Dynamic CpG island methylation landscape in oocytes and its fate in preimplantation embryos

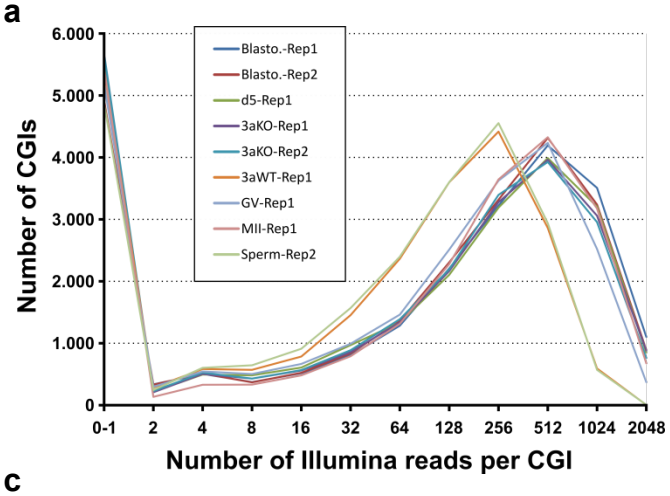
Sébastien A. Smallwood[1], Shin-ichi Tomizawa[1], Felix Krueger[2], Nico Ruf[1], Natasha Carli[1], Anne Segonds-Pichon[2], Shun Sato[3], Kenichiro Hata[3], Simon R. Andrews[2] and Gavin Kelsey[1,4,]

[1]Laboratory of Developmental Genetics and Imprinting, The Babraham Institute, Cambridge, CB22 3AT, UK
[2]Bioinformatics Group, The Babraham Institute, Cambridge, CB22 3AT, UK
[3]Department of Maternal-Fetal Biology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya, Tokyo 157-8535, Japan
[4]Centre for Trophoblast Research, University of Cambridge, Cambridge, CB2 3EG, UK

Supplementary Information.

Supplementary Figure 1: Statistics of RRBS Illumina sequencing. **a,** Illumina sequencing read coverage at CGIs for nine representative RRBS datasets, showing that the majority of CGIs with appropriate MspI fragments are deeply sequenced. **b,** Summary of RRBS outputs for all datasets presenting the number of individual CpGs sequenced with a read depth ≥5 (informative CpGs), the number of informative CpGs located within CGIs, and the number of informative CGIs (for which informative CpGs represent ≥10% of all CpGs per CGI, with a minimum of 5 CpGs). **c,** For nine representative RRBS datasets, the percentage coverage of CGIs by CpGs with a read depth ≥5 is shown, demonstrating that the majority of the informative CGIs are observed at well above the cut-off of 10% of CpGs per CGIs. **d**, Illumina sequencing read coverage at individual CpGs for nine representative RRBS datasets, showing that the majority of CpGs are deeply sequenced. **e,** Summary statistics of two RRBS libraries independently prepared from 15ng or 500pg of the same genomic DNA (GV oocytes). **f,** Correlation of CGI methylation calls between RRBS libraries generated with 15ng or 500pg of the same genomic DNA, demonstrating the reproducibility of our protocol. CGIs with contradictory methylation calls (for example, ≥75% in the 500pg library and ≤25% in the 20ng one) were removed from subsequent analysis. **g,** Correlation of CGI methylation calls between RRBS libraries generated from GV oocytes (GV-Rep1) and *3a+/+* GV oocytes (3aWT-Rep1).

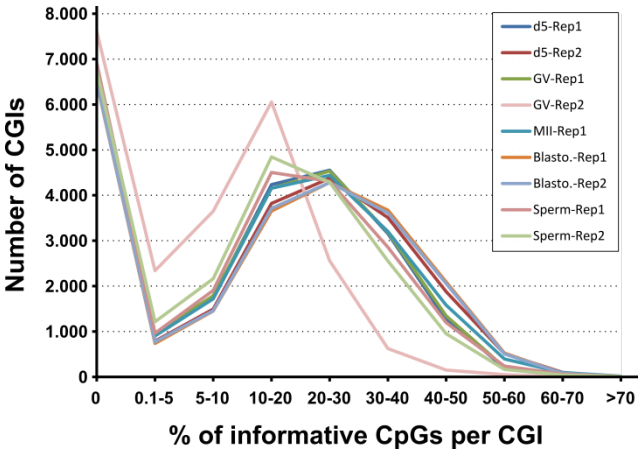
a



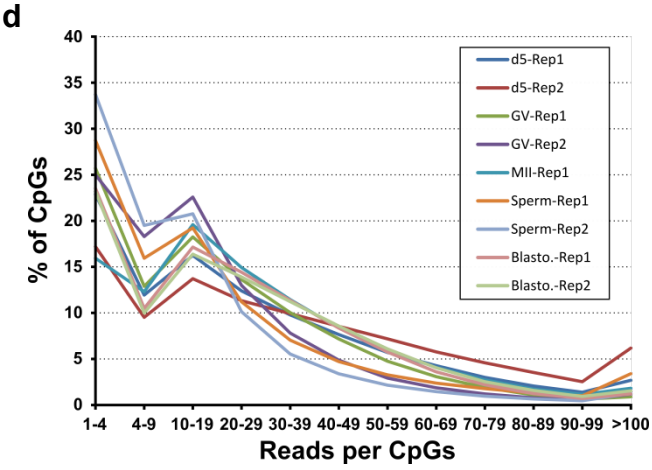
b

Datasets	Number of CpGs (>5reads)	Number of CpGs overlapping CGIs	Number of informative CGIs
d5-Rep1	766,689	647,433 (84.4%)	13,241
d5-Rep2	898,635	736,121 (81.9%)	13,953
GV-Rep1	804,744	650,327 (80.8%)	13,280
GV-Rep2	494,080	358,411 (72.5%)	9,276
MII-Rep1	870,700	685313 (78.7%)	13,660
Blasto.-Rep1	930,316	743,878 (80.0%)	14,088
Blasto.-Rep2	912,147	738,468 (81.0%)	14,023
Sperm-Rep1	905,126	612,123 (67.6%)	12,974
Sperm-Rep2	829,608	577,057 (69.6%)	12,644
3aWT-Rep1	744,982	574,343 (77.1%)	12,526
3aKO-Rep1	801,658	662,807 (82.7%)	13,022
3aKO-Rep2	796,506	659,294 (82.8%)	13,069
3LWT-Rep1	886,516	714,043 (80.5%)	13,650
3LWT-Rep2	1,240,360	886,562 (71.5%)	14,982
3LKO-Rep1	773,081	626,237 (81.0%)	12,784
3LKO-Rep2	1,004,009	722,496 (72.0%)	13,787

c



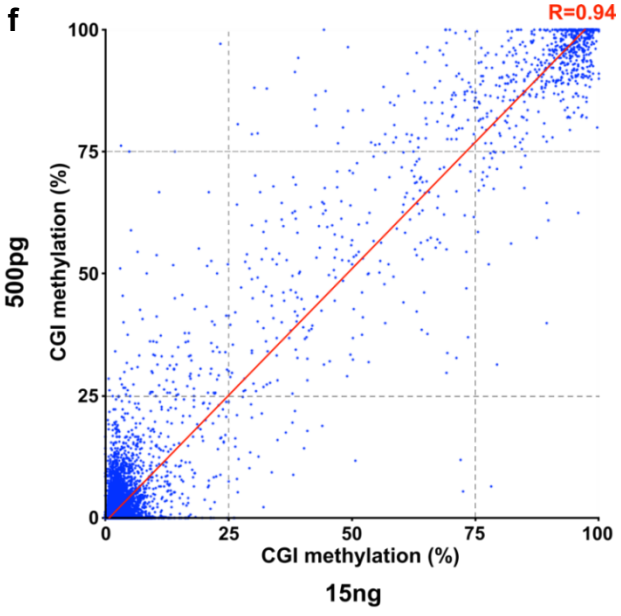
d



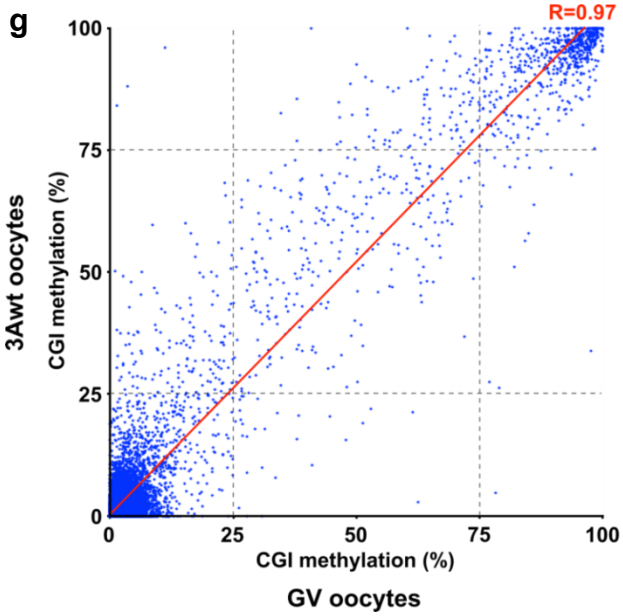
e

	15ng	500pg
number of sequences generated	15,836,627	14,085,820
number of MspI sequences	15,223,791 (96.1%)	12,649,330 (89.8%)
number of aligned sequences	10,583,896 (66.9%)	7,431,567 (52.8%)
number of informative CpGs	804,744	494,080
% of CpGs within CGIs	80.8%	72.5%
number of informative CGIs	13,280	9,276

f

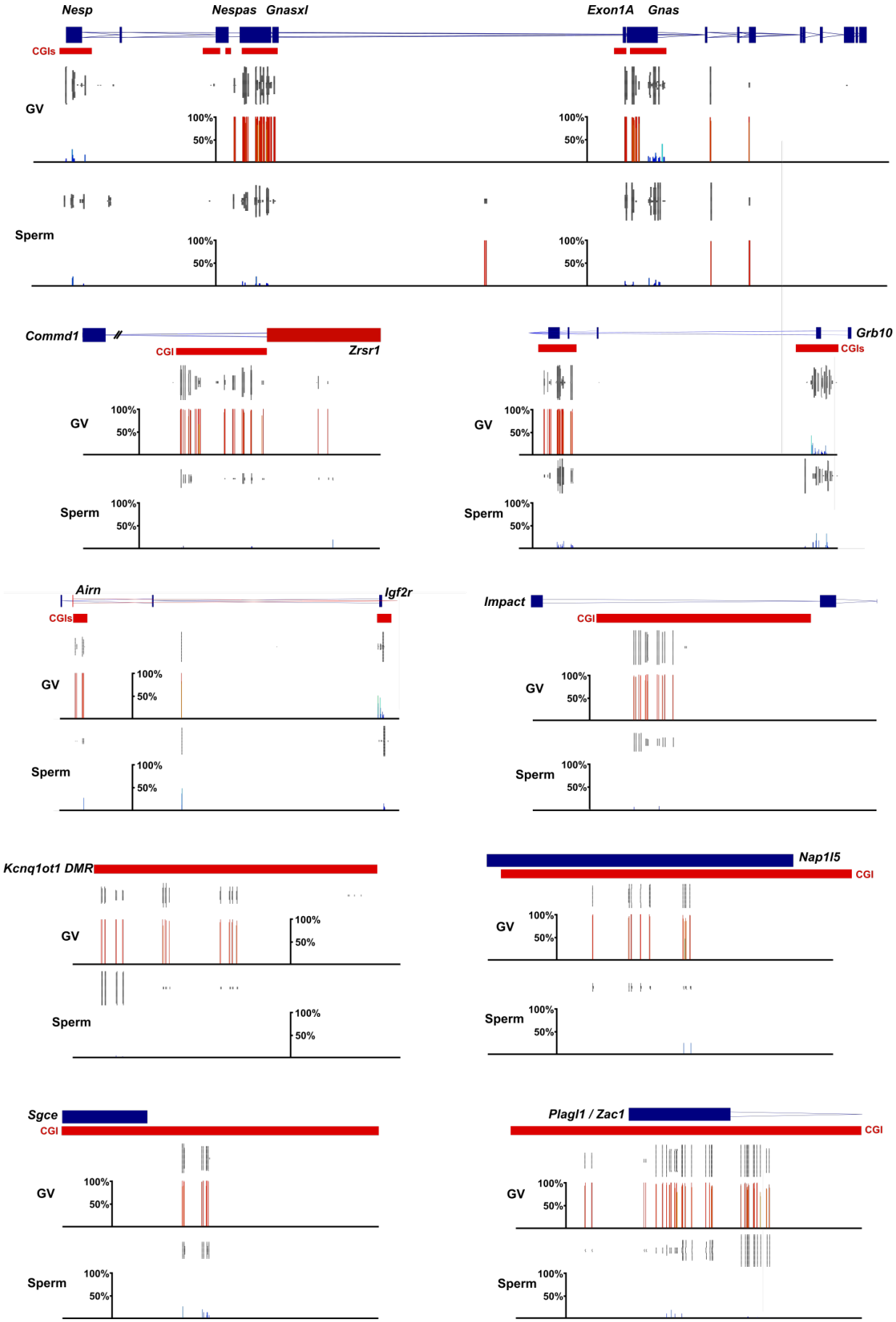


g

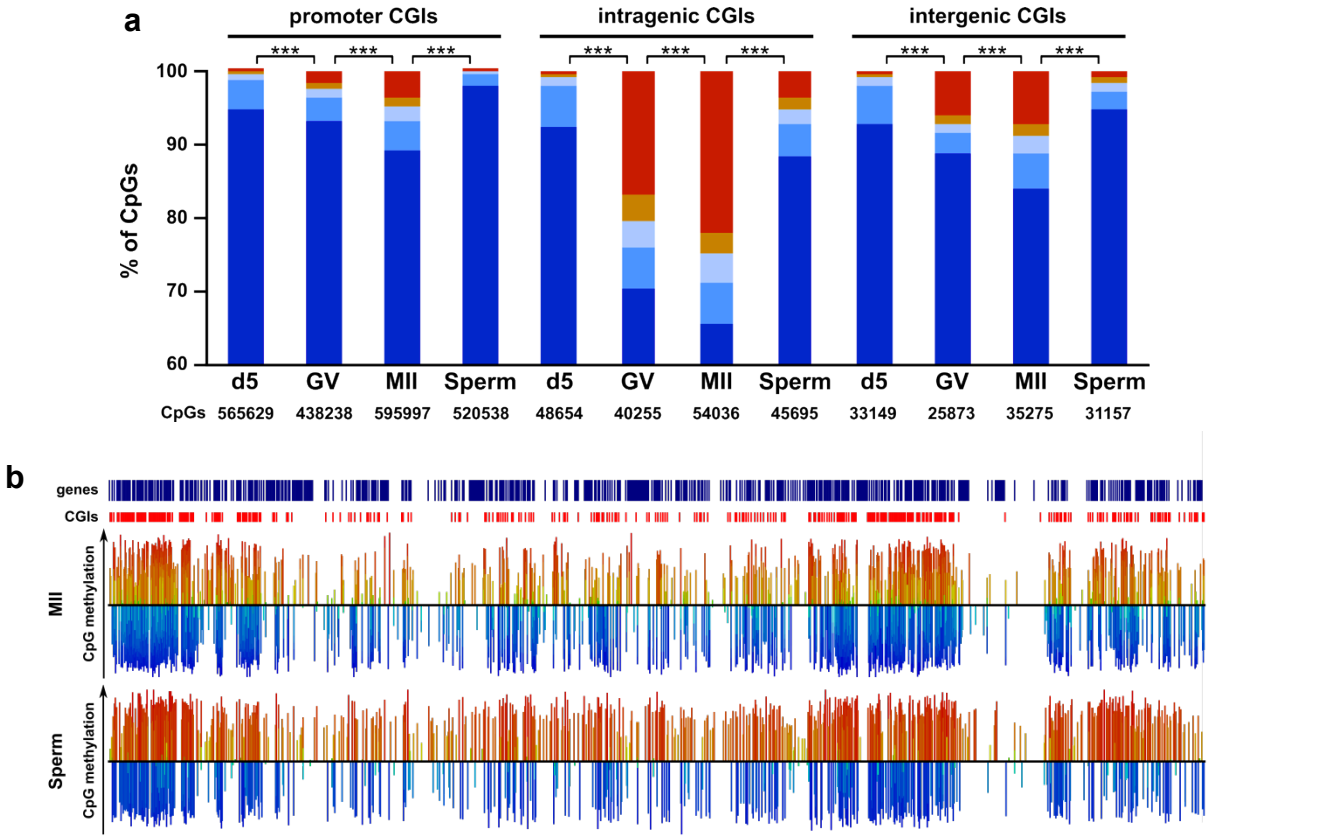


Supplementary Figure 1.

Supplementary Figure 2: CpG methylation identified by RRBS at germline DMRs of known imprinted loci in GV oocytes (GV-Rep1) and sperm (Sperm-Rep1). For each gene, exon organisation and location of CGIs/DMRs are given at the top. Below, the vertical grey lines represent the sequence read depth for each cytosine scored and the coloured lines the methylation level, with range from 0 to 100% from dark blue to dark red. For example, at the *Gnas* cluster, the germline DMRs at *Nespas/Gnasxl* and *Exon1A* exhibit high levels of methylation in oocytes but very low methylation in sperm, whereas the somatic DMR at *Nesp* and the constitutively unmethylated CGI at the *Gnas* promoter exhibit very low methylation in both oocytes and sperm. Similarly, the germline DMRs at *Airn* and *Grb10* show high methylation specifically in oocytes, whilst the CGIs for the *Igf2r* and *Grb10* promoters, respectively, in the same loci are unmethylated in both oocytes and sperm.



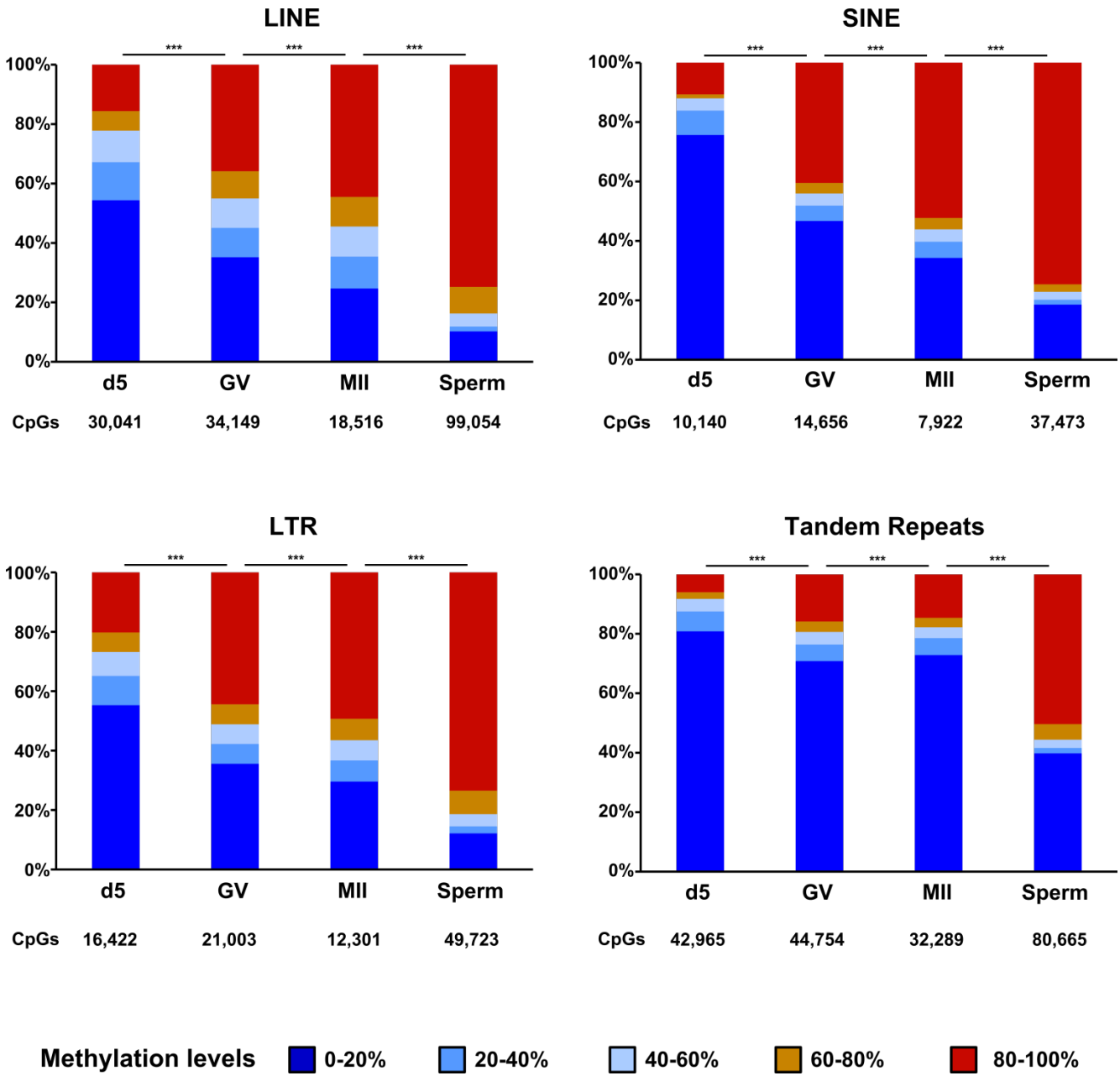
Supplementary Figure 2



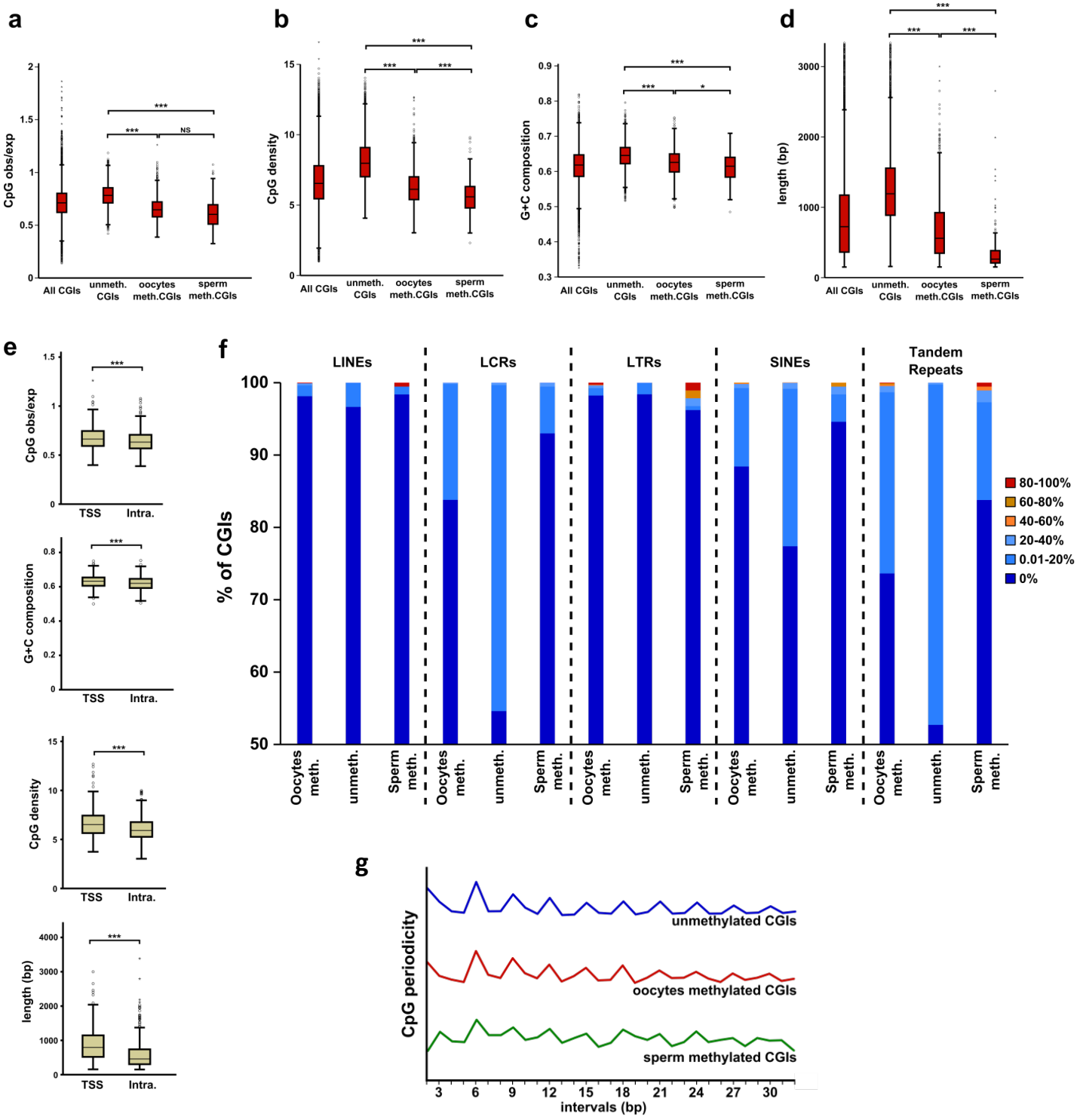
c

	all CGI	oocyte (all)	sperm (all)	oocyte specific	sperm specific	oocyte & sperm
number of CGIs	23,020	1,062	185	687	58	100
promoter (%)	58.7	38.6	14.1	47.3	5.2	14.0
intragenic (%)	23.9	52.3	73.5	42.9	74.1	80.0
intergenic (%)	17.5	9.1	12.4	9.8	20.7	6.0

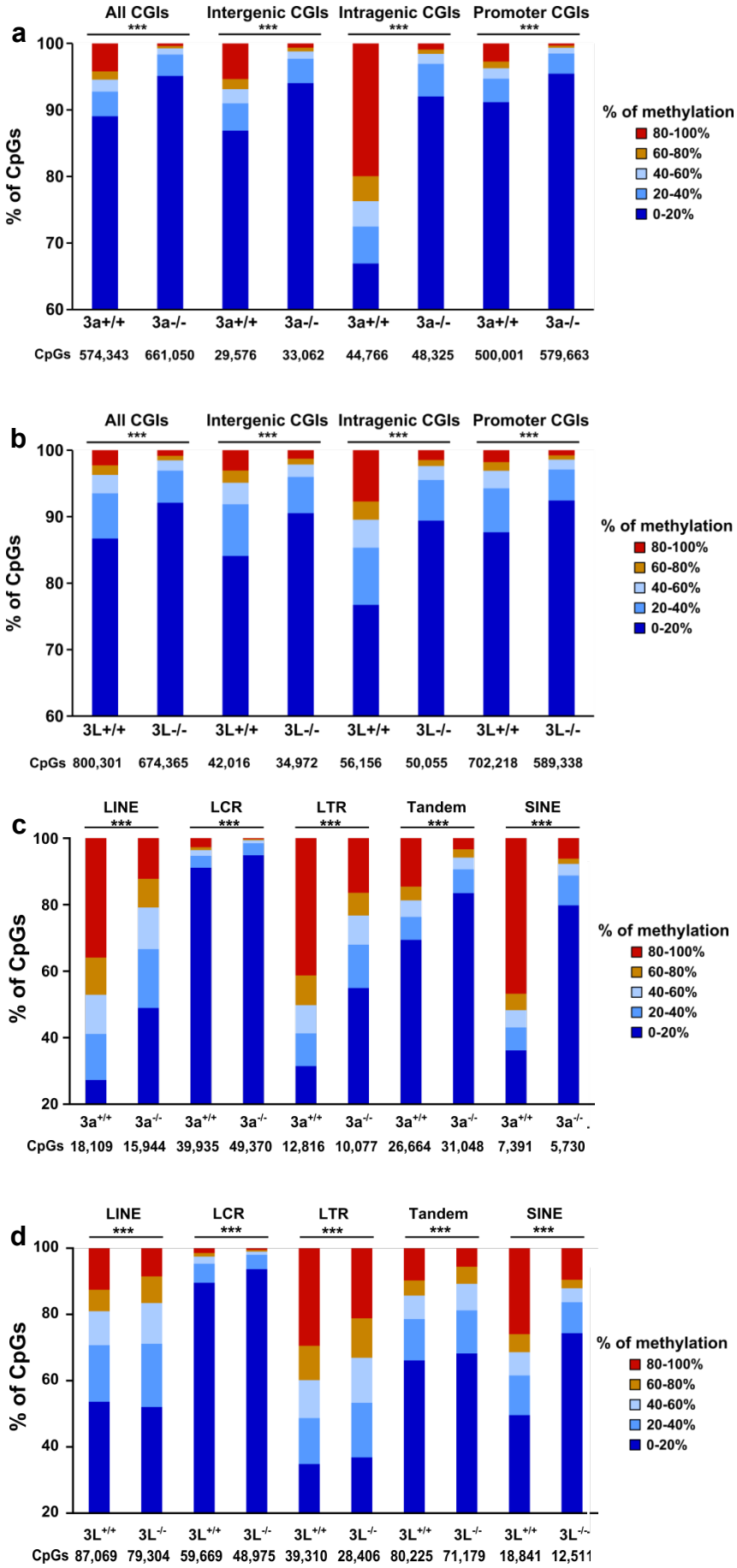
Supplementary Figure 3: **Distribution of CpG and CGI methylation. a,** Distribution of CpG methylation levels within CGIs classified according to their genomic location (promoter, intragenic, intergenic), in immature (d5), mature (GV and MII) oocytes and sperm. Below is indicated the number of CpGs, averaged between biological replicates, analysed in each category. This result indicates that CpGs located within CGIs are less methylated in sperm compared to mature oocytes, independent of CGI location, and that intragenic CGIs are proportionately more methylated than promoter and intergenic CGIs, in both oocytes and sperm (***: $p<0.001$, $\chi 2$ test) **b,** Entire chromosome 19 views of CpG methylation in MII oocytes and sperm. Data are presented as log methylated divided by unmethylated cytosines, such that green to red lines above the line approximate to degrees of methylation >50% and pale to dark blue lines below the line approximate to degrees of methylation <50%. This view reflects the higher CpG methylation level across the genome in sperm compared to MII oocytes **c,** Locations of methylated CGIs with respect to transcription units based on EnsEMBL, Refseq and UCSC annotations; values for promoter, intragenic and intergenic are percentage of total CGIs in each group (first row).



Supplementary Figure 4: CpG methylation in repetitive elements. Distribution of CpG methylation levels within repetitive elements (RE) in d5, GV, MII oocytes, and sperm, demonstrating the higher methylation of REs in sperm compared to mature oocytes. Below each graph the number of CpGs scored within the elements is given (***: $p<0.001$, $\chi 2$ test).

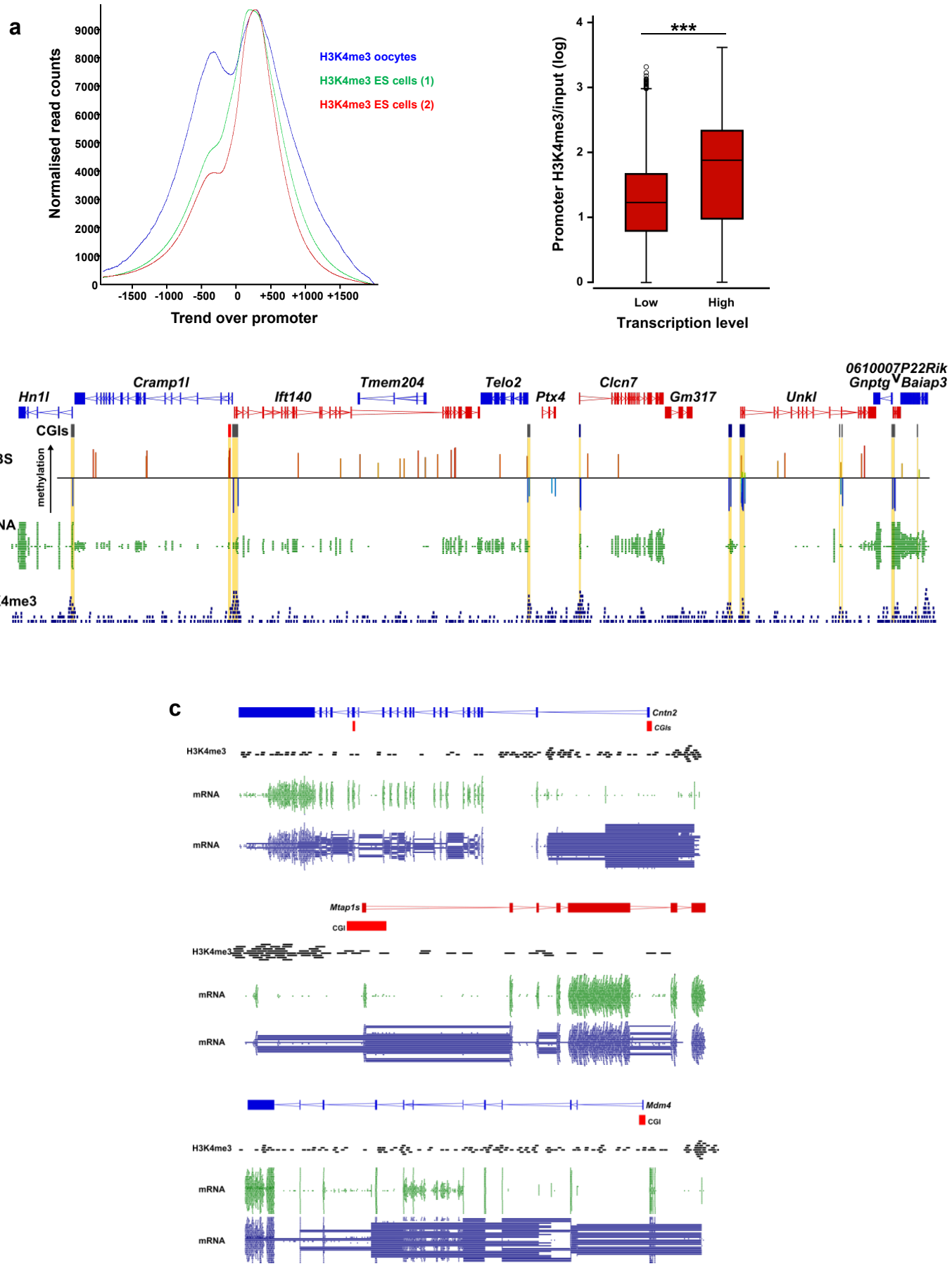


Supplementary Figure 5: Properties of CGIs methylated in oocytes and sperm. **a-d,** Length, CpG and GC contents and CpG density in CGIs methylated in oocytes (1062 CGIs), sperm (185 CGIs) or unmethylated in gametes (7318 CGIs). (NS: non-significant; **: p<0.01; ***: p<0.001, Kruskall-Wallis test) **e,** Comparison of properties of methylated CGIs in oocytes between promoter and intragenic CGIs (***: p<0.001, Mann-Whitney U test) **f,** Repetitive element (RE) contents of CGIs methylated in oocytes, sperm or unmethylated in the germline. The graphs present the % of the CGIs having RE contents of 0%, 0.01-20%, etc. **g,** CpG periodicity plots of CGIs methylated in oocytes, sperm or unmethylated in gametes.



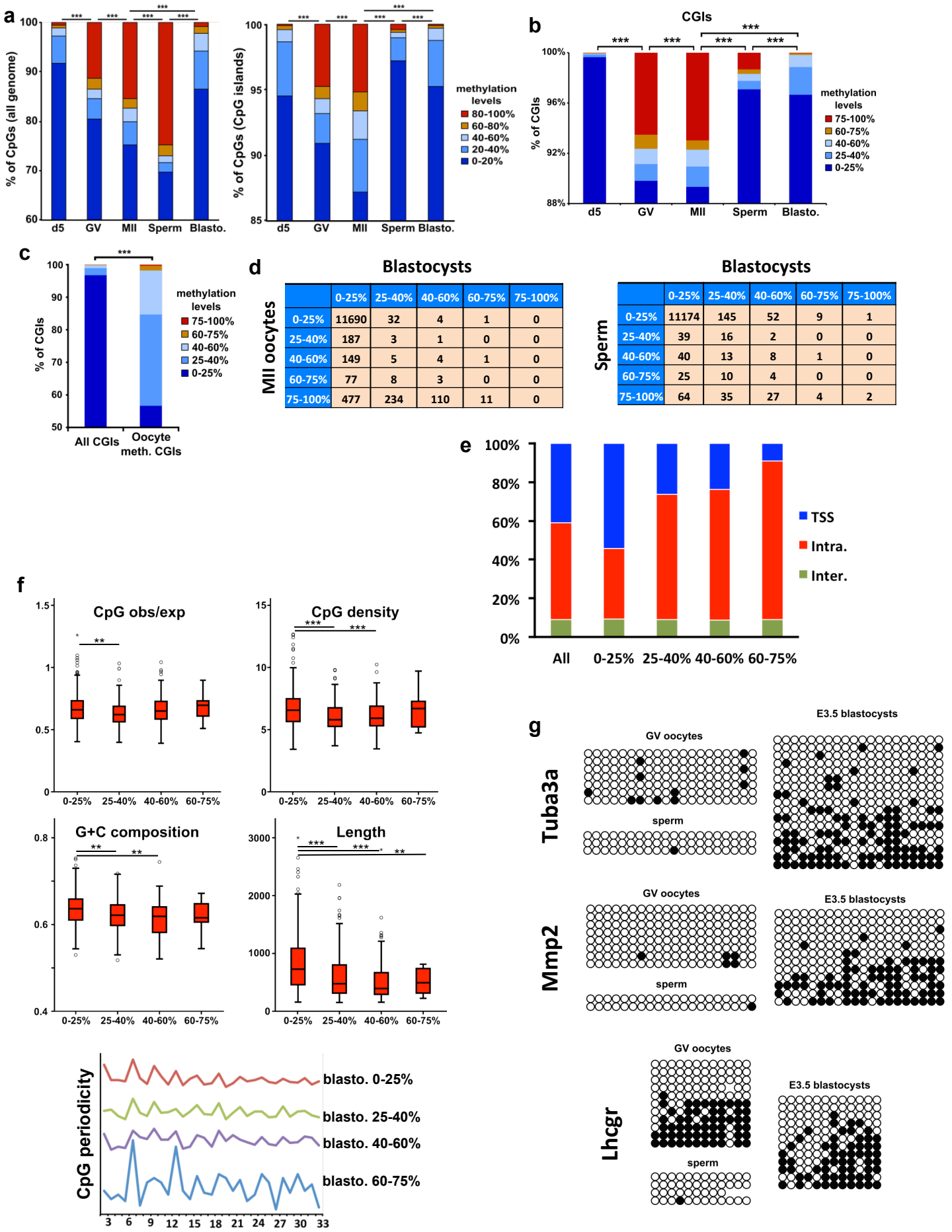
Supplementary Figure 6: Methylation in Dnmt3a and Dnmt3L deficient oocytes. Distribution of methylation levels of CpGs overlapping the different classes of CGIs (promoter, intragenic, intergenic) (**a,b**), or different classes of repetitive elements (RE) (**c,d**), in *Dnmt3a* wild-type (*3a+/+*) and knock-out (*3a-/-*), or *Dnmt3L* wild-type (*3L+/+*) and knock-out (*3L-/-*) GV oocytes. This analysis shows that both DNA methyltransferases have a global impact on CpG methylation, independently of their genomic context. The generally lower level of RE methylation in *Dnmt3L+/+* oocytes and the apparently 'weaker' effect of absence of DNMT3L might be attributed to the *Dnmt3L+/+* and *Dnmt3L-/-* oocytes being collected at a less advanced stage in comparison to *Dnmt3a+/+* and other GV oocyte samples examined by RRBS. Below is indicated the average number of CpGs analyzed in each category between biological replicates (***: p<0.001, χ2 test).

Supplementary Figure 7: H3K4me3 analysis in growing oocytes. a, Characteristics of H3K4me3 micro-ChIP-seq in oocytes. **Left**: sequencing read trend over promoters in the H3K4me3 ChIP-seq performed in oocytes (d15) compared to ES cells. ES cell H3K4me3 ChIP-seq data were obtained from public databases: (1) GSM594581 and (2) GSM535982. **Right**: H3K4me3 enrichment in promoters of low or highly transcribed genes in growing oocytes as determined by mRNA-Seq. (***: $p<0.001$, Mann-Whitney U test) **b,** Correlation between CpG methylation, mRNA expression and H3K4me3 enrichment. A region from chromosome (ch17:25,077,852-25,389,352) is displayed. Gene symbols and organization are given at the top, with genes transcribed from left to right in red and from right to left in blue. CpG methylation (GV oocytes, upper track presented as log of methylated CpGs divided by unmethylated CpGs), mRNA levels (d10 growing oocytes, middle track), and H3K4me3 enrichment (d15 growing oocytes, lower track, corrected for input background) are shown. The CGI marked in red is methylated, those marked blue are unmethylated, and those marked in grey are unknown. **c,** Identification of alternative promoters in growing oocytes. A selection of genes with a promoter methylated CGI for which an alternative upstream promoter and transcription across the CGI were inferred from mRNA-Seq in d10 oocytes is shown. For each gene, exon organization and CGI location are given on top. H3K4me3 reads (d15 growing oocytes) are represented in black. Green bars show mRNA-Seq (single-read) with reads spanning exon:intron junctions being split. Below, in blue, the reads spanning exon:intron junctions are linked by horizontal lines. In all three cases, transcription start sites upstream of the annotated promoter were inferred from the mapped mRNA-Seq reads and were found to be associated with enrichment for H3K4me3 reads.

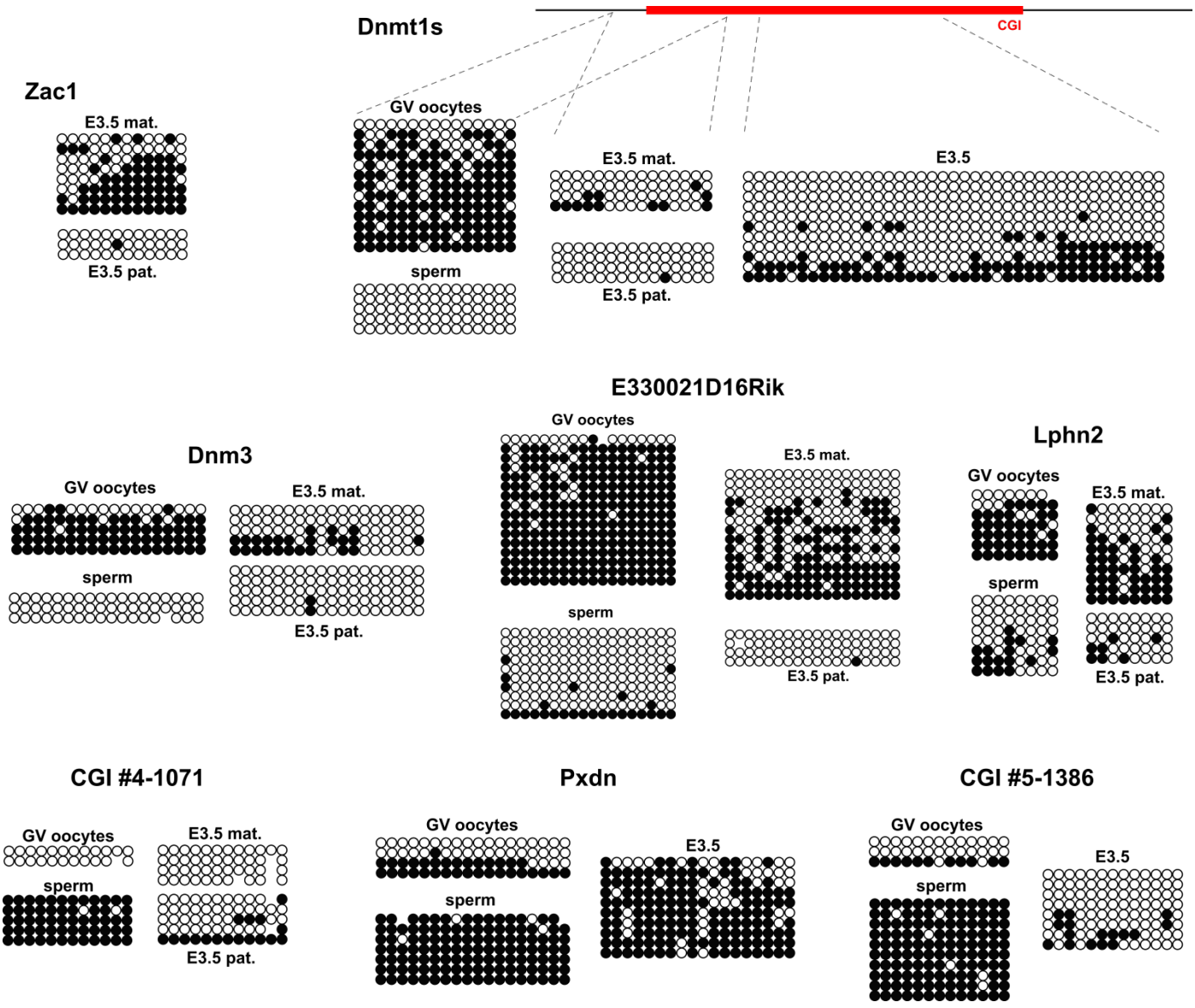


Supplementary Figure 7

Supplementary Figure 8: Fate of gamete-derived CGI methylation in blastocysts. a-b, Distribution of CpG methylation levels across the genome (**a**-left), overlapping CGIs (**a**-right), and CGI methylation levels (**b**) in immature (d5), mature (GV and MII) oocytes, sperm and E3.5 blastocysts. This analysis reflects the overall low level of methylation in blastocysts as a consequence of post-fertilisation reprogramming of the genome. The number of CpGs and CGIs analysed is indicated in Suppl. Fig. 1b (***: $p<0.001$, $\chi 2$ test). **c**, Comparison of the distribution of CGI methylation in blastocysts between all CGIs (n=13875) and the CGIs also methylated in mature oocytes (n=890) (***: $p<0.001$, $\chi 2$ test). **d,** Distribution of CGI methylation level in blastocysts according to their methylation level in gametes (left, MII oocytes *v.* blastocyst comparison; right, sperm *v*. blastocyst comparison) showing that the majority of CGIs methylated between 25-40% in blastocysts are fully methylated in MII oocytes (***: $p<0.001$, $\chi 2$ test) . **e,** Genomic localization of CGIs methylated in MII oocytes (75-100%) classified according to their methylation level in blastocysts (the number of CGIs in each class is given in the bottom row of the table in **d**). **f,** CpG and GC contents, CpG density, length and CpG periodicity of CGIs methylated in MII oocytes (75-100%) classified according to their methylation level in blastocysts (* $p<0.05$; **: $p<0.01$; ***: $p<0.001$, Kruskall-Wallis test). All four classes of CGI have an underlying 3bp periodicity, the more prominent peaks at 6 and 12bp in the 60-75% methylated category may be a reflection of the much lower number of CGIs in this group (n=11). **g,** Bisulphite sequencing in GV oocytes, sperm and E3.5 blastocysts of three CGIs identified by RRBS as examples of CGIs *de novo* methylated in blastocysts. Open circles represent unmethylated CpGs and dark circles methylated CpGs.



Supplementary Figure 8



Supplementary Figure 9: Parental-allele methylation analysis in blastocysts of CGIs methylated in gametes. Bisulphite sequencing in GV oocytes, sperm and hybrid E3.5 blastocysts (C57BL/6J x CAST/Ei) of seven CGIs identified by RRBS as methylated in oocytes or sperm and presenting signs of methylation in blastocysts. When possible methylation in blastocysts was discriminated between parental alleles using SNPs between C57BL/6J and CAST/Ei strains. Bisulphite sequencing of the known maternal germline DMR of the imprinted locus *Zac1* is provided as a control for allele discrimination. Open circles represent unmethylated CpGs and dark circles methylated CpGs.

Supplementary Table 2: CGIs called methylated in MII but not GV oocytes

CGI-ID	Position	Name
#1-53	Chr1:16082890-16083575	4930444P10Rik
#1-500	Chr1:88199490-88200753	B3gnt7
#1-529	Chr1:89405790-89407455	Ngef
#1-854	Chr1:154246932-154247827	Glt25d2
#2-94	Chr2:18602648-18602969	Bmi1
#2-112	Chr2:20441109-20441627	Etl4
#2-145	Chr2:24909714-24910543	Nelf
#2-1124	Chr2:130461621-130461897	Prosapip1
#2-1221	Chr2:146337107-146338192	A230067G21Rik
#2-1387	Chr2:155644583-155645132	BC029722
#3-14	Chr3:8462705-8463080	AC130716.3-1
#3-117	Chr3:32607238-32608341	Actl6a
#3-233	Chr3:54589939-54590325	Smad9
#3-574	Chr3:95491434-95492179	Adamtsl4
#3-788	Chr3:116125897-116127278	Cdc14a
#4-507	Chr4:100240600-100240790	RP23-309M8.3
#4-818	Chr4:125735293-125736156	1810007P19Rik
#4-1102	Chr4:135553348-135553552	RP23-161N17.11
#4-1151	Chr4:137285069-137285413	Rap1gap
#4-1266	Chr4:141794251-141795651	9030409G11Rik
#4-1402	Chr4:151489885-151490406	Tnfrsf25
#5-273	Chr5:36924014-36925294	Tbc1d14
#5-327	Chr5:44491690-44492943	Prom1
#5-996	Chr5:121056572-121056872	Iqcd
#5-1474	Chr5:144131518-144132343	Grid2ip
#5-1478	Chr5:144148813-144149306	Grid2ip
#6-55	Chr6:17230862-17231594	Cav2
#6-101	Chr6:28781396-28784050	Snd1
#6-211	Chr6:40275285-40275830	Agk
#6-239	Chr6:47892689-47893724	RP23-14P17.3
#6-274	Chr6:50546259-50547093	4921507P07Rik
#6-503	Chr6:83958370-83959094	Dysf
#6-637	Chr6:93628414-93629018	Magi1
#6-716	Chr6:113232225-113232817	Cpne9
#6-717	Chr6:113233064-113233468	Cpne9
#7-48	Chr7:4635940-4636155	1500019G21Rik
#7-110	Chr7:7230439-7231264	null
#7-250	Chr7:19742677-19743120	Gipr
#7-887	Chr7:52970340-52970892	Sphk2
#7-1149	Chr7:88049665-88050099	Nmb
#7-1529	Chr7:134164407-134164943	null
#7-1816	Chr7:152392122-152392463	null
#8-2	Chr8:3278976-3279870	Insr
#8-172	Chr8:26628949-26630497	Fgfr1
#8-437	Chr8:72853777-72855732	Gdf1
#8-693	Chr8:87225465-87225908	Lyl1
#8-699	Chr8:87295836-87296350	Nfix
#8-712	Chr8:87417301-87417693	Gcdh
#8-927	Chr8:108137095-108137462	2310066E14Rik
#8-1191	Chr8:124276786-124277800	Jph3
#9-1008	Chr9:108017188-108018041	Bsn
#9-1030	Chr9:108499246-108500191	4933406E20Rik
#10-47	Chr10:11211772-11211987	null
#10-366	Chr10:61440407-61442367	Col13a1
#10-576	Chr10:79433392-79434468	Grin3b
#10-1085	Chr10:126527634-126529266	Agap2
#11-230	Chr11:30654829-30655552	RP23-358M10.3
#11-356	Chr11:50048163-50048348	Mgat4b
#11-889	Chr11:78156836-78157808	Unc119
#11-1056	Chr11:87983090-87983692	Cuedc1
#11-1062	Chr11:88162314-88162477	Msi2
#11-1130	Chr11:95084590-95085222	null
#11-1768	Chr11:117734158-117734890	Tha1
#12-306	Chr12:58676819-58677861	4921506M07Rik
#12-351	Chr12:70891459-70892387	Cdkl1
#12-463	Chr12:81535706-81538109	Wdr22
#13-289	Chr13:44964945-44965478	Jarid2
#13-347	Chr13:49515995-49517368	Ippk
#13-864	Chr13:113779672-113780228	Ccno
#13-897	Chr13:118009260-118010149	Emb
#14-128	Chr14:31362166-31363276	Tkt
#15-533	Chr15:83180100-83181117	Arfgap3
#15-551	Chr15:84276205-84277179	1810041L15Rik
#15-749	Chr15:98439164-98440574	null
#16-164	Chr16:18836548-18837113	2510002D24Rik
#16-592	Chr16:93919450-93920119	Cldn14
#17-474	Chr17:34087856-34088649	B3galt4
#17-562	Chr17:35836836-35838077	Ddr1
#17-938	Chr17:72131437-72132831	Clip4
#18-211	Chr18:36829685-36830101	Sra1
#18-321	Chr18:38791106-38791889	Arhgap26
#18-457	Chr18:61866679-61867649	Pcyox1l
#18-489	Chr18:65243177-65243938	Nedd4l
#18-589	Chr18:76320476-76321086	null
#18-637	Chr18:80432955-80433717	null
#19-238	Chr19:10115401-10117049	Fads3
#X-25	ChrX:7455155-7456002	Pim2
#X-78	ChrX:13063026-13064168	Nyx
#X-507	ChrX:166427875-166428454	Mid1

Supplementary Table 3: EpiGRAPH analysis results. (1/2)

List of attributes found significantly different (Bonferroni) between methylated and unmethylated CGIs in matures oocytes

Rank	Variable Name	Attribute Group Name	Unmethylated CGIs	Methylated CGIs
1	overlapAverageSize	Chromosome Organisation	10705595.19	6461661.881
2	Pat AA skew	DNA Sequence	0.955069816	0.625072647
3	Pat AA std	DNA Sequence	0.063486415	0.055676137
4	Pat AT skew	DNA Sequence	1.060542974	0.657286929
5	Pat CC freq	DNA Sequence	0.221270612	0.206414565
6	Pat CC std	DNA Sequence	0.083932335	0.077408841
7	Pat CG freq	DNA Sequence	0.079160103	0.063244462
8	Pat CG std	DNA Sequence	0.050575328	0.040954385
9	Pat GC std	DNA Sequence	0.052859708	0.048111884
10	Pat TA skew	DNA Sequence	1.063245727	0.787711653
11	Pat CCCG freq	DNA Sequence	0.021659431	0.016151509
12	Pat CCGA freq	DNA Sequence	0.007996026	0.006398444
13	Pat CCGC freq	DNA Sequence	0.023424418	0.016408833
14	Pat CCGG freq	DNA Sequence	0.010874528	0.008504384
15	Pat CGAG freq	DNA Sequence	0.009955094	0.008218715
16	Pat CGCC freq	DNA Sequence	0.021591365	0.015001292
17	Pat CGCG freq	DNA Sequence	0.008801869	0.004856313
18	Pat CGGA freq	DNA Sequence	0.010645563	0.008428885
19	Pat CGGC freq	DNA Sequence	0.019376091	0.014305246
20	Pat CGTC freq	DNA Sequence	0.006285175	0.005296771
21	Pat GCCC freq	DNA Sequence	0.023043787	0.020259311
22	Pat GCGA freq	DNA Sequence	0.00791568	0.006414888
23	Pat GCGC freq	DNA Sequence	0.010591224	0.007348892
24	Pat TAAA freq	DNA Sequence	0.002946379	0.002176811
25	Pat TTAA freq	DNA Sequence	0.00125891	0.000961393
26	Pat plusT std	DNA Sequence	0.081247266	0.074547681
27	CpG content	DNA Sequence	0.079075993	0.063108689
28	CpG obs vs exp ratio	DNA Sequence	0.767805874	0.654946153
29	CpG vs TpG v CpA ratio	DNA Sequence	0.082466959	0.069406064
30	GC content	DNA Sequence	0.642435855	0.623938512
31	length	DNA Sequence	1201.369	688.275
32	overlapAverageSize	DNA Sequence	1791.370683	681.032777
33	rise skew	DNA Structure	0.016095465	-0.055183519
34	roll std	DNA Structure	0.759912131	0.73350349
35	slide std	DNA Structure	0.275075275	0.249973987
36	overlapRegionsCount	Repetitive DNA	0.964534924	0.583640044
37	overlapTotalLength	Repetitive DNA	52.25925152	37.06730018
38	repClass Low complexity overlapRegionsCount	Repetitive DNA	0.469526445	0.186778981
39	repClass Low complexity overlapTotalLength	Repetitive DNA	23.45723149	8.619673696
40	repClass Simple repeat overlapTotalLength	Repetitive DNA	16.34739213	10.24694697
41	repFamily Low complexity overlapRegionsCount	Repetitive DNA	0.469526445	0.186778981
42	repFamily Low complexity overlapTotalLength	Repetitive DNA	23.45723149	8.619673696
43	repFamily Simple repeat overlapTotalLength	Repetitive DNA	16.34739213	10.24694697
44	Pat CA freq	DNA Sequence	0.120096199	0.140957304
45	overlapRegionsCount	Chromosome Organisation	1.115281041	2.18157639
46	overlapRegionsCount	DNA Sequence	1.115281041	2.18157639
47	rise	DNA Structure	3.239647495	3.243896889
48	Pat AG freq	DNA Sequence	0.14283668	0.150035999
49	Pat AGCA freq	DNA Sequence	0.009047981	0.011658973
50	Pat GGCA freq	DNA Sequence	0.010662241	0.013244325
51	Pat CAGC freq	DNA Sequence	0.017620536	0.021077014
52	Pat ACAG freq	DNA Sequence	0.007769387	0.009637252
53	Pat CACA freq	DNA Sequence	0.007901329	0.0102523
54	Pat AC freq	DNA Sequence	0.091446429	0.098133875
55	Pat CATC freq	DNA Sequence	0.005660264	0.007796345
56	Pat CCAG freq	DNA Sequence	0.016639784	0.019102656
57	Pat CTGC freq	DNA Sequence	0.017584873	0.020841528
58	Pat ATGC freq	DNA Sequence	0.004139499	0.005701411
59	Pat ACCA freq	DNA Sequence	0.005804815	0.007339142
60	repClass Simple repeat overlapRegionsCount	Repetitive DNA	0.317235256	0.208833879
61	repFamily Simple repeat overlapRegionsCount	Repetitive DNA	0.317235256	0.208833879
62	Pat ACTG freq	DNA Sequence	0.006553573	0.008121666
63	Pat CCCC freq	DNA Sequence	0.022613672	0.019456372
64	Pat plusA std	DNA Sequence	0.080005144	0.073808545
65	Pat CCAC freq	DNA Sequence	0.01202238	0.014020661
66	Pat CAGG freq	DNA Sequence	0.016043236	0.018243793
67	Pat AAAA freq	DNA Sequence	0.005543419	0.004309761
68	roll skew	DNA Structure	0.497585811	0.430738767
69	Pat AT freq	DNA Sequence	0.023468181	0.026976425
70	Pat CGAC freq	DNA Sequence	0.005373276	0.004546125
71	Pat plusG std	DNA Sequence	0.100160351	0.093670045
72	Pat plusG freq	DNA Sequence	0.3230947	0.310043947
73	G content	DNA Sequence	0.3230947	0.310043947
74	Pat GCAC freq	DNA Sequence	0.009402311	0.011343088
75	Pat TGCA freq	DNA Sequence	0.003953584	0.00512498
76	overlapAverageSize	Repetitive DNA	92.22315864	67.81075747
77	Pat plusC std	DNA Sequence	0.099238371	0.092018063
78	gieStain gneg overlapRegionsCount	Chromosome Organisation	0.535981578	1.165692885
79	Pat GGGA freq	DNA Sequence	0.016798222	0.014991824
80	Pat CATG freq	DNA Sequence	0.002137313	0.003168427

Supplementary Table 3: EpiGRAPH analysis results. (2/2)

List of attributes found significantly different (Bonferroni) between methylated and unmethylated CGIs in matures oocytes

Rank	Variable Name	Attribute Group Name	Unmethylated CGIs	Methylated CGIs
81	Pat plusT skew	DNA Sequence	0.481806397	0.258310883
82	Pat GCCA freq	DNA Sequence	0.012036689	0.013984565
83	Pat ACAT freq	DNA Sequence	0.002526196	0.003743113
84	Pat plusA freq	DNA Sequence	0.179310416	0.1887325
85	Pat CTAA freq	DNA Sequence	0.002882143	0.002400916
86	tilt	DNA Structure	0.094447479	0.086074636
87	Pat GC freq	DNA Sequence	0.105110503	0.099703355
88	Pat plusA skew	DNA Sequence	0.451252956	0.255388923
89	Pat TA std	DNA Sequence	0.023764427	0.021898814
90	Pat plusT freq	DNA Sequence	0.178253729	0.187328988
91	Pat TCCA freq	DNA Sequence	0.00870772	0.01005179
92	Pat AATA freq	DNA Sequence	0.001868996	0.001564623
93	Pat AATT freq	DNA Sequence	0.001148681	0.000969246
94	Pat ATGA freq	DNA Sequence	0.002950817	0.004179948
95	Pat AAAT freq	DNA Sequence	0.003139183	0.002678546
96	Pat ATTA freq	DNA Sequence	0.001436496	0.001281888
97	Pat CTCA freq	DNA Sequence	0.008387812	0.00972922
98	Pat ATAA freq	DNA Sequence	0.001789872	0.001512015
99	Pat CAGA freq	DNA Sequence	0.009901594	0.011217872
100	Pat ACGC freq	DNA Sequence	0.006914657	0.006136754
101	tilt skew	DNA Structure	0.14849511	0.189319768
102	Pat GTTA freq	DNA Sequence	0.002183103	0.001942821
103	slide	DNA Structure	-0.332162756	-0.319615375
104	Pat AGCG freq	DNA Sequence	0.010963887	0.009995597
105	Pat GAAA freq	DNA Sequence	0.006524919	0.005753072
106	Pat CACC freq	DNA Sequence	0.012434776	0.013930578
107	Pat CTAG freq	DNA Sequence	0.002670177	0.002328242
108	slide skew	DNA Structure	-1.73035479	-1.592463996
109	Pat CGAA freq	DNA Sequence	0.00375897	0.003354616
110	Pat CTCC freq	DNA Sequence	0.020551259	0.019032286
111	SASA skew	DNA Structure	-0.048356662	-0.024561818
112	Pat AGTG freq	DNA Sequence	0.007021139	0.008182328
113	repClass SINE overlapTotalLength	Repetitive DNA	9.794820036	6.757944301
114	twist skew	DNA Structure	-0.333170286	-0.282763399
115	twist std	DNA Structure	0.909709936	0.892561103
116	Pat TATA freq	DNA Sequence	0.000485684	0.000491342
117	repClass SINE overlapRegionsCount	Repetitive DNA	0.138495281	0.099704182
118	Pat AAAG freq	DNA Sequence	0.006873441	0.006293014
119	Pat GACC freq	DNA Sequence	0.00943862	0.008752898
120	Pat ATGG freq	DNA Sequence	0.005362876	0.006386437
121	Pat GGCC freq	DNA Sequence	0.010378475	0.009544409
122	Pat CA skew	DNA Sequence	0.202998138	0.063710259
123	Pat GACA freq	DNA Sequence	0.006430135	0.007473398
124	Pat ACCG freq	DNA Sequence	0.007112833	0.006672139
125	Pat AAAC freq	DNA Sequence	0.004736722	0.004321643
126	Pat CACG freq	DNA Sequence	0.007186778	0.008267702
127	Pat ATCA freq	DNA Sequence	0.002878723	0.003695563
128	Pat AGCT freq	DNA Sequence	0.004947546	0.005751943
129	shift skew	DNA Structure	0.41544756	0.37805301
130	twist	DNA Structure	35.07221165	35.03861044
131	Pat GTGA freq	DNA Sequence	0.00645994	0.007453373
132	Pat AATC freq	DNA Sequence	0.003016919	0.002785156
133	Pat ACAC freq	DNA Sequence	0.006651948	0.00779771
134	Pat AG skew	DNA Sequence	0.142491037	0.013093388
135	Pat CA std	DNA Sequence	0.047198052	0.046493703
136	Pat CTGA freq	DNA Sequence	0.008107034	0.00905534
137	Pat ACCT freq	DNA Sequence	0.007197272	0.007944991
138	SASA	DNA Structure	0.423686073	0.417767217
139	Pat AC skew	DNA Sequence	0.360008131	0.235919926
140	Pat TCGA freq	DNA Sequence	0.001413107	0.001439582
141	Pat TGAA freq	DNA Sequence	0.004469861	0.005352819
142	Pat AACG freq	DNA Sequence	0.003222458	0.002999586
143	Pat AG std	DNA Sequence	0.048149472	0.0470305
144	Pat GGAC freq	DNA Sequence	0.01092201	0.010373888

Supplementary Table 4

Bisulphite Sequencing PCR primers

CpG island	Forward	Reverse
Syt2	ATTTTTTAGAGATTTGGTGTAGTGGTT	ATCATTTCAAAAATTACAAACTAAAC
Tuba3a	GAGGAAGTTTTAGTTTTAGAGGAGA	CCAATCAAAAAATACCAACACTACA
Mmp2	GGGATTGTTAGGATTTGTAAGTATT	ACCCAACAAACAACACAAAAC
Lhcgr	TGTTTTTTAGTAATGGGGTTTAATAT	TACACTCTCCAAAATTATCAAAATC
Zac1	TGGGTGCCCCAGTTGCAGCCAGAGATGCAG	CAAAAAAACACAAATCACCTCTTC
Dnmt1s	CAGAGTTTTTTTGATAGTTTTAGGAT	AATATACAAAAAATACCCCTCTCCC
	GAGGGTTGATTTTTTTGTTTTTTTA	TCCCCCACTCTCTTACCCTATATA
Dnm3	ATTGTGTGATTTGGTAGGTAGTTTAG	CCAATAAAATTAAAAACTCTAAAAAAAA
E330021D16Rik	GATTAGTTATTTGAGTTGTTGATTAAG	CTAAAATTCCTAACATCCATCTTCC
Lphn2	GTGGAGAATTATGGGTTAATAGAGT	AAATATACCCAACCTTAAAACTAACC
CGI 4-1071	GGAAGGTTATATGTGATTTTAAGTTTT	ATCTTCCAAAAATATTTCTCTAACC
Pxdn	GGTTATGATTTTTTTTAATGATTTT	AAACTTACCAAACCTCTACACAATACC
CGI 5-1386	TTTTTTTTGTAATTGGTAATTTTGTTATAT	AACCTCTCCTAACTACCCAATTCTC