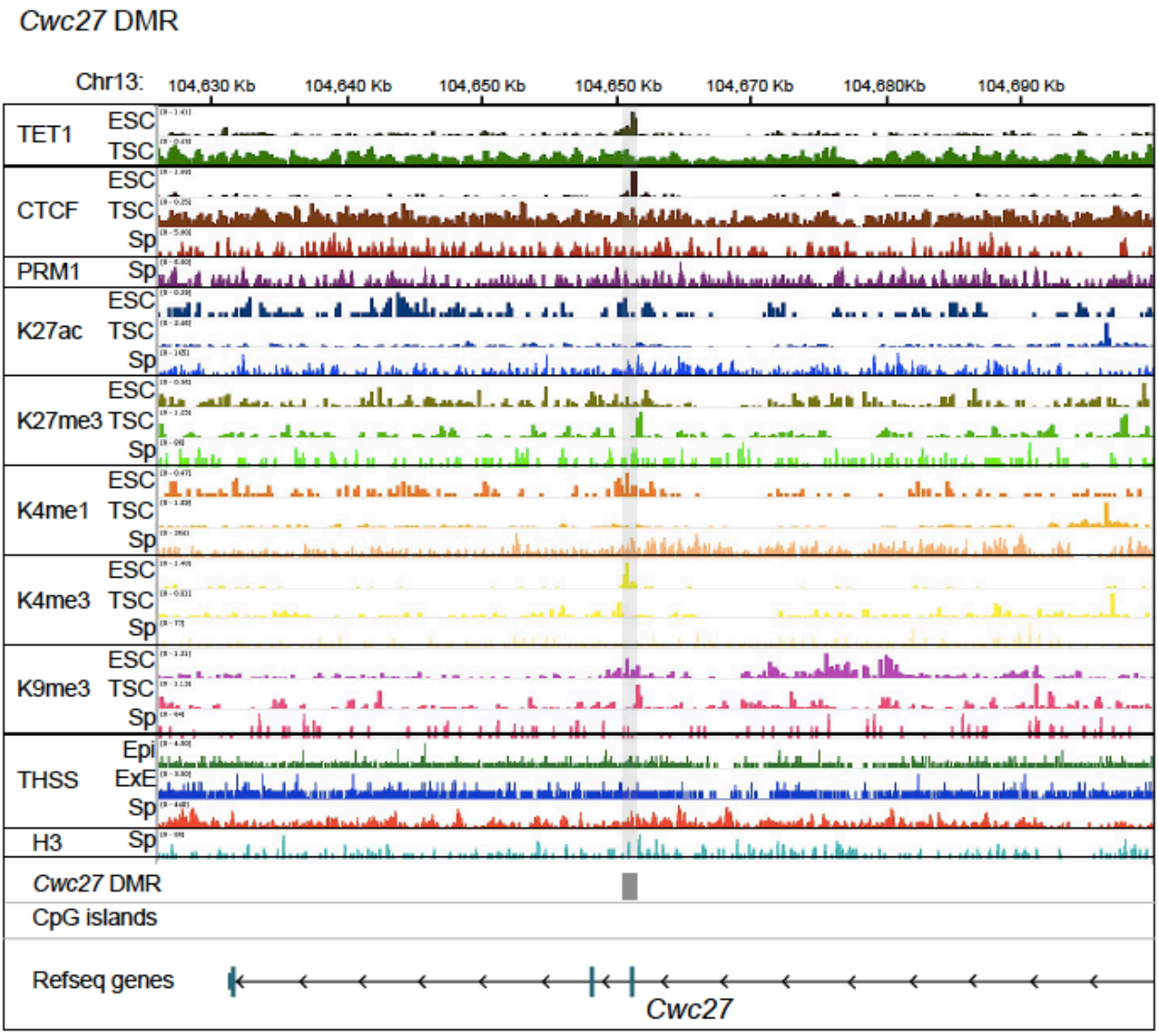
a, Using published ChIP-seq data sets in wildtype CD1 spermatozoa[37], mean enrichment of selected histone modifications and DNA binding proteins in the genomic regions identified as differentially methylated in sperm of all *Mtrr* genotypes combined (N= 379 DMRs; red line) was determined and compared to the baseline genome (blue line). **b-c**, Using published ATAC-seq data sets in (**b**) wildtype CD1 spermatozoa[37] and (**c**) wildtype B6D2F1 mouse epiblast and extraembryonic ectoderm (ExE) at embryonic day 6.5[38], mean enrichment of Tn5 transposase sensitive site (THSS) in the genomic regions identified as differentially methylated in sperm of all *Mtrr* genotypes combined (N= 379 DMRs; red line) compared to the baseline genome (blue line). DMRs in the region surrounding the *Mtrr* gene-trap insertion site were not included in either analysis. Dotted lines indicate the start and end of the DMR. Six kilobases (Kb) of DNA surrounding the DMR was also considered.



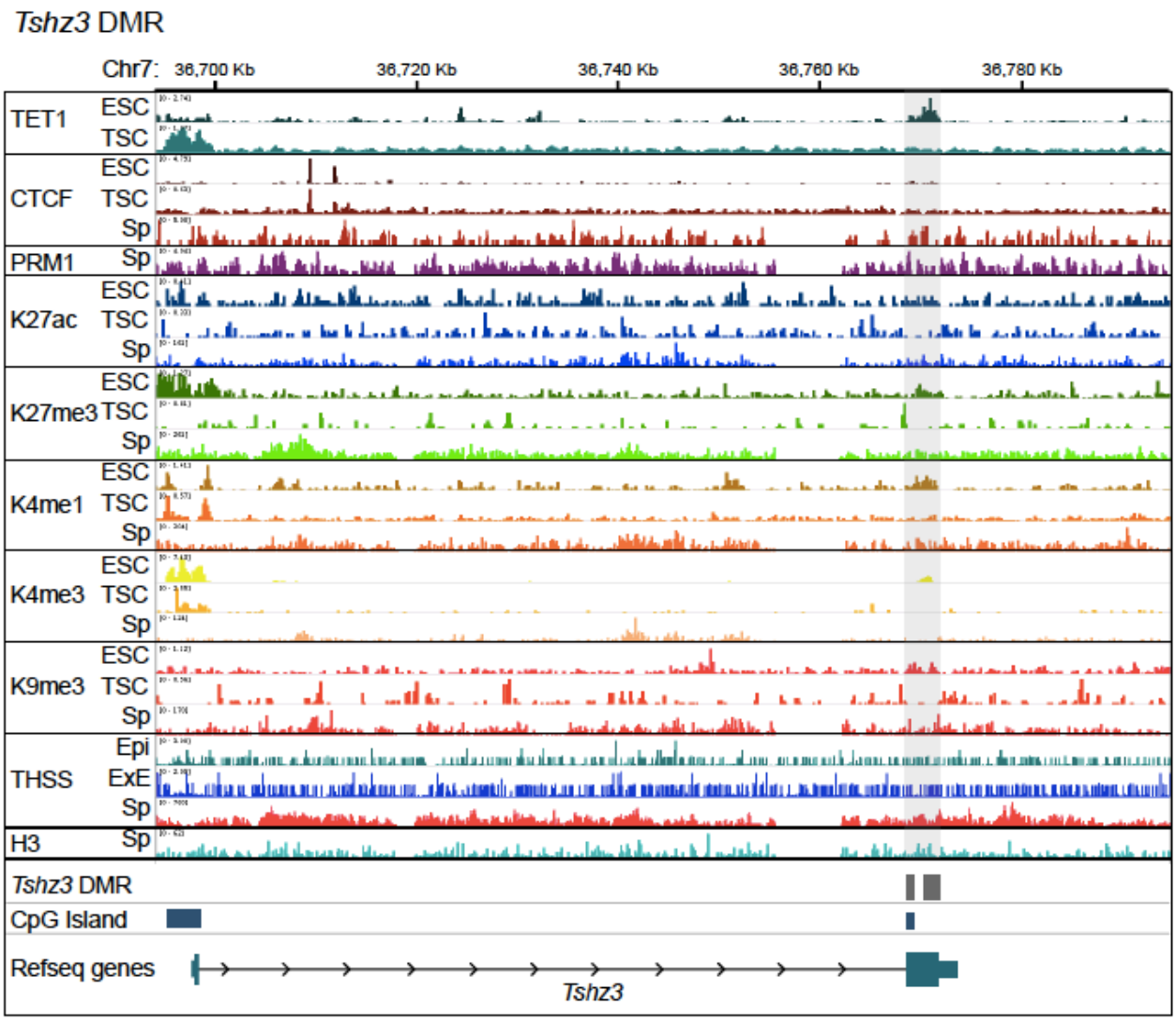
Supplementary Fig. 8. Analysis of DNA methylation and gene expression at sperm DMRs in *Mtrr*$^{gt/gt}$ tissue.

a-j, Schematic drawings of each DMR assessed indicate its relationship to the closest gene and are followed by graphs showing the average percentage of methylation at individual CpGs for the corresponding DMR as determined by bisulfite pyrosequencing. CpG methylation at several sperm DMRs from F0 *Mtrr*$^{+/gt}$ males was assessed in C57Bl/6J control (black lines) and *Mtrr*$^{gt/gt}$ (blue lines) embryos and placentas at E10.5. N=4 placentas and N=6-8 embryos assessed per genotypic group. Data is shown as mean ± sd for each CpG site. Two-way ANOVA, with Sidak's multiple comparisons test, performed on mean methylation per CpG site per genotype group. *p<0.05, **p<0.01, ***p<0.001. **k-m,** RT-qPCR analysis of mRNA expression of genes proximal to

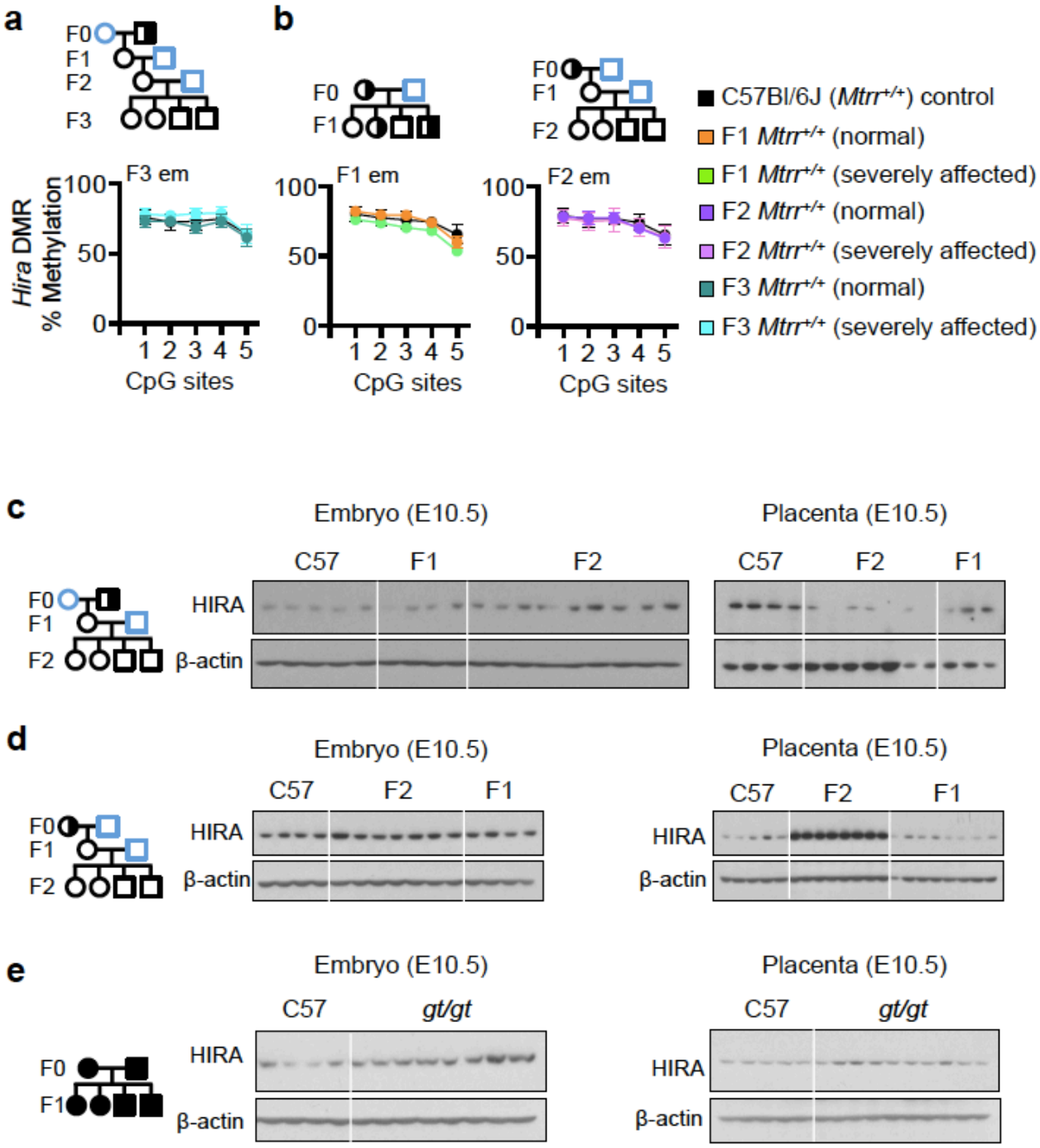
[Supplementary fig. 8 continued] sperm DMRs in *Mtrr*$^{gt/gt}$ (**k**) embryos and (**l**) placentas at E10.5, and (**c, f**) *Mtrr*$^{gt/gt}$ adult livers. C57Bl/6J control tissues (black bars) were assessed as controls. Tissue from (**k-m**) phenotypically normal *Mtrr*$^{gt/gt}$ individuals (blue bars) and (**l**) severely affected (white bars) was assessed. See also **Supplementary fig. 2b,e** for pedigrees used in this analysis. Data is plotted as mean ± sd, and is presented as relative expression to C57Bl/6J levels (normalized to 1). N=4-8 individuals per group. Independent t tests or one-way ANOVA with Dunnett's multiple comparisons tests were performed. *p<0.05, **p<0.01, ***p<0.001.



Supplementary Fig. 9. Epigenetic signature of the intragenic *Cwc27* DMR in wildtype ESCs, TSCs and spermatozoa.
Enrichment of DNA binding proteins (TET1, CTCF, PRM1, H3) and histone modifications (H3K27ac, H3K27me3, H3K4me1, H3K4me3, H3K9me3) in the *Cwc27* locus on Chr13 (~37,000 kb downstream of *Mtrr* gene) using published ChIP-seq data sets in wildtype embryonic stem cells (ESCs), trophoblast stem cells (TSCs)[42], and spermatozoa (Sp)[37]. Tn5 transposase sensitive sites (THSS) were also determined using published ATAC-seq data sets of normal B6D2F1 mouse epiblast and extraembryonic ectoderm (ExE) at embryonic day 6.5[38] or normal CD1 spermatozoa[37]. Grey box and shading indicate region of *Cwc27* DMR identified in sperm of $Mtrr^{+/gt}$ males. Schematic of protein encoding *Cwc27* partial transcript is shown at the bottom.



Supplementary Fig. 10. Epigenetic signature of the intragenic *Tshz3* DMR in wildtype ESCs, TSCs and spermatozoa.
Enrichment of DNA binding proteins (TET1, CTCF, PRM1, H3) and histone modifications (H3K27ac, H3K27me3, H3K4me1, H3K4me3, H3K9me3) in the *Tshz3* locus on Chr7 using published ChIP-seq data sets in wildtype embryonic stem cells (ESCs), trophoblast stem cells (TSCs)[42], and CD1 spermatozoa (Sp)[37]. Tn5 transposase sensitive sites (THSS) were also determined using published ATAC-seq data sets of wildtype B6D2F1 mouse epiblast and extraembryonic ectoderm (ExE) at embryonic day 6.5[38] or CD1 spermatozoa[37]. Grey box and shading indicate region of *Tshz3* DMR identified in sperm of $Mtrr^{+/gt}$ males. Schematic of protein encoding *Tshz3* transcript is shown.



Supplementary Fig. 11. Analysis of *Hira* DMR methylation and HIRA protein expression. **a-b**, Bisulfite pyrosequencing data showing average percentage methylation at individual CpGs in the *Hira* DMR (DMR A10) in (**a**) F3 wildtype embryos (em) at E10.5 derived from F0 $Mtrr^{+/gt}$ males, (**b**) F1 and F2 wildtype embryos at E10.5 derived from F0 $Mtrr^{+/gt}$ females. C57Bl/6J embryos were used as controls. Phenotypically-normal and severely affected conceptuses were assessed. Data is represented as average methylation (± sd) per CpG site. N=3-8 individuals per group. Two-way ANOVA, with Sidak's multiple comparisons test, performed on mean methylation per CpG site. **c-e,** Western blots showing HIRA protein in (**c-d**) wildtype F1 and/or F2 embryos and placentas at E10.5 derived from (**c**) F0 $Mtrr^{+/gt}$ males or (**d**) F0 $Mtrr^{+/gt}$ females, and (**e**) Mtrrgt/gt embryos and placentas from $Mtrr^{gt/gt}$ intercrosses. β-actin was used as a loading control. N=3-10 conceptuses were assessed per pedigree/genotype. Pedigree legend: circle, female; square, male; blue outline, C57Bl/6J control; black outline, $Mtrr^{gt}$ mouse line; white filled, $Mtrr^{+/+}$; half-white/half-black filled, $Mtrr^{+/gt}$; black filled, $Mtrr^{gt/gt}$. See also **Fig. 6b,g-h**.

Supplementary file 1. Primer sequences

qPCR primers				
Gene name	**Forward primer (5' → 3')**	**Reverse primer (5' → 3')**	**Primer conc. (nm)**	**Ref.**
Cwc27	TGATAATGGCAGCCAGTTTTTCT	CTGTCAGGCGTAGCATGTTGT	200	-
Cts8	TCCTGTGAAGAATCAGGGCAC	GTGCTCAGTGGGACCAGTTT	200	-
Dynlt1a	GAAGACTTCCAGGCCTCCG	GGTTGACTTTGCTGTGCTGG	100	-
En2	GCTGAGTTTCAGACCAACAGGTA	GCTTGTTCTGGAACCAAATCTT	200	-
Exoc4	CACAGCCTACAGGGGCATTG	TTGGCAGCGATTTCAAGAGTC	200	-
Gapdh	CATGGCCTTCCGTGTTCCT	GCGGCACGTCAGATCCA	variable	-
Gas1	CCTCTGCACCACGTGTCTTA	TGGCAGTACCGAGCTTTAGG	200	-
Hira (set 1)	CTCCATCTTGTCAGGAAGTGAT	GTTCCTGGCACTCAGTAAAGAG	200	-
Hira (set 2)	TATGAAAACGCCCCGCTCTT	ATTCGCGCTATTGGGCATCT	200	-
Hprt	CAGGCCAGACTTTGTTGGAT	TTGCGCTCATCTTAGGCTTT	variable	85
Hsd17b3	CTGAGCACTTCCGGTGAGAG	ATAAGGGGTCAGCACCTGAAT	200	-
IAP-GAG	AACCAATGCTAATTTCACCTTGGT	GCCAATCAGCAGGCGTTAGT	200	26
IAP-3'LTR	GCACATGCGCAGATTATTTGTT	CCACATTCGCCGTTACAAGAT	100	26
LINE1-5'UTR	GGCGAAAGGCAAACGTAAGA	GGAGTGCTGCGTTCTGATGA	100	26
LINE1-ORF2	GGAGGGACATTTCATTCTCATCA	GCTGCTCTTGTATTTGGAGCATAGA	200	26
March5	TTCACCAGGCTTGTCTCCA	GCATCACTGTCACTGCTCCA	150	-
Nsun2	GGTAAACCATGACGCATCC	CCTTCCACTGTGAGACTCC	100	86
Ptch1	CCTCTGCTCCTTGATTGGCA	TCCCCAGTCCTGTCCTCAAA	200	-
SINEB1	TGAGTTCGAGGCCAGCCTGGTCTA	ACAGGGTTTCTCTGTGTAGCCCTG	100	26
Srd5a1	CTTGAGCCAGTTTGCGGTGTA	GCCTCCCCTGGGTATTTTGTATC	200	-
Tbpba	ACTGGAGTGCCCAGCACAGC	GCAGTTCAGCATCCAACTGCG	200	-
Tshz3	GCGCGCAGCAGCCTATGTTTC	TCAGCCATCCGGTCACTCGTC	300	-
Uqcrb	TCTCAGGTCAAAATGGCGGG	ATCATCTCGCATTAACCCCAGT	200	-

iPCR Tag primers			
ID	**Primer sequence**	**Barcode**	**Sequence obtained**
iPCRtag1	CAAGCAGAAGACGGCATACGAGATAACGTGATGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	AACGTGAT	ATCACGTT
iPCRtag2	CAAGCAGAAGACGGCATACGAGATAAACATCGGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	AAACATCG	CGATGTTT

iPCR Tag primers (continued)			
ID	**Primer sequence**	**Barcode**	**Sequence obtained**
iPCRtag3	CAAGCAGAAGACGGCATACGAGATATGCCTAAGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	ATGCCTAA	TTAGGCAT
iPCRtag4	CAAGCAGAAGACGGCATACGAGATAGTGGTCAGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	AGTGGTCA	TGACCACT
iPCRtag5	CAAGCAGAAGACGGCATACGAGATACCACTGTGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	ACCACTGT	ACAGTGGT
iPCRtag6	CAAGCAGAAGACGGCATACGAGATACATTGGCGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	ACATTGGC	GCCAATGT
iPCRtag7	CAAGCAGAAGACGGCATACGAGATCAGATCTGGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	CAGATCTG	CAGATCTG
iPCRtag8	CAAGCAGAAGACGGCATACGAGATCATCAAGTGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC	CATCAAGT	ACTTGATG

Bisulfite pyrosequencing primers

DMR name	Coordinates	Forward primer (5' → 3')	Reverse primer (5' → 3')	Sequencing primer
11	Chr1: 65,104,501-65,105,000	[biotin]-ATAGATAGTGAAAGGATGGAGAGTAAGAA	CTCCAACCACTAAAAACTTAACTCA	CCACTAAAAACTTAACTCAATC
17	Chr2: 144,308,501-144,309,000	AGTTTTTTTGGGTGTGAGAATTAT	[biotin]-ATCCAAAACCACTAAAAAACACAAC	TGGGTGTGAGAATTATT
29	Chr19: 36,911,501-36,912,000	AGGGGAGAATTTTGAATGTGAT	[biotin]-ACCCCACAACTTTCTTATAACTT	GGTTATAGGGGGGAT
60	Chr10: 122,886,001-122,886,500	[biotin]-AGATGTAAAAGAAAGGAAGGTAGT	CAATCCCCCATTCAATACAAAAA	ACAAAAATACCCTCCC
177	Chr19: 37,247,501-37,248,000	[biotin]-AATTTAAGTGAAGAATTTTGGTTTTATG	AACCCTAAATATTCTCCTTTACTCAAC	ATTCTCCTTTACTCAACT
181	Chr19: 37,280,001-37,280,500	AGGAAGTATTTGAGATGTTTAGAGTT	[biotin]-CCACTACAATATATCCTTAACCCTACT	TTATAGGATTATAGGATGTGT
185	Chr3: 122,504,501-122,505,000	[biotin]-TGTTTTATTTGTAGTTGGAGAAGTAAG	ACCTCTCAAAAATCCTATCCATAATAT	CACCCCCTCACTAATCCTCTA
189	Chr4: 156,135,001-156,136,000	TGGGAGTTAGTTATTGGTTTAGTTTGAGG	[biotin]-AACCCAATACTAATCCACCTTTACA	TGAGGTTGGTATAGGAA
220	Chr6: 136,907,001-136,907,500	AGTTATGTGTAGGGTATTAATTTAATGT	[biotin]-TCTCCCCATACCTAACTTCACACT	AGATGTATGTTTTTGAAGTTTA
269	Chr6: 33,270,501-33,271,000	[biotin]-TAGGAAATAAAAATGTGAGGGGTAAT	ACTACTACTCTATCCCTTTTATAACA	ACTCTATCCCTTTTATAACAA
274	Chr7: 75,821,501-75,822,000	[biotin]-GGATGTAGATGAATTTGGAAATTTTAGA	ACTAAAACTCCACTATTCCTTTCCACAATT	CTATTCCTTTCCACAATTAC
278	Chr8: 116,801,001-116,801,500	GTTTGAATTTTTTGAGGGTTATTTAAGGTT	[biotin]-TCCCCCTTTTAAAACTATCTCTTCT	TTTGAGTTTAGTTTTTTATATTGTG

Bisulfite pyrosequencing primers (continued)				
DMR name	**Coordinates**	**Forward primer (5' → 3')**	**Reverse primer (5' → 3')**	**Sequencing primer**
279	Chr8: 119,794,501-119,795,000	GTTTTTTTAGTAGAGTTGGGAGTTT	[biotin]-AAACAAACTAAACCTAAACAAATATAACT	GGGAGTTTTGTTTTTTTTAGAT
280	Chr10: 4,354,501-4,355,000	AGAAGGGGTTATAGGAAGTATTTTAGG	[biotin]-CCCAAAAAATAATCCCTTCCCTTTTC	GGGTTATAGGAAGTATTTTAGGA
281	Chr7: 75,821,501-75,822,000	AGGATGGGGGAATAAAATGATG	[biotin]-ACCCAAACCTATATAATAATAACTTTCC	GGGGAATAAAATGATGTG
282	Chr10: 82,975,501-82,976,500	GAATTGTTTAGGGAGGATTTTTTTATAGT	[biotin]-CCAACATTCTACAAAATCTACTACTCC	TGTTGTGTAGATAGATATTTTGAT
289	Chr11: 44,696,501-44,697,000	GGTTGTGGGTTGTTATTGTAGTTAAG	[biotin]-ATCAAATTTTCTCTTTCCCCATTC	GTTATTGTAGTTAAGGGGT
A10	Chr16: 18,975,501-18,977,000	TGTTATAGGTTGGAGGGTATGAGTT	[biotin]-ATCCAATCCTTCTCAACAATCTCTCTC	AGGAGTAGAGAGTTTTAG
B19	Chr19: 37,252,001-37,252,500	[biotin]-AATTGGAGAATAGGATTTTTAGTGTTTATA	AACAATAATAAATACTTAAACTACA	TCCTAAATCTCCCTATAACACCATCAAT
B21	Chr19: 37,260,501-37,261,000	AGTATAAGTGATGGAGTTTTAGATAAAT	[biotin]-AATCTTAAACCAATACAAATCCATACT	TGATGGAGTTTTAGATAAATTAG
C7	Chr19: 37,238,001-37,239,000	GGTAGTTTTGGATATGTAAGAGTTG	[biotin]-CATATCCCCTCTCTCTTTATCTTTTTA	GGTTTATAGAGTTGAGAATTGTG
D20	Chr13: 104,660,501-104,661,000	GAAATGTTTGAGAGAATGGGTTTTAG	[biotin]-ATACAATAACCTTAACCACTCATAAC	AAATTTTAGTTTAAGGTTTTATATA
D41	Chr19: 37,247,501-37,248,000	[biotin]-AAGTGAAGAATTTTGGTTTTATGTAGAAT	AACCCTAAATATTCTCCTTTACTCAACT	CTCCTTTACTCAACTCT
D87	Chr5: 15,671,001-15,671,500	GGGTGTTGTTTTAGAGGTTTTGTATTTTTG	[biotin]-ATCACCTCATCCTACTAAATACTATAAC	AGATTAATTTTTTAAGTTTTGTTG
E28	Chr14: 13,512,001-13,512,500	TAGGGAGATTGATTGTGTTTTATTAGAT	[biotin]-ACTACCCTCTTCCTCTCTTACTCCTAAAT	AGATTGATTGTGTTTTATTAGATT
E50	Chr15: 78,741,501-78,742,000	AGGTAGATTAGGTTGAATATTTATAGAGA	[biotin]-AAAAACAAAATCTATAAAACCCTAAATCC	GGGTTGAAATTAAAGATAGG
E52	Chr17: 6,324,501-6,325,500	AGTTTTGGAGTTTAGGTTAGGT	[biotin]-ATAACCATATCAATAAAACCTTCTTAACA	AGTTTAGGTTAGGTTAAAG
E66	Chr17: 6,562,501-6,563,000	GGAGAGGGTTGTTTTTTTTGGGATAAGT	[biotin]-ATTTACCCTTCTCCCCACTACCA	GTGTTTTTGTTTTTAGTTAGTTATT
E74	Chr18: 47,505,501-47,506,000	GTAAGTTGGAGGTAGATATAGTTAGAAGT	[biotin]-ACTCCATTTACAACTCTTCACCATACT	ATAAAAGTTTTTAAATTTATTTGGG

Bisulfite pyrosequencing primers (continued)				
DMR name	**Coordinates**	**Forward primer (5' → 3')**	**Reverse primer (5' → 3')**	**Sequencing primer**
E81	Chr2: 135,170,501-135,171,000	TGTAAGGTTTAAGATAGGAGATGTTATT	[biotin]-TACAAACCCATATAACTACAAAACTCAT	TTTAGTTGGTGAGATTGA
E109	Chr5: 28,168,501-28,169,000	TAGAAGTTTTATTGGTTAGGTGTAAGTTA	[biotin]-CAAAAAAATCTTTAAACACCAAAAAACATC	ATGGATGTGAAGAGG
E112	Chr6: 33,270,501-33,271,000	[biotin]-TAGGAAATAAAAATGTGAGGGGTAAT	ACTACTACTCTATCCCTTTTATAACA	ACTCTATCCCTTTTATAACAAT
E114	Chr7: 16,633,501-16,634,000	GTTTTTGGGTTTTAGTAATAGTGTATGG	[biotin]-AACACATATAACTTCCCACACATC	GTTTTTAGAGAGTAGTTTAGG
E115	Chr7: 36,770,501-36,772,000	AGTTTGATATTGAAGGGGTTATTTTGGA	[biotin]-AACATAATATACTTCCCTTCTAAAATCT	GGTTATTTTGGAAGAGGT

Primer concentration: PCR primers, 250 nm; sequencing primers, 417 nm.

Supplementary file 2. Accession numbers of published ChIP-seq and ATAC-seq data sets used in this study.

ID	Experiment	Tissue	Target	Accession	Ref.	Data
1	ChIP-seq	ESC	H3K27ac	GEO: GSM1000099	42	raw/processed*
2	ChIP-seq	ESC	H3K4me3	GEO: GSM769008	42	raw/processed*
3	ChIP-seq	ESC	H3K9me3	GEO: GSM1000147	42	raw/processed*
4	ChIP-seq	ESC	H3K4me1	GEO: GSM769009	42	raw/processed*
5	ChIP-seq	ESC	H3K27me3	GEO: GSM1000089	42	raw/processed*
6	ChIP-seq	ESC	CTCF	GEO: GSM918748	42	raw/processed*
7	ChIP-seq	ESC	TET1	GEO: GSM611192	42,43	raw/processed*
8	ChIP-seq	TSC	H3K27ac	GEO: GSM1035380, GSM1035381	46	raw/processed*
9	ChIP-seq	TSC	H3K4me3	GEO: GSM1035382	46	raw/processed*
10	ChIP-seq	TSC	H3K9me3	GEO: GSM1035383, GSM1035384	46	raw/processed*
11	ChIP-seq	TSC	H3K4me1	GEO: GSM1035385	46	raw/processed*
12	ChIP-seq	TSC	H3K27me3	GEO: GSM1035386, GSM1035387	46	raw/processed*
13	ChIP-seq	TSC	CTCF	GEO: GSM967658	44	raw/processed*
14	ChIP-seq	TSC	TET1	GEO: GSE109545	45	raw/processed*
15	ChIP-seq	Sperm	H3K27ac	GEO: GSM2088387, GSM2401435	37	processed
16	ChIP-seq	Sperm	H3K4me3	GEO: GSM2088391, GSM2401439	37	processed
17	ChIP-seq	Sperm	H3K9me3	GEO: GSM2088388, GSM2401436	37	processed
18	ChIP-seq	Sperm	H3K4me1	GEO: GSM2088390, GSM2401438	37	processed
19	ChIP-seq	Sperm	H3K27me3	GEO: GSM2088386, GSM2401434	37	processed
20	ChIP-seq	Sperm	CTCF	GEO: GSM2088382, GSM2088383, GSM2088384	37	processed
21	ChIP-seq	Sperm	H3	GEO: GSM2088392	37	processed
22	ChIP-seq	Sperm	PRDM1	GEO: GSM2088400, GSM2401441	37	processed
23	ATAC-seq	Sperm	THSS	GEO: GSM2088376, GSM2088377, GSM2088378	37	processed
24	ATAC-seq	ExE	THSS	GEO: GSM2229962, GSM2229963	38	processed
25	ATAC-seq	Epiblast	THSS	GEO: GSM2229960, GSM2229961	38	processed

*The source given in the table is the link for the raw fastq files, which was reused by ref [42]. We used the processed file provided by the authors of ref [42].
ESC, embryonic stem cells; ExE, extraembryonic ectoderm; Sperm, spermatozoa; TSC, trophoblast stem cells