Research

Authors for correspondence:
Yongchun Zuo
e-mail: yczuo@imu.edu.cn
Guangpeng Li
e-mail: gpengli@imu.edu.cn

†These authors contributed equally to this study.

THE ROYAL SOCIETY
PUBLISHING

DNA methylation subpatterns at distinct regulatory regions in human early embryos

Rongsong Luo[1,2,†], Chunling Bai[1,2,†], Lei Yang[1,2], Zhong Zheng[1,2], Guanghua Su[1,2], Guangqi Gao[1,2], Zhuying Wei[1,2], Yongchun Zuo[1,2] and Guangpeng Li[1,2]

[1]State Key Laboratory of Reproductive Regulation and Breeding of Grassland Livestock, and [2]College of Life Sciences, Inner Mongolia University, Hohhot 010070, People's Republic of China

RL, 0000-0002-1822-8260

DNA methylation has been investigated for many years, but recent technologies have allowed for single-cell- and single-base-resolution DNA methylation data-sets and more accurate assessment of DNA methylation dynamics at the key genomic regions that regulate gene expression in human early embryonic development. In this study, the region from upstream 20 kb to downstream 20 kb of RefSeq gene was selected and divided into 12 distinct regions (up20, up10, up5, up2, 5'UTR, exon, intron, 3'UTR, down2, down5, down10 and down20). The candidate promoter region (TSS ± 2 kb) was further divided into 20 consecutive subregions, which were termed 'bins'. The DNA methylation dynamics of these regions were systematically analysed along with their effects on gene expression in human early embryos. The dynamic DNA methylation subpatterns at the distinct genomic regions with a focus on promoter regions were mapped. For the 12 distinct genomic regions, up2 and 5'UTR had the lowest DNA methylation levels, and their methylation dynamics were different with other regions. The region 3'UTR had the highest DNA methylation levels, and the correlation analysis with gene expression proved that it was a feature of transcribed genes. For the 20 bins in promoter region, the CpG densities showed a normal distribution pattern, and the trend of the methylated CpG counts was inverse with the DNA methylation levels, especially for the bin 1 (downstream 200 bp of the TSS). Through the correlation analysis between DNA methylation and gene expression, the current study finally revealed that the region bin −4 to 6 (800 bp upstream to 1200 bp downstream of the TSS) was the best candidate for the promoter region in human early embryos, and bin 1 was the putative key regulator of gene activity. This study provided a global and high-resolution view of DNA methylation subpatterns at the distinct genomic regions in human early embryos.

1. Introduction

In mammalian reproduction, a sperm fertilizes an egg, and then the zygote under-goes cleavage and develops into a morula; the inner cells of the morula give rise to the inner cell mass (ICM), from which the embryo proper is derived, whereas the outer cells differentiate into the trophectoderm (TE), which gives rise to extraem-bryonic tissues [1]. This developmental process requires extensive erasure of epigenetic marks, and the dynamics of pre-implantation are driven by the need to reprogram haploid parental epigenomes for the morula to reach a totipotent state in which epigenetic programmes essential for the first-cell lineage commitment and differentiation are established. DNA methylation is the well-studied epigenetic mark and plays multiple critical roles, including transcriptional silencing, genomic imprinting, X chromosome inactivation and repression of transposable elements [2–6]. DNA methylation predominantly occurs on the cytosine in CpG

dinucleotides that are stably distributed throughout mammalian genomes [7–9]. Furthermore, DNA methylation is involved in diverse key cellular processes, such as early embryogenesis, stem cell differentiation, regulation of neuronal development and cancer development [9–11]. During early embryogenesis, cell lineage commitments are accompanied by extensive epigenetic remodelling [5,12]. The greatest reductions in DNA demethylation occur in the zygote and the 2-cell stage during pre-implantation. After implantation, remethylation occurs gradually until the original levels are reached [13,14].

In recent years, advanced techniques have rapidly propelled the field of epigenetics forward and have allowed for genome-wide epigenetic analyses of ever-decreasing amounts of biological materials. These techniques have provided insight into global DNA methylation reprogramming at single-base resolution using single cells at multiple developmental stages during early embryogenesis [6,15–19], allowed for refined assessment of DNA methylation dynamics across genomes during specific developmental stages by evaluating specific classes of DNA elements [13,17,20,21], and allowed for methylation comparisons between early mouse and early human embryos [22]. Although the DNA sequences are identical for all of the cells within a body, epigenetic information differs in every cell and is responsible for the maintenance of different cell types within an organism. However, the extents of methylated DNA sequences differ between cell types and developmental stages. Embryos may display similar phenotypes, but with varying methylations on genomic regions associated with cell proliferation and differentiation [23,24]. Comprehensive analyses of genome-wide DNA methylation variations revealed that DNA methylation is not always negatively correlated with gene expression. In fact, DNA methylation of some loci correlated positively with gene expression [25–27], and for these loci many questions remain to be answered. For example, little is known about the quantitative and specific dynamics of genome-wide DNA methylation, and it is still unknown whether DNA methylation of specific genomic subregions changes during early embryogenesis.

Therefore, genome-wide single-cell- and single-base-resolution datasets and more accurate assessments of DNA methylation dynamics could provide an in-depth understanding of the extent of methylation throughout the genome, as well as the DNA methylation of specific genomic regions involved in the regulation of early embryonic development. In this study, the ±20 kb region of RefSeq genes were selected and divided into 12 regions, and then the promoter regions were divided into 20 consecutive subregions, termed 'bins'. The DNA methylation dynamics of these regions were systematically analysed along with their effects on gene expression. Finally, we mapped the dynamic DNA methylation subpatterns at these genomic subregions. This study provided a high-resolution and global view of DNA methylation subpatterns at the distinct genomic regions in human early embryos.

2. Results

2.1. The dynamics of global DNA methylation in human early embryos

Analysis of global DNA methylation level indicated two distinct clusters: the pre-implantation stages and the post-implantation stage (figure 1a). The DNA methylation level decreased from the zygote to the two-cell stage, remained relatively constant until the ICM of blastocyst, and then dramatically increased at post-implantation stage (figure 1b). When DNA methylation level was calculated at the chromosomal level, similar trends were observed (electronic supplementary material, figure S1). To further understand the methylation dynamics of the differentially methylated genes, all the RefSeq genes were classified as hyper-methylated, moderately methylated or hypo-methylated genes. The data showed that most of the genes were moderately methylated or hypo-methylated at the pre-implantation stages. But for the post-implantation embryos, most of the genes were hyper-methylated or moderately methylated. The number of hyper-methylated genes in post-implantation embryos was significantly higher than pre-implantation embryos (figure 1c). During the pre-implantation developmental stages, most of the genes maintained their hypo-methylation statuses, only a few genes changed from a hypo-methylated state to a hyper-methylated state or from a hyper-methylated state to a hypo-methylated state. When the methylation states of embryos at any of the pre-implantation stages were compared to the post-implantation stage, 396–1130 genes changed from the hypo-methylated state to the hyper-methylated state, whereas only a few genes changed from the hyper-methylated state to the hypo-methylated state (figure 1d).

2.2. DNA methylation dynamics at the 12 distinct genomic regions

To gain insight into the dynamics of DNA methylation at the 12 genomic regions, we analysed the distribution of DNA methylation across CpG clusters within a certain region. From up20 to down20, all the developmental stages shared the same DNA methylation trend, the changes of DNA methylation levels were first decreased then increased gradually (figure 2a). For the intergenic regions up20 to up2 and down2 to down20, the up2 region had the lowest DNA methylation level, whereas the other intergenic regions had relatively equivalent DNA methylation levels. For the gene body regions, the DNA methylation level was the lowest at the 5'UTR and the highest at the 3'UTR. However, the DNA methylation levels of the exon and intron were nearly identical (figure 2a,b). In addition, the DNA methylation dynamics of these regions (except up2 and 5'UTR) were similar to the global DNA methylation dynamics (figures 1b and 2b).

Since all the embryos at different developmental stages had similar DNA methylation dynamics from up20 to down20, we then classified the genes into hyper-, moderate- and hypo-methylated states according to the DNA methylation level of each region. The result showed that the up2 and 5'UTR regions were hypo-methylated in all the developmental stages. All the other regions were moderate- or hypo-methylated in pre-implantation embryos, and the numbers of moderate- or hypo-methylated genes did not vary significantly. However, in post-implantation embryos, these regions were almost hyper-methylated (figure 2c). Notably, the up2 and 5'UTR had the high CpG counts with low DNA methylation levels, but it was totally the opposite for the 3'UTR (figure 2d).

2.3. DNA methylation dynamics at the 20 consecutive bins within the promoter region

DNA methylation of the promoter region regulates gene expression. In most cases, DNA methylation leads to gene
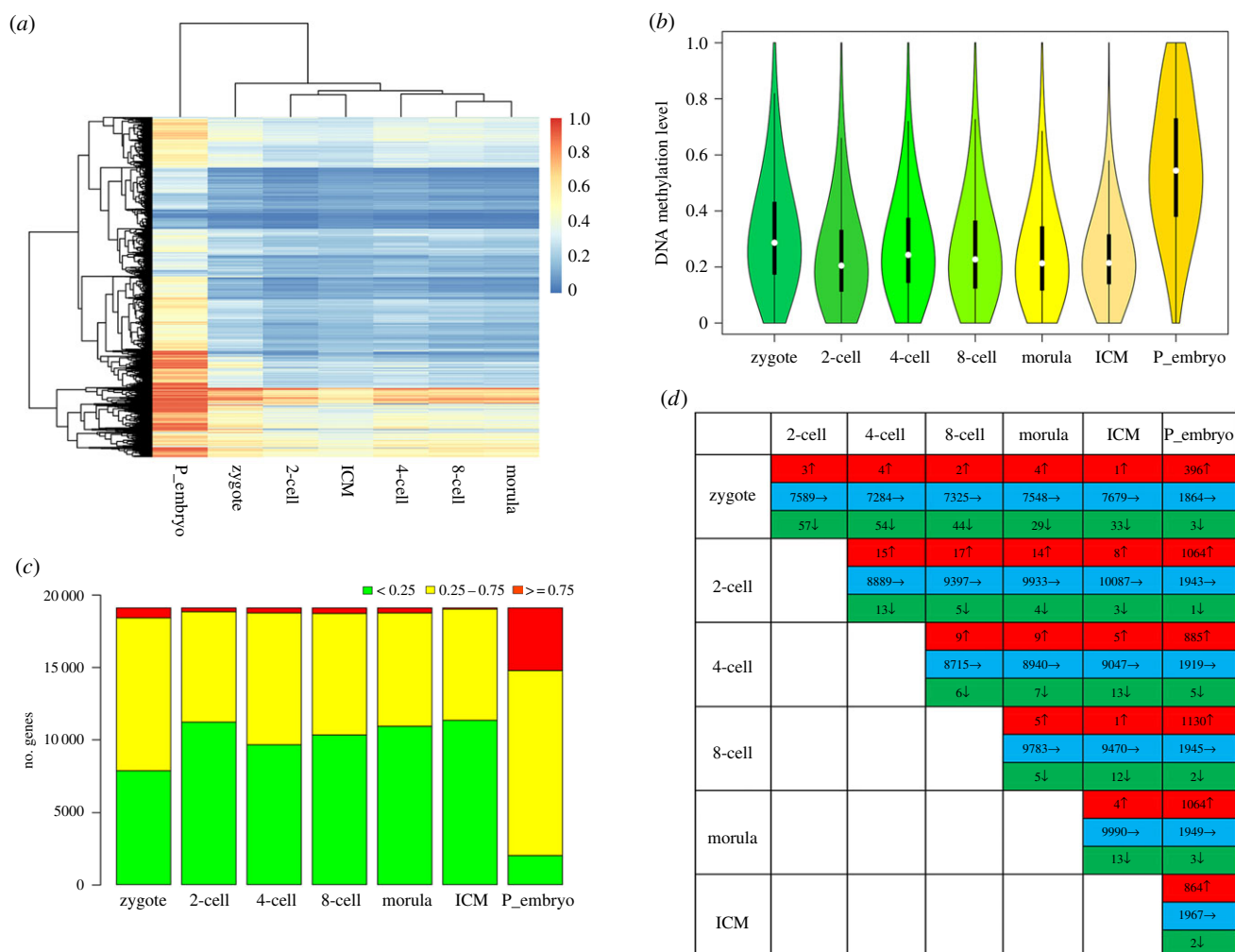
Figure 1. The dynamics of global DNA methylation levels in human early embryos. (*a*) The heat map of DNA methylation of pre- and post-implantation human embryos exhibited two distinct clusters. (*b*) Violin plots of the global DNA methylation dynamics of human embryos. (*c*) Distribution of genes with hyper, moderate or hypo methylation levels at different developmental stages. Red, yellow and green represent genes with hyper, moderate and hypo methylation levels, respectively. (*d*) Differences in the number of hyper- and hypo-methylated genes between pairs of developmental stages. Red, blue and green represent genes whose DNA methylation status went from hypo-methylated to hyper-methylated, maintained hypo-methylation, and went from hyper-methylated to hypo-methylated, respectively. Where the P_embryo is represent the post-implantation embryos.

Panel (d):

	2-cell	4-cell	8-cell	morula	ICM	P_embryo
zygote	3↑ / 7589→ / 57↓	4↑ / 7284→ / 54↓	2↑ / 7325→ / 44↓	4↑ / 7548→ / 29↓	1↑ / 7679→ / 33↓	396↑ / 1864→ / 3↓
2-cell		15↑ / 8889→ / 13↓	17↑ / 9397→ / 5↓	14↑ / 9933→ / 4↓	8↑ / 10087→ / 3↓	1064↑ / 1943→ / 1↓
4-cell			9↑ / 8715→ / 6↓	9↑ / 8940→ / 7↓	5↑ / 9047→ / 13↓	885↑ / 1919→ / 5↓
8-cell				5↑ / 9783→ / 5↓	1↑ / 9470→ / 12↓	1130↑ / 1945→ / 2↓
morula					4↑ / 9990→ / 13↓	1064↑ / 1949→ / 3↓
ICM						864↑ / 1967→ / 2↓

silencing. To obtain detailed information regarding DNA methylation in promoter region, the TSS $\pm$ 2 kb was defined as the promoter region and was subdivided into 20 consecutive bins, which were classified into three subtypes—high, intermediate and low CpG (HCG, ICG, and LCG, respectively)—based on their CpG density according to the reference sequence (figure 3*a*; electronic supplementary material, figure S2). The data showed that bin -10 and bin -9 were LCG and the other bins were HCG (bin -2, -1, 1, 2, 3, 4) or ICG (the rest of bins). Notably, bin1 which is 200 bp downstream of the TSS had the highest CpG density (figure 3*b*). Then we pooled and counted the number of methylated bins (had methylated CpGs in a certain bin), and the corresponding DNA methylation level were calculated. For the three subtypes of bins, the HCG bins had the lowest DNA methylation levels, while the ICG bins had moderate DNA methylation levels and LCG bins had the highest DNA methylation levels (electronic supplementary material, figure S4).

The CpG counts of the 20 consecutive bins displayed a regular pattern that increase firstly and then decrease. We then asked whether the CpG counts were associated with the methylation level. The results indicated that the trend of the CpG counts was consistent with the number of methylated bins but was inverse

with its DNA methylation levels (figure 3*b,c*). For the 20 bins, bin -3 to 5 were hypo-methylated in all developmental stages (figure 3*d*), while the other bins had the similar DNA methylation dynamic with the dynamics of genome-wide and the 10 distinct genomic regions (figure 1*b*, 2*b*, 3*d*). All the 20 consecutive bins had similar DNA methylation features for each developmental stage (figure 3*e*). These data indicated that in human early embryos, those bins in promoter region with higher CpG densities had lower DNA methylation levels, whereas the bins with lower CpG densities had higher DNA methylation levels. Moreover, the DNA methylation dynamics in different tissues had the same pattern in these bins (electronic supplementary material, figure S3).

2.4. Relationships between DNA methylation at distinct genomic regions and gene expression

To gain high-resolution information about DNA methylation and regulation of gene expression, the correlations between DNA methylation at the distinct genomic regions and gene expression were analysed along with previously published single-cell transcriptome data [28]. For the gene body regions,
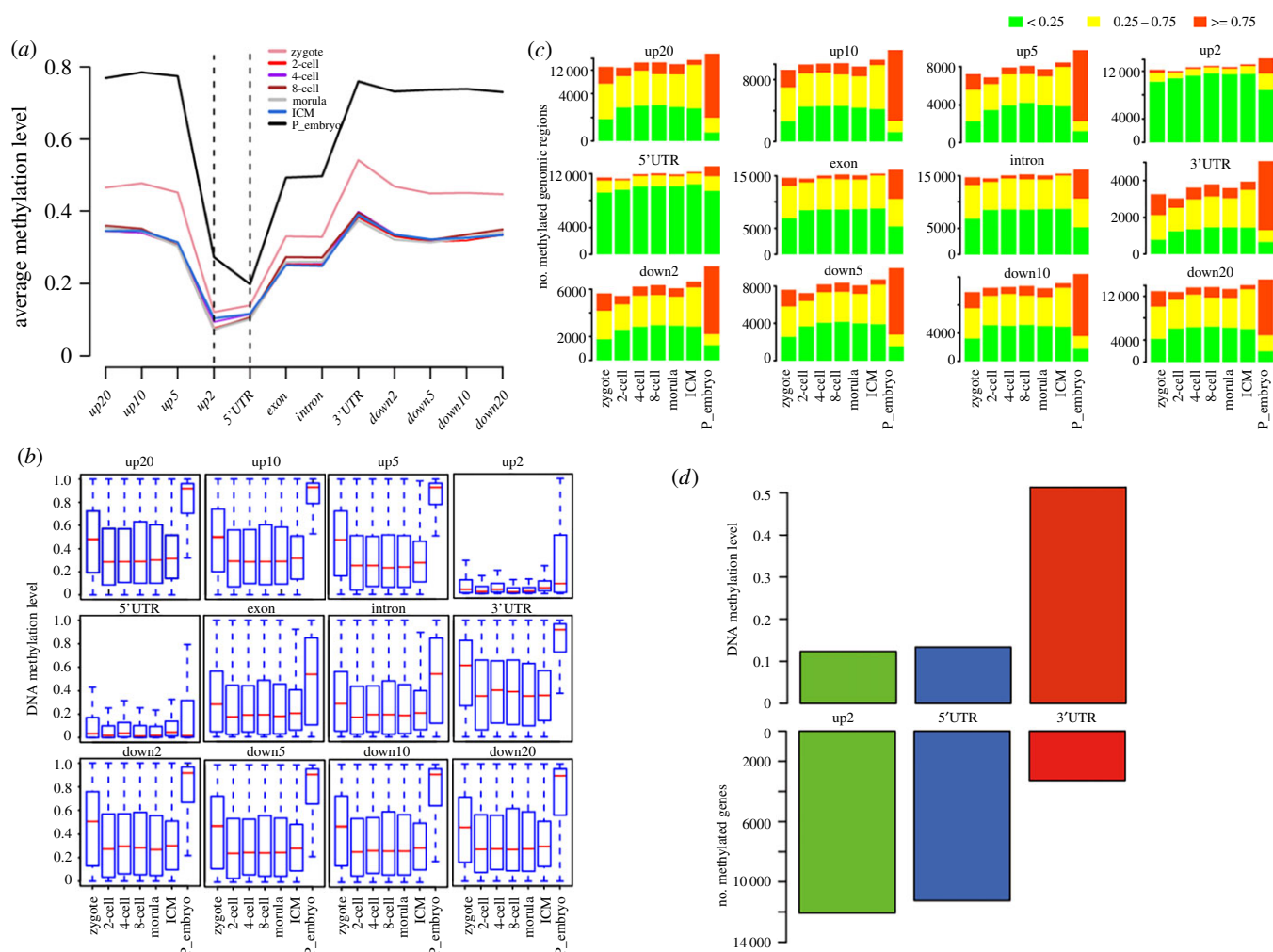
Figure 2. The characteristics of DNA methylation level at the 12 genomic regions in human early embryos. (*a*) Average DNA methylation levels of the 12 genomic regions for each developmental stage. (*b*) Box plots showing the DNA methylation dynamics for each genomic region. (*c*) The number of genes with hyper, moderate or hypo methylation levels at the 12 genomic regions. (*d*) The numbers of methylated genes and DNA methylation levels at the genomic regions up2, 5′UTR and 3′UTR.

the DNA methylation levels and gene expression were negatively correlated in 5′UTR, but positively correlated in 3′UTR. For exon and intron, the DNA methylation levels were positively correlated with gene expression in two-cell, four-cell and eight-cell stages, and negatively correlated in the zygote, morula, ICM and post-implantation stages. It was noteworthy that the correlations were weak from zygote to morula stages. For the intergenic regions, the DNA methylation levels and gene expression were negatively correlated in up2 but positively correlated in the other regions (figure 4*a*; electronic supplementary material, table S1).

We further evaluated the correlations between DNA methylation and gene expression at those bins in promoter region. The results showed that DNA methylation of bins −4 to 6 were negatively correlated with gene expression (figure 4*b*). The principal component analysis (PCA) also demonstrated that these bins clustered together (figure 4*c*). But this situation was not so suitable for somatic tissues (electronic supplementary material, table S1).

2.5. Gene ontology analysis of differentially methylated genes

Considering bin 1 had the highest CpG density and the lower DNA methylation level, there was a strong negative correlation

between DNA methylation and gene expression in all the developmental stages (figure 4*d*). Thus, bin 1 was selected for further analysis and the GO analysis was conducted with the differentially methylated genes within bin 1. A total of 363 genes demonstrated decreased DNA methylation and increased gene expression from ICM to post-implantation embryos (electronic supplementary material, table S2). These genes were enriched for terms associated with cell proliferation and differentiation which could better reflect the transition from pre-implantation to post-implantation, such as cell development, organogenesis, proton transporting and response to hypoxia (electronic supplementary material, table S3).

Since all the developmental stages shared the same DNA methylation trend from up20 to down20, we then analysed CD81 gene, which codes for a protein that mediates signal transduction and plays important role in the regulation of cell development, activation, growth and motility [29]. We highlighted the DNA methylation dynamics at the ±20 kb region of CD81 and performed a more detailed analysis of the region 200 bp downstream of the TSS during all the developmental stages. Similar dynamics of DNA methylation were observed compared with above-mentioned results. Finally, a comprehensive landscape was mapped to demonstrate the DNA methylation levels and the correlations between DNA methylation and gene expression in all the
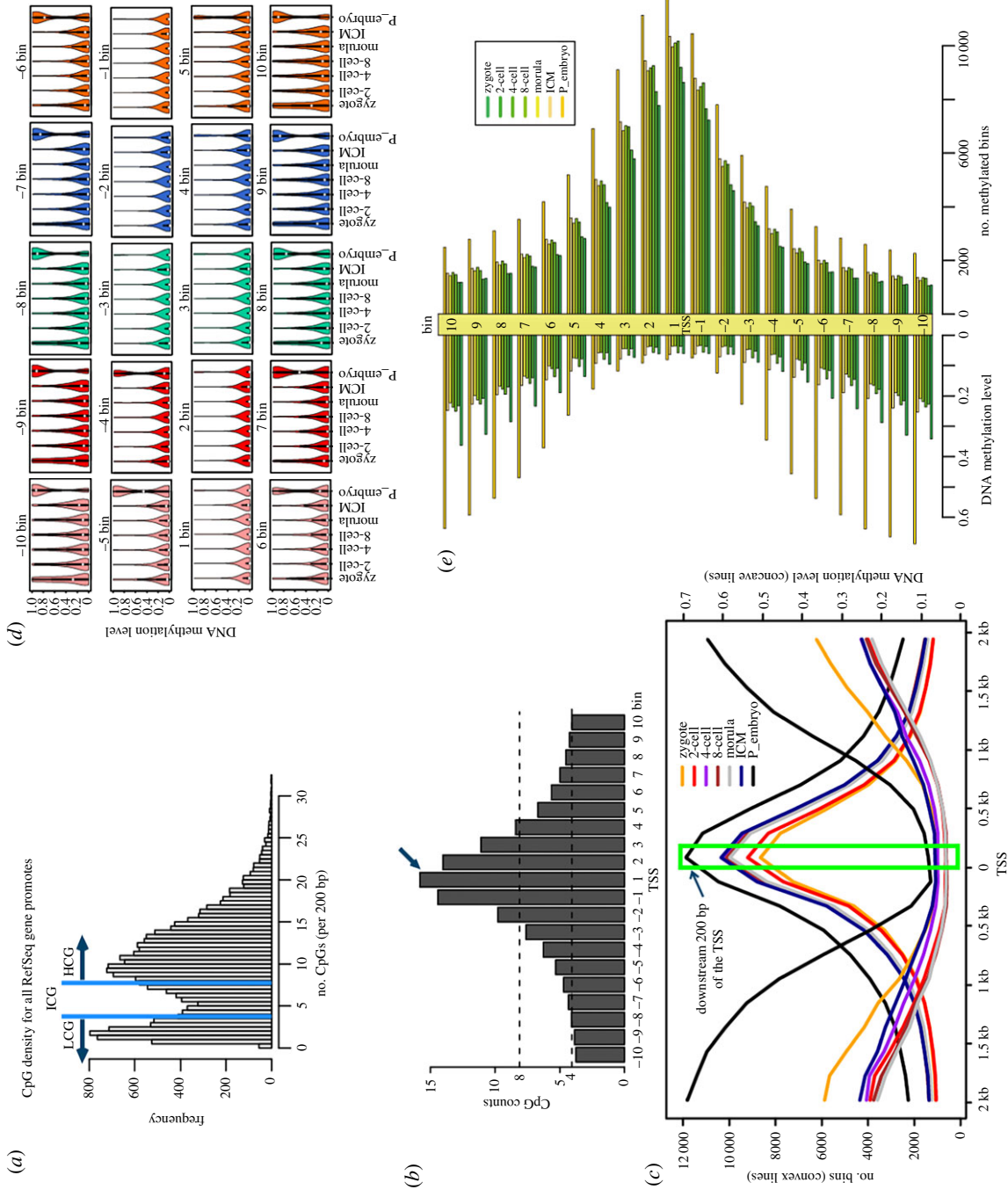
Figure 3. The CpG densities and DNA methylation level of the 20 consecutive bins within promoter region. (*a*) The classification of the CpG densities for all the RefSeq gene promoters. HCG, more than 8 CpG per 200 bp. ICG, 4–8 CpG per 200 bp. LCG, less than 4 CpG per 200 bp. (*b*) The distribution of CpG counts at the 20 bins. (*c*) A line chart showing the status of DNA methylation during human early embryonic development. (*d*) Dynamics of DNA methylation at the 20 bins. (*e*) Comparison of the numbers of methylated bins and the corresponding DNA methylation levels.
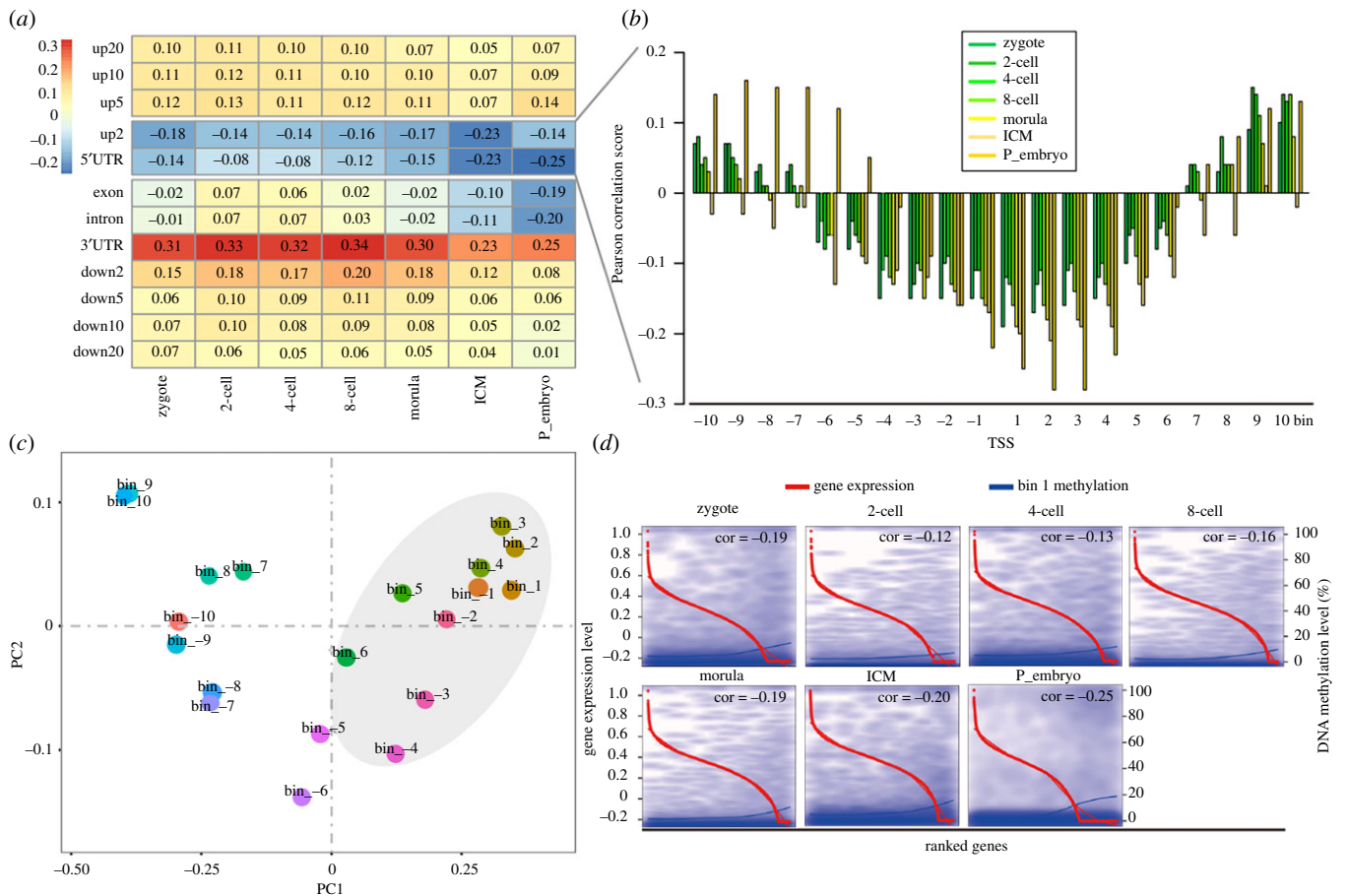
Figure 4. The relationships between DNA methylation level at distinct genomic regions and gene expression. (*a*) The relationships between the DNA methylation level of the 12 genomic regions and gene expression. (*b*) The relationships between the DNA methylation level and gene expression for the 20 consecutive bins. (*c*) PCA showed that the bins from −4 to 6 cluster to gather as depicted by the grey area. (*d*) A scatter plot shows the DNA methylation level and the relative expression level for bin1 of the RefSeq genes. The red and blue curves represent gene expression levels and DNA methylation levels, respectively. Pearson correlation coefficients are marked at the top right corner of each panel. The horizontal axis from left to right below each box represents the RefSeq gene expression level from high to low, respectively.

regions and bins in human early embryos (electronic supplementary material, figure S5).

3. Discussion

DNA methylation has been investigated for many years and is generally considered as a silencing epigenetic mark that regulates gene expression and maintains genome stability [2,30–33]. With the development of high-throughput sequencing technology, genome-wide DNA methylation patterns can be mapped, which facilitates research on the biological functions of DNA methylation in embryonic development, cell differentiation and cancer development.

In present study, we performed an integrated analysis between DNA methylation at distinct genomic regions and gene expression in human early embryos, and mapped DNA methylation patterns at these distinct genomic regions. DNA methylation levels have been shown to correlate with CpG density within distinct genomic regions [34]. The 5′UTR and up2 regions have higher CpG densities and relatively low DNA methylation levels, and are negatively correlated with gene expression. It has been suggested that DNA methylation of the exon and the intron may affect regulation of alternative transcription, exon usage and splicing [35,36]. This analysis indicated that the DNA methylation of the exon and the

intron were both relatively stable and correlated weakly positively or negatively with gene expression, thus they may play the same roles in human early embryos. The 3′UTR that flanks the TES is a region where microRNAs bind to regulate gene activity [37–40]. In our study, the 3′UTR had the highest level of DNA methylation and showed a strong positive correlation with gene expression, indicating that the hypermethylation of the 3′UTR is a feature of transcribed genes. Therefore, the DNA methylation of 3′UTR could be used as a reporter to determine whether a gene is actively transcribed or not. Previous studies suggested that DNA methylation in intergenic regions regulates microRNA expression [41] and that high DNA methylation levels in intergenic regions stabilizes the genome [36]. The DNA methylation levels of the intergenic regions except up2 were relatively high in present study, which may contribute to genomic stability and conservation. Our results regarding the distribution of DNA methylation are consistent with the idea that DNA methylation contributes an additional level of stability to the epigenetic status, and also indicate that the role of DNA methylation has region-specific effects.

As a crucial regulator of gene expression, DNA methylation at the promoter region was thought to suppress gene expression, whereas DNA methylation in the vicinity of the TSS was thought to block transcription initiation [27,28,42,43]. However, the gene promoter region is a fuzzy concept with

varied definitions [13,22,44]. In present study, the 20 consecutive bins within the promoter region had high CpG densities, but exhibited low DNA methylation levels, especially for bins −4 to 6 (800 bp upstream to 1200 bp downstream of the TSS). The DNA methylation levels of these bins were negatively correlated with gene expression during all the developmental stages, indicating that this region was essential for regulating gene expression. This region may be used as the best candidate for the promoter region in human early embryonic development. Furthermore, this situation was not completely suited to different tissues in our study; more data should be added for further comparative analysis. GO enrichment analysis were conducted with the differentially methylated genes identified by the bin 1 between the ICM of blastocyst and post-implantation showed that the enriched genes played key roles in the pre-implantation to post-implantation transition. A previous study indicated that methylation of the region downstream of the promoter may regulate usage of alternative promoters [45]. Since the bins located downstream of the TSS are part of the first exon, the first exon may play an important role in regulating gene activity in human early embryos.

Recently, several genomic chromatin researches for mammalian early embryonic development have been published [46–48]. Gao et al. [49] mapped the chromatin accessibility landscape in human early embryos for the first time, from the perspective of development and evolution to understand human chromatin in human early embryonic development. They analysed the correlations between DHSs and DNA methylation, and their results showed that high CpG promoter regions tend to establish DHSs, whereas just opposite in the low CpG promoter regions [49]. The integrated analysis between DNA methylation and chromatin accessibility based on the same different genomic regions might increase our understanding of epigenetics.

The epigenetic remodelling that occurs during early embryonic development is complex and dynamic. With the development of new technologies and the use of big data, we can perform integrated multi-omics analysis in the future, including DNA methylation, histone modifications, chromatin accessibility and gene expression, to explore lineage- and species-specific epigenetic remodelling from multiple dimensions to improve our understanding of early developmental events. This study demonstrated a method to detect specific methylation across distinct genomic regions and provided a better understanding of DNA methylation dynamics in human early embryonic development and reprogramming.

4. Methods

4.1. Data collection and processing

Human genome assembly (hg19) and gene annotations were downloaded from the UCSC genome browser (http://genome.ucsc.edu/). For genes that have alternative transcripts with the same TSS, only one transcript was chosen, and genes that begin with the mature messenger RNA (MM) were analysed further. Genome-wide DNA methylation datasets at single-cell and single-base resolution of human early embryos at pre-implantation (zygote, 2-cell, 4-cell, 8-cell, morula and ICM of blastocyst), post-implantation stage and different tissues (somatic cell at 7 weeks, 10 weeks and 19 weeks, liver at 6 weeks, PGC at 11 weeks) were

downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE49828 and GSE63818 [13,21]. The datasets for each developmental stage were composed of one to three biologically replicated samples.

4.2. Annotation of genomic regions

To evaluate DNA methylation dynamics at distinct genomic regions in human early embryos, the region ±20 kb of RefSeq genes was selected and divided into gene body and intergenic regions based on gene structure and annotations. The gene body—the genomic region from TSS (transcript start site) of a gene to its TES (transcript end site)—was composed of the 5′UTR (untranslated region), exon, intron and the 3′ UTR. In this study, the 'exon' includes all the exons within a gene and so does the 'intron'. The intergenic region was defined as the complement of the gene body and comprises the upstream 20 kb and the downstream 20 kb of the gene. The intergenic region was then divided into eight regions, including the region 20 kb upstream of the TSS (up20) and the region 0 kb downstream of the TES (down2) and so on. Finally, the DNA methylation dynamics of the 12 distinct genomic regions of the RefSeq genes were analysed individually (figure 5a).

To explore the detailed DNA methylation dynamics of the promoter region, the TSS ± 2 kb region was defined as the promoter region and further binned it into 20 consecutive bins (200 bp for each) based on gene structure and annotations (figure 5b). The number of CpG dinucleotides in each bin was counted. According to the CpG counts, the bins were classified as HCG, ICG and LCG based on thresholds described previously [10].

4.3. Integrated analysis of genome-wide DNA methylation and gene expression

The DNA methylation of the gene body and intergenic regions were summed to determine the global DNA methylation levels of the RefSeq genes. Upon calculation of the DNA methylation levels of these genomic regions, only the CpG sites with read coverage greater than five were analysed further. The average of the DNA methylation level of all the CpGs captured in a region was interpreted as the average DNA methylation level of the region. Analysis of DNA methylation of single CpG was also limited to the CpGs with at least fivefold read coverage. We defined the DNA methylation level of each sample as the arithmetic average of the DNA methylation levels of all the RefSeq genes, whereas the DNA methylation level of each reported developmental stage was defined as the arithmetic average of all the biological replicates.

An integrated analysis of DNA methylation and gene expression was conducted to assess the relationship between DNA methylation and gene expression. The log2 of gene expression levels (reads per kilobase of transcript per million mapped reads, RPKM) of RefSeq genes was calculated, and genes with an RPKM equal to 0.0 were reset to the minimum value of non-zero. The Pearson correlation coefficient (r) between gene expression and DNA methylation was calculated for each genomic region and bin. The genome-wide single-cell gene expression data of human early embryos and different tissues were used for an associated analysis with
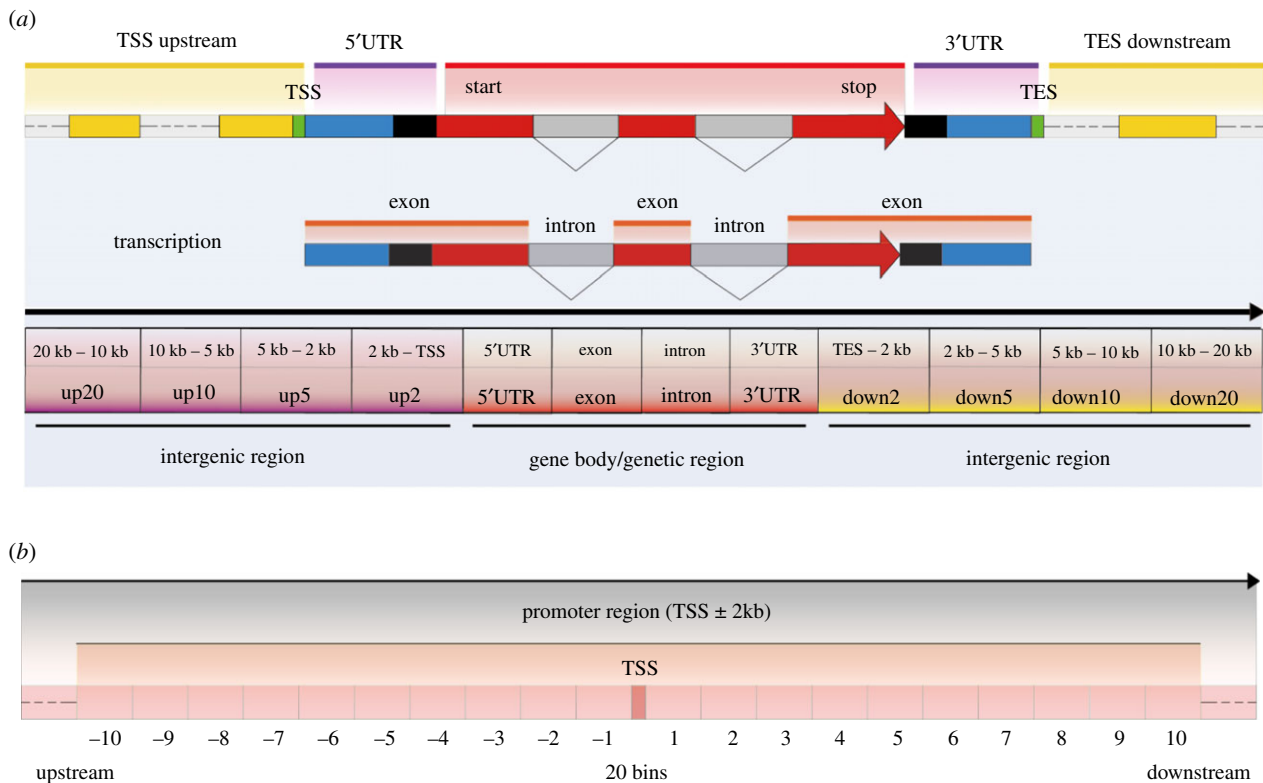
Figure 5. Overview of the distinct genomic regions and the bins within promoter region. (a) The region from 20 kb upstream to 20 kb downstream of RefSeq gene was selected and divided into 12 distinct regions, including the gene body region, which we termed the 5′UTR to 3′UTR regions, and the intergenic region, which we termed the up20 to up2 and the down2 to down20 regions. (b) The TSS $\pm$ 2 kb was defined as candidate promoter region and was further divided into 20 consecutive subregions, termed bins.

DNA methylation. The data calculations and the integrated analysis were performed using customized Python scripts.

4.4. Functional annotation of differentially methylated genes

GO analysis was conducted for the following genes: (i) the DNA methylation level of bin 1 decreased more than twofold from ICM to post-implantation (*t*-test, $p < 0.5$) and for genes in which the absolute DNA methylation value was greater than the mean DNA methylation value; (ii) the gene expression level increased from ICM to post-implantation. The GO analysis was performed with DAVID online (http://david.abcc.ncifcrf.gov/) [50]. A representational item was defined as a modified Fisher's exact *p*-value with an adjustment for multiple tests using the Benjamini

method. Data analysis and visualization were performed using customized R scripts.

References

1. Yamanaka Y, Ralston A, Stephenson RO, Rossant J. 2006 Cell and molecular regulation of the mouse blastocyst. *Dev. Dyn.* **235**, 2301–2314. (doi:10.1002/dvdy.20844)

2. Bird A. 2002 DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21. (doi:10.1101/gad.947102)

3. Hellman A, Chess A. 2007 Gene body-specific methylation on the active X chromosome. *Science* **315**, 1141–1143. (doi:10.1126/science.1136352)

4. Li E, Beard C, Jaenisch R. 1993 Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365. (doi:10.1038/366362a0)

5. Hackett JA, Surani MA. 2013 DNA methylation dynamics during the mammalian life cycle. *Phil. Trans. R Soc. B* **368**, 20110328. (doi:10.1098/rstb.2011.0328)

6. Lister R, Pelizzola M, Dowen RH. 2016 Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322. (doi:10.1038/nature08514)

7. Suzuki MM, Bird A. 2008 DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476. (doi:10.1038/nrg2341)

8. Ziller MJ, Gu H, Müller F, Donaghey J. 2013 Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477–481. (doi:10.1038/nature12433)

9. Smith ZD, Meissner A. 2013 DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220. (doi:10.1038/nrg3354)

10. Xie W, Schultz MD, Lister R, Hou Z. 2013 Epigenomic analysis of multi-lineage differentiation of human embryonic stem cells. *Cell* **153**, 1134–1148. (doi:10.1016%2Fj.cell.2013.04.022)

11. Jones PA, Baylin SB. 2007 The epigenomics of cancer. *Cell* **128**, 683–692. (doi:10.1016/j.cell.2007.01.029)

12. Zernicka-Goetz M, Morris SA, Bruce AW. 2009 Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. *Nat. Rev. Genet.* **10**, 467–477. (doi:10.1038/nrg2564)

13. Guo H, Zhu P, Yan L. 2014 The DNA methylation landscape of human early embryos. *Nature* **511**, 606–610. (doi:10.1038/nature13544)

14. Santos F, Hendrich B, Reik W, Dean W. 2002 Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241**, 172–182. (doi:10.1006/dbio.2001.0501)

15. Borgel J, Guibert S, Li Y, Chiba H. 2010 Targets and dynamics of promoter DNA methylation during early mouse development. *Nat. Genet.* **42**, 1093–1100. (doi:10.1038/ng.708)

16. Smallwood SA et al. 2011 Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat. Genet.* **43**, 811–814. (doi:10.1038/ng.864)

17. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. 2012 A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344. (doi:10.1038/nature10960)

18. Zhu P et al. 2018 Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat. Genet.* **50**, 12–19. (doi:10.1038/s41588-017-0007-6)

19. Stelzer Y, Shivalila CS, Soldner F, Markoulaki S, Jaenisch R. 2015 Tracing dynamic changes of DNA methylation at single-cell resolution. *Cell* **163**, 218–229. (doi:10.1016/j.cell.2015.08.046)

20. Wang L, Zhang J, Duan JL, Gao XX, Zhu W. 2014 Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**, 979–991. (doi:10.1016/j.cell.2014.04.017)

21. Fan G et al. 2015 The transcriptome and DNA methylome landscapes of human primordial germ cells. *Cell* **161**, 1437–1452. (doi:10.1016/j.cell.2015.05.015).

22. Smith, ZD, Chan, MM, Humm, KC. 2014 DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615. (doi:10.1038/nature13581)

23. Kitamura E et al. 2007 Analysis of tissue-specific differentially methylated regions (TDMs) in humans. *Genomics* **89**, 326–337. (doi:10.1016/j.ygeno.2006.11.006)

24. Illingworth R et al. 2008 A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **6**, e22. (doi:10.1371/journal.pbio.0060022)

25. Jones PA. 1999 The DNA methylation paradox. *Trends Genet.* **15**, 34–37. (doi:10.1016/S0168-9525(98)01636-9)

26. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schübeler D. 2005 Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853–862. (doi:10.1038/ng1598)

27. Jones PA.2012 Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492. (doi:10.1038/nrg3230)

28. Yan L et al. 2013 Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139. (doi:10.1038/nsmb.2660)

29. Levy S, Todd SC, Maecker HT. 1998 CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. *Annu. Rev. Immunol.* **16**, 89–109. (doi:10.1146/annurev.immunol.16.1.89)

30. Cedar H, Bergman Y. 2012 Programming of DNA methylation patterns. *Annu. Rev. Biochem.* **81**, 97–117. (doi:10.1146/annurev-biochem-052610-091920)

31. Jaenisch R, Bird A. 2003 Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33**, 245–254. (doi:10.1038/ng1089)

32. Reik W, Dean W, Walter J. 2001 Epigenetic reprogramming in mammalian development. *Science* **293**, 1089–1093. (doi:10.1126/science.1063443)

33. Zhang X et al. 2006 Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* **126**, 1189–1201. (doi:10.1016/j.cell.2006.08.003)

34. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M, Schübeler D. 2007 Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457–466. (doi:10.1038/ng1990)

35. Maor GL, Yearim A, Ast G. 2015 The alternative role of DNA methylation in splicing regulation. *Trends Genet.* **31**, 274–280. (doi:10.1016/j.tig.2015.03.002)

36. Yegnasubramanian S, Wu Z, Haffner MC, Esopi D. 2011 Chromosome-wide mapping of DNA methylation patterns in normal and malignant prostate cells reveals pervasive methylation of gene-associated and conserved intergenic sequences. *BMC Genomics.* **12**, 313. (doi:10.1186/1471-2164-12-313)

37. Malumbres M, de Castro IP, Santos J, Piqueras JF, Pellicer A. 1999 Hypermethylation of the cell cycle inhibitor p15INK4b 3'-untranslated region interferes with its transcriptional regulation in primary lymphomas. *Oncogene* **18**, 385–396. (doi:10.1038/sj.onc.1202299)

38. Maussion G, Yang J, Suderman M, Diallo A, Nagy C, Arnovitz M, Mechawar N, Turecki G. 2014 Functional DNA methylation in a transcript specific 3'UTR region of TrkB associates with suicide. *Epigenetics* **9**, 1061–1070. (doi:10.4161/epi.29068)

39. Maussion G, Yang J, Yerko V, Barker P, Mechawar N, Ernst C, Turecki G. 2012 Regulation of a truncated form of tropomyosin-related kinase B (TrkB) by Hsa-miR-185* in frontal cortex of suicide completers. *PLoS ONE* **7**, e39301. (doi:10.1371/journal.pone.0039301)

40. Smith JF et al. 2007 Identification of DNA methylation in 3' genomic regions that are associated with upregulation of gene expression in colorectal cancer. *Epigenetics* **2**, 161–172. (doi:10.4161/epi.2.3.4805)

41. Pheiffer C, Erasmus RT, Kengne AP, Matsha TE. 2016 Differential DNA methylation of microRNAs within promoters, intergenic and intragenic regions of type 2 diabetic, pre-diabetic and non-diabetic individuals. *Clin. Biochem.* **49**, 433–438. (doi:10.1016/j.clinbiochem.2015.11.021)

42. Baubec T, Schübeler D. 2014 Genomic patterns and context specific interpretation of DNA methylation. *Curr. Opin Genet. Dev.* **25**, 85–92. (doi:10.1016/j.gde.2013.11.015)

43. Brenet F, Moh M, Funk P, Feierstein E, Viale AJ, Socci ND, Scandura JM. 2011 DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS ONE* **6**, e14524. (doi:10.1371/journal.pone.0014524)

44. Elango N, Yi SV. 2008 DNA methylation and structural and functional bimodality of vertebrate promoters. *Mol. Biol. Evol.* **25**, 1602–1608. (doi:10.1093/molbev/msn110)

45. Maunakea AK et al. 2010 Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**, 253–257. (doi:10.1038/nature09165)

46. Wu J et al. 2016. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature* **534**, 652–657. (doi:10.1038/nature18606)

47. Lu F, Liu Y, Inoue A, Suzuki T, Zhao K, Zhang Y. 2016. Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell* **165**, 1375–1388. (doi:10.1016/j.cell.2016.05.050)

48. Wu J et al. 2018. Chromatin analysis in human early development reveals epigenetic transition during ZGA. *Nature* **557**, 256–260. (doi:10.1038/s41586-018-0080-8)

49. Gao L et al. 2018. Chromatin accessibility landscape in human early embryos and its association with evolution. *Cell* **173**, 248–259. (doi:10.1016/j.cell.2018.02.028)

50. Huang DW, Sherman BT, Lempicki RA. 2009 Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57. (doi:10.1038/nprot.2008.211)