Journal of Genetics and Genomics
Manuscript Draft

Manuscript Number: JGG-D-16-00396R1

Title: Genome-wide abnormal DNA methylome in human blastocyst

Article Type: Research paper

Keywords: abnormal; DNA methylome; Human.

Manuscript Region of Origin:

Abstract: Proper reprogramming of parental DNA methylomes is essential for mammalian embryonic development. However, it is unknown whether abnormal methylome reprogramming occurs and associates with the failure of embryonic development. Here we analyzed the DNA methylomes of 57 blastocysts and 29 trophectoderm samples with different morphological grades during assisted reproductive technology practices (ART). Our data reveals that the global methylation levels of high-quality blastocysts are similar (0.30±0.02), while the methylation levels of low-quality blastocysts are divergent and away from those of high-quality blastocysts. The proportion of blastocysts with a methylation level falling within the range of 0.30±0.02 in different grades correlates with the live birth rate for that grade. Moreover, abnormal methylated regions associate with the failure of embryonic development. Furthermore, we can use the methylation data of cells biopsied from trophectoderm to predict the blastocyst methylation level as well as to detect the aneuploidy of the blastocysts. Our data suggest DNA methylome may be a potential biomarker in blastocyst selection in ART.

***Detailed Response to Reviewers**

Reviewer #1*:*

In this manuscript, Li et al have analyzed the whole genome DNA methylation of 57 blastocysts and 29 trophectoderm samples in different morphological grades during assisted reproductive technology practices (ART). They found that the global methylation levels of high-quality blastocytes were within a narrow range around 30%, while the global methylation levels of low-quality blastocysts went more divergent. They also showed that the methylome of trophectoderm cells can predict the blastocyst methylation level and detected the aneuploidy of the blastocysts by using the DNA methylation data. This study provides a comprehensive view of DNA methylation status of the low-quality blastocysts comparing with the morphologically high-quality blastocytes. Several aspects should be improved before it can be published on JDD.

Answer: We thank this reviewer for the positive remarks regarding to the findings that we reported in this study.

Major:

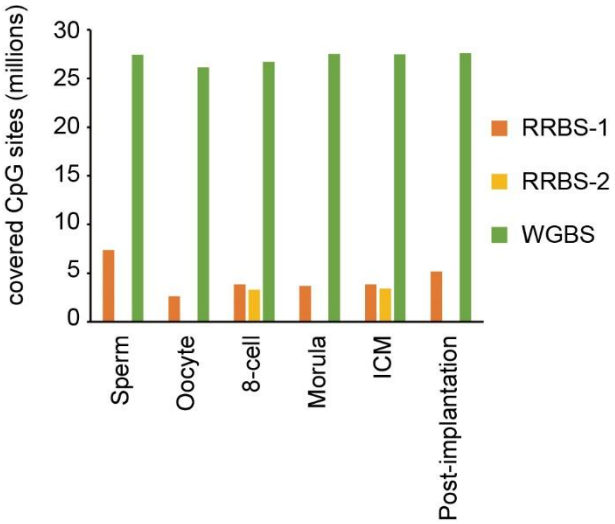
1. *The authors have adapted the conventional WGBS to low amount of input DNA material. They should provide more information about the performance of the method.*

For each sample, how many raw sequences have been obtained? How about the mapping efficiency? How many CpGs are covered in different depth?

Answer: We have included the raw reads number and mapping efficiency information in Supplemental Table 1. We also compared CpGs coverage and methylation level in different depth in Supplemental Table 2.

How about the distribution of these CpGs in different genomic region? Whether the method have no bias toward the CpG-enriched region, as expected and to be different from the PBAT- and RRBS-based methods.

Answer: Theoretically, there is no bias in our method as we did not do any enrichment. The procedure of our method followed the traditional whole genome bisulfite sequencing except that we omitted several steps for purification. We have compared our method with RRBS in another unpublished work, which indicated that our method can recover much more CpGs than RRBS (see below).



In addition, if the authors have tested the low detection limit of the method "as low as single-cell ", they should show the relevant data. If not, it should be careful to claim this. Nevertheless, the method is good if it can work for low amount DNA.

Answer: Sorry for the mistake. It should be "as low as 5-cells" as the cell numbers biopsied from blastocyst in Fig. 3 were around 5 cells.

2. The authors found that the methylation levels of low-quality blastocysts are "divergent". However, the description "divergent" is relatively unclear, if there is no special mechanism for it. Can I understand that the low-quality blastocysts can be divided into two groups as hypermethylated and hypomethylated? And the two groups may be caused by different mechanisms ? If so, the authors should analyze the hypermethylated and hypomethylated blastocysts separately.

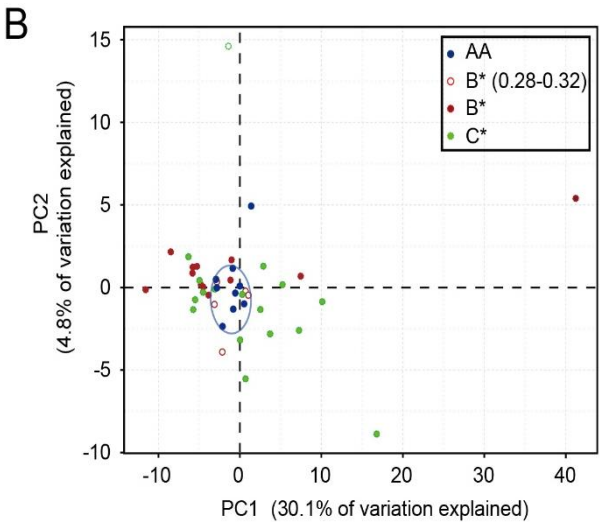
Answer: We described the methylation levels of low-quality blastocysts as "divergent" mainly because they showed a broader range of methylation level than in high-quality blastocysts. We think the word “divergent” is reasonable for the description.
For the mechanisms, we hypothesis that the divergent methylation status may be caused by insufficient or excessive demethylation as DNA methylome undergoes globally demethylation during early embryonic development. We will dissect the mechanism in following study.

3. In Figure1B, S1C and S1D, statistical analysis should be performed to examine which genomic regions showed significant differences between the AA and CC blastocysts.

Answer: Thanks for the suggestion. We included statistical analysis in the revised Fig. 1C, S1C and S1D.

4. The authors should perform clustering or other analysis to examine whether the blastocysts of different morphological grades can be separated based on the DNA methylation.

Answer: Thanks for the suggestion. We performed the PCA analysis (Fig. 1B), which showed that the AA blastocysts were closely clustered together (blue circle), indicating variations among AA blastocysts are small. For middle grade blastocysts, most of the blastocysts with methylation level falling within the range of 0.30±0.02 are similar to AA blastocysts (within blue circle), but not for other middle-grade blastocysts. As for the low-quality blastocysts, the majority of them distributed outside of the blue circle, and the variations of them were much larger.



5. Figure 2 and S2. As mentioned above, it seems to be better to separately compare the hypermethylated and hypomethylated CC blastocysts with the AA blastocysts for DMR identification. Also, the DMRs themselves should be divided into hypermethylated or hypomethylated groups for GO analysis. What are the lengths of the DMRs? Further, I suggested that the authors give a heatmap to show the methylation levels of these DMRs among these blastocysts.

Answer: Thanks for the suggestions. We have divided the DMRs into hyper and hypomethylated groups for GO analysis (Figure 2A). We also showed the length distribution of these DMRs in Figure S2A.

6. Figure 2B and C. How many DMRs (hypermethylated or hypomethylated) overlap the enhancer region and how about the GO term of the associated genes? Besides showing an example, the authors should give a global view of the interesting relationship between DMRs and enhancers.

Answer: Thanks for the suggestions. We analyzed the overlap between enhancer and DMRs (Figure S2D). There were 944 enhancers overlap with identified DMRs. The GO term of these enhancers was in Figure S2E.

7. The English writing of the manuscript should be improved.

Answer: We have carefully revised the writing of the manuscript.

Minor:

1. *Figure S1C, What is LADS? Lamina-associated domains?*

Answer: Yes, it is lamina-associated domains. We added the description in the revised legend.

2. *Figure 1B, Please add the corresponding identifier of the blastocysts to the different color lines.*

Answer: We have added the ID of these balstocysts in the figure legend.

3. *Figure 2C, what is the relationship between the blastocysts tested for RNA expression and the blastocysts tested for DNA methylation. Are they the same blastocysts or no*

Answer: They were different blastocysts.

Reviewer #2:

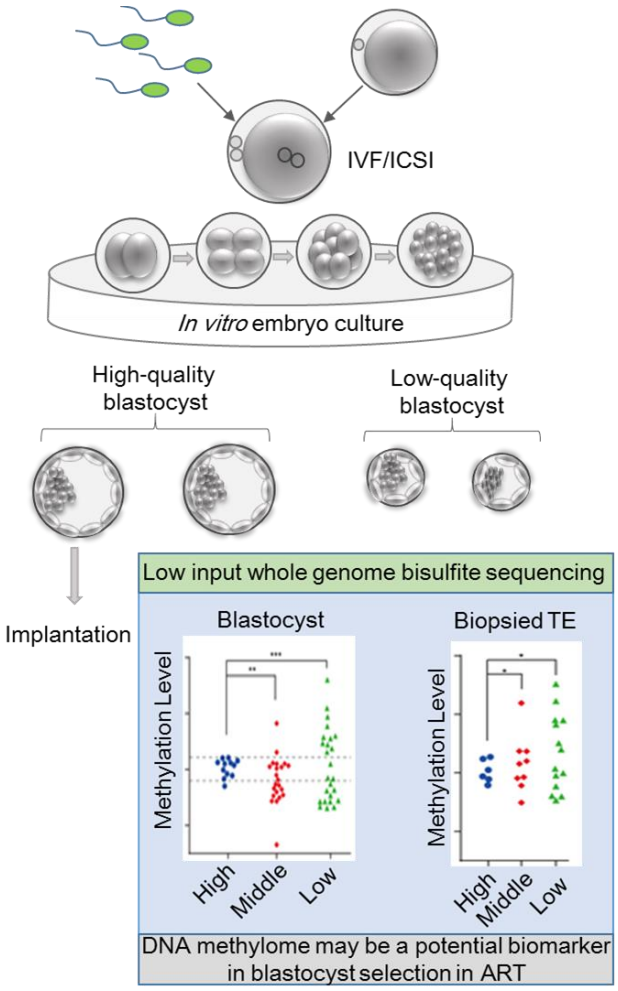
Dr. Li and colleagues provided a large genome-wide epigenetic profiling research to the cells in the different stage of embryonic development. The study was performed rigorously and the findings sound quite interesting. I recommend to publish the manuscript if the authors could make the following revision.

Answer: We thank this reviewer for the positive remarks regarding to the technical process and findings that we reported in this study.

Major Compulsory Revisions

1, A schematic overview or graphical abstract of the study was encouraged to be provided as the Figure 1, including the sample, analysis pipeline and main idea and found of the study.

Answer: Thanks for the suggestion. We have included the graphical abstract in revised Figure 4.



2, In the abstract, the value of 0.30±0.02 is not clear, what's that? Beta/M and SD/IQR/95%CI?

Answer: Sorry for the unclear writing. The value should be mean±SD.

3, "For PCA analysis, the methylomes with the genomic coverage higher than 15% were included". It is not clear what does the authors want to indicate.

Answer: In the revised version, we added PCA analysis. We only used the methylomes with the genomic coverage more than 15%, as lower coverage cannot be clustered accurately.

4, "Sequencing reads were trimmed to remove the reads containing adapters and low Quality", it is not clear how the author to do it.

Answer: Sorry for the unclearance. Data trimming was done by software Trimmomatic (Bolger et al., 2014) with the default parameter, which could remove Illuminar sequencing adaptor and reads quality lower than QC30. We included the detailed information in the revised manuscript as below: “Sequencing reads were trimmed with default setting of Trimmomatic software to remove the reads containing adapters and low-quality”.

5, "The chromosome number was deducted by R package HMMcopy". Deducted should be detected? It is not clear which dataset were applied for this analysis, WGBS or RNA-seq?

Answer: Thanks for the reminder. It should be “deduced” but not “deducted”. WGBS data were used for this analysis.

6, in "Global abnormality of DNA methylome in human blastocyst", whether the average methylation level is from same genomic positions?

Answer: As mentioned in the manuscript, “Average methylation level in each stage was the mean of methylation levels of all CpG sites.” We did not use the common shared CpGs across all different samples in this paper.

7, Why use three different notes, AB, BA, or BB, to indicate middle grades?

Answer: In IVF practice, blastocyst morphology is important to describe blastocyst quality and predict the implantation rate. Gardner morphological blastocyst grading system has been very useful in classifying the degree of blastocyst expansion as well as the morphological appearance of the inner cell mass (ICM) and the trophectoderm (TE) cells. The ICM score is listed second as, A. many cells forming a cohesive epithelium, B. few cells forming a loose epithelium and C. very few large cells. The final score is the TE score (A. tightly packed, many cells, B. loosely grouped, several cells and C. very few cells). A more detailed description is here: http://www.advancedfertility.com/blastocystimages.htm
In our practices, we first choose blastocysts with “AA” score, which will give highest success rate (Figure S3B). We also choose blastocysts with “AB, BA or BB” score to

implant, while the success rate becomes lower (Figure S3B). Therefore, we assign these blastocysts as “middle grades” blastocysts.

8, For figure 1A and 1B (Supplementary Figure S1), It would be great if the author could add the boxplot outside the dots so that the quantile of the data distribution could be provided.

Answer: Thanks for the suggestion. We have included the quantile information in revised Figure 1A, 1C, S1C, and S1D.

9, For figure 1B, more details should be provided, such as CpG shore, CpG shelf, main histone main TF-binding regions from ENCODE Project.

Answer: We had the CpG shore and TF-binding regions data in our Supplemental Figure 1C. As histone modifications are much dynamic and tissue specific, we did not find the correlated datasets from ENCODE for human blastocyst.

10, For supplementary Figure 1B, why only the Figures of AA and CC were shown?

Answer: These were representative figures for different grades based on Gardner morphological blastocyst grading system. As Gardner grades is commonly used in ART field and there are lots of representative pictures about this (i.e. http://www.advancedfertility.com/blastocystimages.htm), so we just put two most representative grades.

11, In Supplementary Figure 1D, what does it mean for "DNA" (the first column)?

Answer: It means DNA repeat elements in RepeatMasker.

12, the definition or the database of the above genomic regions was not defined or provided. Please provided the database (include version) or definition.

Answer: All the genomic datasets except enhancer were downloaded from UCSC table browser. Enhancer dataset was obtained from the putative enhancers in human embryonic stem cell (hESC) (Xie et al., 2013). All the genomic regions were under human hg19 reference. We have revised these information in the manuscript.

13, "we regarded the methylome of AA blastocysts as good epigenomic status and used the average methylation level of AA blastocysts (0.30±0.02) as the control. Interestingly, the proportion of blastocysts with a methylation level falling within the range of 0.30±0.02 in different grades is correlated with the live birth rate for that grade". Not clear what's 0.02 represent.

Answer: Sorry for the unclearance. It means standard deviation.

14, Supplementary Figure 1C and 1D should have corresponding supplementary tables to show the exact values. Note: all the supplementary tables could be place in one excel in different sheets.

Answer: We have included all the values for these figures in Supplementary Table 3.

15, Figure S2B is not clear. What does the lines indicate? Different samples? it seems the number of the lines less than the sample size in the present study, right? Or else it is a schematic diagram？ Then why not provided a classic example with the current data in a classic DMR (genes)?

Answer: Sorry for the unclearance. The lines indicate different samples used for DMR analysis. We included the sample information in the legend for Figure S2B.

16, "Thus, hypothetically, the DNA methylome of a few TE cells removed from blastocysts could predict the methylome level of AA blastocysts (0.30)". What's this 0.30 represents?

Answer: As mentioned in the manuscript, it represents the global methylation level of high-grade blastocysts.

17, Why only parts of the sample have "BSCR"? What's the difference between bisulfite conversion rate calculated by spike-in unmethylated lambda DNA and non-CpG methylation? Or measured by methylation level of chrM?

Answer: We are not sure what the reviewer mean here for “only parts of the sample have “BSCR”. Actually, we have included “BSCR” for all the samples in Supplemental Table 1. Bisulfite conversion rate calculated by spike-in unmethylated lambda DNA was commonly used in the field. Non-CpG cannot be used in our case as non-CpG can also be methylated in blastocyst stage. Although mitochondrial DNA have been used to calculate BSCR in some cases, the accuracy is still under debate as conflicting data regarding the possible presence of methylated cytosine within mitochondrial DNA (mtDNA) have been reported (Bellizzi et al., 2013). Thus, in our paper, we used spike-in unmethylated lambda DNA to calculate the BSCR.

18, "Notably, the P value of the homogeneity of variance between high-grade and low-grade TEs is 0.03, and the P value between high-grade and middle-grade is 0.05". Which matric/measurement does the author use to represent "homogeneity"?
Answer: We performed Bartlett's test by R function bartlett.test, which is used to test whether samples have equal variances. Equal variances across samples is called homogeneity of variances.

19, The following corresponding supplementary Tables should be provided for DMR, Figure 1A, 1B (value and variation), Figure 2A (genomic regions). In additional, the author

forget to mention which databased was chose to do the enrichment analysis.

Answer: We have provided these tables in revised manuscript. All the genomic datasets except enhancer were downloaded from UCSC table browser. Enhancer dataset was obtained from the putative enhancers in human embryonic stem cell (hESC) (Xie et al., 2013). All the genomic regions were under human hg19 reference. We have revised these information in the manuscript.

References:

Bellizzi, D., D'Aquila, P., Scafone, T., Giordano, M., Riso, V., Riccio, A., and Passarino, G. (2013). The Control Region of Mitochondrial DNA Shows an Unusual CpG and Non-CpG Methylation Pattern. DNA Research *20*, 537-547.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D.*, et al.* (2013). Epigenomic Analysis of Multi-lineage Differentiation of Human Embryonic Stem Cells. Cell *153*, 1134-1148.

Genome-wide abnormal DNA methylome in human blastocyst

Guoqiang Li[1,2,#], Yang Yu[1,#], Yong Fan[3#], Congru Li[2,4#], Xiaocui Xu[2,4], Jialei Duan[2], Rong Li[1], Xiangjin Kang[3], Xin Ma[2,4], Xuepeng Chen[2,4], Yuwen Ke[2], Jie Yan[1], Ying Lian[1], Ping Liu[1], Yue Zhao[1], Hongcui Zhao[1], Yaoyong Chen[3], Xiaofang Sun[3], Jianqiao Liu[3], Jie Qiao[1*], Jiang Liu[2,4*]

1 Ministry of Education Key Laboratory of Assisted Reproduction, and Beijing Key Laboratory of Reproductive Endocrinology and Assisted Reproductive Technology, Center of Reproductive Medicine, Peking University Third Hospital, Beijing, 100191, China
2 CAS Key Laboratory of Genome Sciences and Information, Collaborative Innovation Center of Genetics and Development, Beijing Institute of Genomics, CAS, Beijing 100101, China
3 Key Laboratory for Major Obstetric Diseases of Guangdong Province，The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, 510150, China
4 University of Chinese Academy of Sciences, Beijing, 100029, China
# These authors contributed equally to this work
*Correspondence: jie.qiao@263.net (J.Q.), liuj@big.ac.cn (J.L.)

Abstract

Proper reprogramming of parental DNA methylomes is essential for mammalian embryonic development. However, it is unknown whether abnormal methylome reprogramming occurs and associates with the failure of embryonic development. Here we analyzed the DNA methylomes of 57 blastocysts and 29 trophectoderm samples with different morphological grades during assisted reproductive technology practices (ART). Our data reveals that the global methylation levels of high-quality blastocysts are similar (0.30±0.02), while the methylation levels of low-quality blastocysts are divergent and away from those of high-quality blastocysts. The proportion of blastocysts with a methylation level falling within the range of 0.30±0.02 in different grades correlates with the live birth rate for that grade. Moreover, abnormal methylated regions associate with the failure of embryonic development. Furthermore, we can use the methylation data of cells biopsied from trophectoderm to predict the blastocyst methylation level as well as to detect the aneuploidy of the blastocysts. Our data suggest DNA methylome may be a potential biomarker in blastocyst selection in ART.

Introduction

Epigenetic information plays critical roles during animal development (Baubec et al., 2015; Lee et al., 2015; Pujadas and Feinberg, 2012; Schubeler, 2015). The plasticity of epigenome enables cell differentiation, organogenesis, and animal development. On the other hand, proper epigenomic pattern at certain developmental stage is also required to ensure the totipotency of early embryos in animals(Baubec et al., 2015). Genome-wide DNA demethylation occurs during early embryogenesis in both human (Guo et al., 2014; Smith et al., 2014) and mouse (Smith et al., 2012; Wang et al., 2014). The manually disturbing the DNA methylome reprogramming by genetic knock-out of DNA methyltransferases (DNMTs) or Tet3 in mouse results in a failure of embryonic development (Gu et al., 2011; Li et al., 1992; Okano et al., 1999), indicating the importance of DNA methylome in mammalian development.

At present, approximately 12% of women of childbearing age in the United States have used assisted reproductive technology (ART) (Centers for Disease Control and Prevention, 2014; Women's and Health, 2012). During natural pregnancy, the fertilized human eggs often abnormally develop, resulting in the miscarriage (Evers, 2002; Macklon et al., 2002). Likewise, during ART practice, significant proportions of human early embryos defectively develop and fail in producing the birth alive (Heitmann et al., 2013; Hill et al., 2013). Genetic instability is speculated to be part of the reason, and indeed many kinds of mutations and chromosome errors are found in a proportion of these defective human embryos(Handyside et al., 1992; Wells and Delhanty, 2001). The frequency of aneuploidy in blastocysts significantly increased for women older than 35 (Franasiak et al., 2014a; Franasiak et al., 2014b). To avoid the genetic defects in early human embryos, preimplantation genetic diagnosis (PGD) has been widely used through the removal of a few cells from an embryo by biopsy and subsequent genetic analysis (Handyside et al., 1992; Wells and Delhanty, 2001). PGD for patients with good prognosis have resulted in the increased implantation and delivery rates (Forman et al., 2013; Scott Jr et al., 2013). Nevertheless, the majority of human embryos fail in producing birth alive for no known reasons. Although DNA methylation plays important roles during embryogenesis, it is unknown whether

abnormal methylome reprogramming occurs during human embryonic development.

Material and Methods

Human samples

The human tissue collection and study procedure in this study were approved by the Institutional Review Board at Peking University Third Hospital (Research license 2012SZ015). The methods closely followed the guidelines legislated and posted by the Ministry of Health of the People's Republic of China. The patients were informed of all details of the procedure, including sample utility and research destination. Patients voluntarily signed an informed consent document. Human embryos at the blastocyst stages were donated by the couples who had conceived at least one healthy baby by assisted reproductive technology (ART) treatment. These donor couples, whose infertility is purely due to female tubal factors, had a healthy baby through the IVF cycle already. They then donated the surplus frozen embryos for research with the written informed consents signed by them. Embryos were then graded according to Gardner morphological blastocyst grading system before collected for further methylation study.

The human embryos prepared for blastocyst biopsy were treated firstly at early blastocyst stage. One hole was made in the zona pellucida of these embryos opposite to ICM using laser pulse (ZILOS-tk, Hamilton Thorne Biosciences, Inc., Beverly, MA, USA). On day 6, the fully developed blastocysts were grouped and transferred into biopsy buffer medium (G-MOPES PLUS, Vitrolife). The ICM was fixed at the 9 o'clock position by the holding pipette, and the hached trophectoderm (TE) cells were stretched from the hole by a biopsy pipette with a 25 μm diameter. A few cells were removed from the blastocysts using laser pulse. The biopsied TE cells and the surplus blastocyst were then processed to methylome analysis.

Whole Genome Bisulfite Sequencing

Rather than mapping the methylome by pooling multiple blastocysts together, we investigated the methylome of each blastocyst individually. DNA Methylation

libraries of human single blastocyst were constructed with our modified library generation method termed as “One Tube” method. Briefly, single blastocyst was lysed and then fragmented by sonication. The fragmented DNA was end-repaired, dA-tailed and ligated to cytosine methylated Illumina Truseq adapter. Bisulfite conversion reaction was performed directly on the ligation mix with 0.5% unmethylated lambda DNA spiked-in. PCR amplified library was purified and sequenced on Hiseq2000 or Hiseq3000 platform (Illumina).

mRNA Sequencing

Blastocyst RNA was amplified and reverse-transcripted with REPLI-g WTA Single Cell Kit (Qiagen). RNA-seq Libraries were constructed with NEBNext® Ultra™ RNA Library Prep Kit for illumine (NEB) according to manufacture instruction. QC-passed libraries were then sequenced on Hiseq2500 platform with pair-end module. Methyl-seq and mRNA-seq data have been deposited in the GEO data repository under accession number GSE65736.

Sequencing Data Processing

Sequencing reads were trimmed to remove the reads containing adapters and low quality. Sequencing reads were trimmed with default setting of Trimmomatic software (Bolger et al., 2014) to remove the adapters and low-quality reads. Trimmed reads were aligned to human reference hg19 by using Bismark (v12.5) (Krueger and Andrews, 2011). PCR duplications were removed with Picard (http://broadinstitute.github.io/picard/) and the overlapped regions in uniquely mapped paired reads were clipped with clipOverlap function of BamUtil (http://genome.sph.umich.edu/wiki/BamUtil: clipOverlap). CpG and non-CpG methylation level were extracted with mpileup function of Samtools (v0.1.19) (Li et al., 2009). Strands were merged to calculate the CpG methylation level per site. Average methylation level in each stage was the mean of methylation levels of each site. Principal Component Analysis (PCA) is often used to emphasize grouping structure in the data. We performed PCA analysis on the samples’%-methylation

profiles by R package methylKit (Akalin et al., 2012). For PCA analysis, the methylomes with the genomic coverage higher than 15% were included. For DMR analyses in Fig. 2 and Fig. S2, we used the R package bsseq, which is a smoothing local likelihood method that shows precise results even with low coverage data as well as have the ability to handle biological replicates (Hansen et al., 2012). DMRs contain at least 5 CpGs and the difference level between two groups higher than 0.2 were used for further analyses.

For mRNA-seq data, adapter-containing and low quality reads were trimmed and then aligned with TopHat (Trapnell et al., 2009). The unique reads were used to calculate the Fragments Per Kilobase of transcript per million mapped reads (FPKM) with Cufflinks v2.0.2 (http://cufflinks.cbcb.umd.edu). Differential gene expression between high- and low-quality blastocysts was calculated with DESeq using the default parameters (Anders and Huber, 2010).

The chromosome number was deducted by R package HMMcopy (Ha et al., 2012), which make copy number estimations based on read depth of the whole genome data in fixed interval with additional GC and mappability correction. The copy numbers were then segmented and classified with a robust Hidden Markov Model. Here we divided the genome into non-overlapping windows of 1Mb, and assigned the median autosomal read count corresponds to copy number 2. Chromosomal gain (copy number >2) and loss (copy number <2) were seen as horizontal green bars above and below, respectively, the copy number state of 2. Embryos were diagnosed as normal or euploid if the generated plot showed no gain or loss.

Results

Global abnormality of DNA methylome in human blastocyst

Since DNA methylome reprogramming is highly associated with embryonic development in animals (Jiang et al., 2013; Wang et al., 2014), the precise reprogramming is, theoretically, important in determining embryonic condition. To investigate whether the abnormal methylome reprogramming occurs in human embryos, we analyzed the methylomes of blastocyst of various grades based on Gardner morphological blastocyst grading system including the ICM grade and the trophectoderm (TE) grade (Hardarson et al., 2012). DNA methylomes at base-resolution were examined by using as low as single blastocyst (**Fig. S1A**). We analyzed the methylomes of total 57 blastocysts with different morphological grades including high morphological grade (AA) blastocysts, middle grades (AB, BA, or BB) blastocysts and low grades (CC, BC or CB) blastocysts individually (**Table S1**). Two representative images of high grade (AA) and low grade blastocysts were presented in Fig. S1B**.** We calculated average methylation levels of each blastocysts in different sequencing depth and found that the methylation levels were consistent in different depth (**Table S2**), which indicated that our method could capture the accurate methylation level with low input samples. Our data reveals that the global methylation levels of 12 AA blastocysts are similar (0.30±0.02, 95% CI 0.29-0.31), ranging from 0.27 to 0.32 (**Fig. 1A** and **Table S1**). Surprisingly, the methylation levels of 12 (among 20) middle quality blastocysts do not fall within the range of 0.30±0.02 (**Fig. 1A** and **Table S1**), but are variant. Furthermore, the methylation levels of 22 low quality blastocysts (among 25) differ from AA blastocysts (**Fig. 1A**). The methylation levels of low-quality blastocysts are divergent, ranging from 0.23 to 0.46. Compared to high-quality blastocysts, low-quality blastocysts are more inconsistent in terms of methylation level (**Fig. 1A**). To bring out strong patterns in our dataset, we applied principal component analysis (PCA) to analyze the variations of DNA methylation pattern among different blastocysts. The data also showed that the AA blastocysts are closely clustered together (**Fig. 1B**, blue circle), indicating variations among AA blastocysts are small (**Fig. 1B**). For middle grade blastocysts, most of the blastocysts

with methylation level falling within the range of 0.30±0.02 are similar to AA blastocysts (**Fig. 1B**, within blue circle), but not for other middle-grade blastocysts. As for the low-quality blastocysts, the majority of them distributed outside of the blue circle, and the variations of them were much larger (**Fig. 1B**). These data demonstrate that the divergence of DNA methylome of blastocysts correlates with the morphological grades.

Additionally, the methylation state of functional elements in high-quality blastocysts is similar, while the methylation pattern of functional elements in each blastocyst of low-quality blastocysts displays a different state (**Fig. 1C, Fig. S1C, S1D, and Table S3**). Taken together, these data show that genome-wide abnormality of DNA methylome frequently occurs during human early embryonic development.

Abnormal methylated regions associate with pathways regulating embryonic development

To gain a close-up view of the differences observed above, we investigated differentially methylated regions (DMRs) between high-quality blastocysts and low-quality blastocysts (**Table S4**). The average length of DMRs are around 2kb (**Fig. S2A**). A significant proportion of DMRs are located in CpG islands (CGIs) or CGI shores (**Fig. S2B and S2C**). The functional enrichment analyses demonstrated that the genes with promoters located in DMRs are enriched in many fundamental pathways critical for embryonic development, including cell cycle, DNA metabolism, chromosome localization and DNA modification (**Fig. 2A**). Genes associated with intergenic DMRs are enriched in specific developmental categories, such as neuron development and spinal cord patterning (**Fig. 2A**).

Interestingly, our data show that DNA methylation in enhancer regions are often falsely reprogramed in low-quality embryos. Gene ontology (GO) enrichment analyses for enhancers that locate within DMRs (**Fig. S2D**) showed the enriched categories in developmental and metabolism pathways (**Fig. S2E**). Fig. 2B shows that enhancer for *IDH2* gene, a metabolism enzyme which can regulate the oxidization of 5mC(Kaelin and McKnight, 2013), is hyper-methylated in CC blastocysts (**Fig. 2B**).

In consistence, the expression of IDH2 in three CC blastocysts is much lower than that in three AA blastocysts (**Fig. 2C**). An example for a cell cycle gene, *CDK10*, is also differentially methylated in enhancers, and differentially expressed between high and low-quality blastocysts (**Fig. S2F and S2G**). Furthermore, consistent with that divergence of DNA methylomes in low-quality blastocysts (**Fig. 1B**), the global transcriptomes of CC blastocysts are also revealed divergent and different patterns from AA blastocysts (**Fig. S2H**). Taken together, our results suggest that abnormal methylome may affect the developmental potential of early embryos.

Methylome status associates with the live birth rate following ART treatment

In ART treatments, AA blastocysts are ideal for embryonic transfer, which produce 39% live birth rate in Peking University Third Hospital. However, only a minority (33%) of the fresh elective single embryonic transfers (eSet) are AA blastocysts in our hospital, while the majority of the blastocysts for embryonic transfer (52%) are middle-quality B* blastocysts (AB, BA and BB) (**Fig. S3A**). The live birth rate is approximately up to 28% for either AB/BA blastocysts or BB blastocysts (**Fig. S3B**). A small proportion of patients used the low-quality C* blastocysts (BC and CB), resulted in nearly 4% live birth rate (**Fig. S3B**). Previously, the tested CC grade blastocysts were unable to produce any live birth(Heitmann et al., 2013; Hill et al., 2013), thus they were not used for embryonic transfer in our hospital. Considering that AA blastocysts have the highest live birth rate and uniform methylome, we regarded the methylome of AA blastocysts as good epigenomic status and used the average methylation level of AA blastocysts (0.30±0.02) as the control. Interestingly, the proportion of blastocysts with a methylation level falling within the range of 0.30±0.02 in different grades is correlated with the live birth rate for that grade (**Fig. 3A**, Pearson Correlation Coefficient, r=0.93). The data suggest that DNA methylome status associates with the live birth rate in ART.

DNA methylome examination of cells biopsied from blastocyst

PGD has been used for more than 20 years to screen genetic diseases during ART

practice. The removal of cells by biopsy from the trophectoderm has been used for the diagnosis of genetic mutations or chromosomal errors before embryonic transfer to the uterus (Handyside et al., 1992; Wells and Delhanty, 2001). Therefore, we use the cells removed by biopsy from TE in blastocyst to analyze the DNA methylome. The methylomes of 29 biopsied TE samples are profiled, 26 of them have paired methylome data of the blastocysts (**Table S1**). Our data show that the methylation levels of TE from high-quality and middle-quality blastocysts are similar to the paired blastocysts (**Fig. S3C**). The methylation levels of most TEs from low-quality blastocysts are comparable to the levels of the paired blastocysts, except the levels of a few TEs are higher than the levels of the paired blastocysts, but the level are still significant different from control level (0.30±0.02) (**Table S1, Fig. S3D**). Thus, our results show that DNA methylome of a few cells from TE in blastocyst with high and middle morphological grades can indicate the methylome of entire blastocyst. Notably, the P value of the homogeneity of variance between high-grade and low-grade TEs is 0.03, and the P value between high-grade and middle-grade is 0.05. These data indicate that that the methylation status of TE is also correlated with the morphological grades of blastocysts (**Fig. 3B**), which is similar to the results observed from blastocysts (**Fig. 1A**). Therefore, as for the high-quality and middle-quality blastocysts, we are able to use the methylome of a few biopsied cells from TE to predict the DNA methylome of the blastocyst before the embryonic transfer to the uterus during ART. Since low-quality blastocysts will not be used in ART practices, it is unnecessary to examine its methylomes. Thus, hypothetically, the DNA methylome of a few TE cells removed from blastocysts could predict the methylome level of AA blastocysts (0.30).

Aneuploidy analyses from methylome data

Aneuploidy is frequently observed in human embryos. PGD has been widely used through the removal of a few cells from an embryo by biopsy and subsequent genetic analysis (Handyside et al., 1992; Wells and Delhanty, 2001). PGD for patients with good prognosis have resulted in the increased implantation and delivery rates (Forman

et al., 2013; Franasiak et al., 2014a; Franasiak et al., 2014b; Scott Jr et al., 2013). PGD for Aneuploidy Screening (PGD-AS) has been applied in ART practices by array-based comparative genomic hybridization (arrayCGH) and single-nucleotide polymorphism (SNP) array approaches (Dahdouh et al., 2015). More recently, next-generation sequencing has been introduced into IVF field(Fiorentino et al., 2014). DNA methylation data have also been used to evaluate the copy number of chromosomes (Feber et al., 2014; Oda et al., 2009). Therefore, we aimed to use our DNA methylome data to analyze the chromosome copy number variations (CNVs) in blastocysts and the biopsied TE cells. Notably, aneuploidy was present in 22 blastocysts (**Fig. 3C, Table 1**), about 30% of all examined samples, which is consistent with the ratio discovered by traditional PGD (Fragouli et al., 2011). Our data show that some blastocysts with normal methylation level still present with aneuploidy (**Fig. 3D**), indicating that aneuploidy does not affect DNA methylation. Notably, the results in paired blastocyst and biopsied TE samples are highly consistent (**Table 1**), indicating that epigenomic examination on biopsied TE cells is predictive to the entire blastocyst. In summary, we can use DNA methylome from biopsied samples to predict the DNA methylome pattern and chromosome pattern, which may be useful for TE selection in the future.

Discussion

In this study, we performed whole genome bisulfite sequencing for 57 human preimplanted blastocysts as well as 29 biopsied TE samples individually. We found that global methylation level and pattern were dramatically unstable in low-quality embryos compared to high-quality embryos (**Fig. 1**). We show that DNA methylome examination in biopsied TE can predict the entire blastocyst in both epigenetic status and chromosome aneuploidy (**Fig. 4**).

DNA methylation reprogramming is highly associated with animal embryonic development (Jiang et al., 2013; Wang et al., 2014). The manually disturbing the DNA methylome reprogramming by genetic knock-out of DNA methyltransferases (DNMTs) or Tet3 in mouse results in a failure of embryonic development(Gu et al.,

2011; Li et al., 1992; Okano et al., 1999), indicating the importance of DNA methylome in mammalian development. The global methylome abnormality frequently takes place in human blastocyst, and associate with low live birth rate. Previous studies in human diseases including cancer show that the global methylation level in pathological tissue has limited (or no) change compared to normal tissue (Baylin and Jones, 2011; Hansen et al., 2011; Shen and Laird, 2013). Considering that genetic manipulation of *DNMTs* in mouse can lead to the failure of embryonic development (Li et al., 1992; Okano et al., 1999), the global abnormal methylome may be an important factor that results in the failure of human embryonic development. Indeed, falsely reprogrammed methylation regions are highly enriched in development related pathways (**Fig. 2**), suggesting that the abnormal methylome may affect the developmental potential of embryos.

Right now, we are not sure what the mechanism behind the abnormal methylome is. Various pathways in different embryos may involve the causes of improper reprogramming. We have shown that aneuploidy does not associate with DNA methylation in blastocyst (**Fig. 3D**). The advanced maternal age is associated with the rate of aneuploidy in blastocysts (Franasiak et al., 2014a; Franasiak et al., 2014b). Probably, epigenetic status in blastocysts may also be affected by the maternal age. There is the likelihood that an improper micro-environment ultimately leads to an altered methylome in blastocysts since the environment has the ability to influence epigenetic states (Jirtle and Skinner, 2007). Frozen-thawed embryo transfer cycles may affect the stability of some proteins in early embryos, which could also have the impact on the DNA methylome patterns. It is our hope that the optimization of embryo culture conditions during in vitro fertilization will facilitate methylation reprogramming and will improve the efficiency of ART treatment soon. In the future, more studies should be done to investigate the cause of the abnormality of DNA methylomes in human early embryos.

Right now, the outcome at birth is defined as a successful pregnancy. However, the consequences of ART practice for the baby's later life remains to be considered. Precious work has suggested that the intrauterine environment is associated with

epigenetic programming of the fetal metabolism and predisposition to chronic metabolic disorders (Lehnen et al., 2013) later in life. Epimutation, especially these in imprinting disorders, is wildly observed after ART (Tee et al., 2013). Therefore, a full understanding of the cellular and molecular biology of human reproduction must include a study of epigenetics and genomic imprinting. Furthermore, epigenetic variation during ART practice should be taken into consideration before embryo transferring.

PGD has been widely used to detect genetic mutations or chromosome errors before embryonic transfer to the uterus during ART clinic practices (Handyside et al., 1992; Wells and Delhanty, 2001). We show that DNA methylome examination can also detect aneuploidy in blastocysts. Our data further demonstrates that embryos with better methylome status associate with higher live birth rate (**Fig. 3A**). Furthermore, our data suggest that epigenomic examination in blastocyst can determine the epigenomic status and euploid chromosome of the blastocyst prior to the embryonic transfer. Therefore, methylome examination in blastocyst may have advantage compared to traditional PGD method, and improve the efficiency of ART. Currently, 12% of women use ART services in the United States. In the future, proper clinic trial is needed to evaluate the value of DNA methylome examination method. We wish that the methylome examination in blastocyst might improve the efficiency of ART practices.

Acknowledgments

This work was supported by grants from CAS Strategic Priority Research Program (XDB13040000); the 973 Program of China (2014CB943203, 2015CB856200, 2011CB510101, and 2011CB944504); from National Natural Science Foundation of China to (91219104, 31425015, 31200958, 31371521, 31230047, and 81370766); Beijing Nova Program (xxjh2015011); Zhujiang Science and Technology Star Project of Guangzhou (2012J2200006). We thank the sequencing facility and High Performance Computer Platform in BIG, CAS.

Author Contributions

All authors contributed to the design of the study and critically revised the report for content. JQ and JL were responsible for the conception and overall supervision of the trial. YY and YF managed the samples gathering, with assistance from RL, XK, JY, YL, PL, YZ and HZ. GL and CL performed the experiments and data analyses, with assistant from XX, XM, JD and XP. XH, XS, and JL helped design the original protocol. YK and JQ provided technical advice and supervised the peer counseling intervention. GL, YY, and JL wrote the first draft of the report and were responsible for subsequent collation of inputs and redrafting. JQ and JL are guarantors for the report.

References

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A., and Mason, C.E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biology *13*, 1-9.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology *11*, 1-12.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schubeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243-247.

Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome - biological and translational implications. Nature reviews Cancer *11*, 726-734.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Centers for Disease Control and Prevention, A.S.f.R.M., Society for Assisted Reproductive Technology. (2014). 2012 assisted reproductive technology fertility clinic success rates report.

Dahdouh, E.M., Balayla, J., Audibert, F., Wilson, R.D., Brock, J.A., Campagnolo, C., Carroll, J., Chong, K., Gagnon, A., Johnson, J.A.*, et al.* (2015). Technical Update: Preimplantation Genetic Diagnosis and Screening. J Obstet Gynaecol Can *37*, 451-463.

Evers, J.L. (2002). Female subfertility. Lancet *360*, 151-159.

Feber, A., Guilhamon, P., Lechner, M., Fenton, T., Wilson, G.A., Thirlwell, C., Morris, T.J., Flanagan, A.M., Teschendorff, A.E., Kelly, J.D.*, et al.* (2014). Using high-density DNA methylation arrays to profile copy number alterations. Genome Biology *15*, 1-13.

Fiorentino, F., Biricik, A., Bono, S., Spizzichino, L., Cotroneo, E., Cottone, G., Kokocinski, F., and Michel, C.-E. (2014). Development and validation of a next-generation sequencing–based protocol for 24-chromosome aneuploidy screening of embryos. Fertility and sterility *101*, 1375-1382.e1372.

Forman, E.J., Hong, K.H., Ferry, K.M., Tao, X., Taylor, D., Levy, B., Treff, N.R., and Scott, R.T., Jr. (2013). In vitro fertilization with single euploid blastocyst transfer: a randomized controlled trial. Fertility and

sterility *100*, 100-107 e101.

Fragouli, E., Alfarawati, S., Daphnis, D.D., Goodall, N.-n., Mania, A., Griffiths, T., Gordon, A., and Wells, D. (2011). Cytogenetic analysis of human blastocysts with the use of FISH, CGH and aCGH: scientific data and technical evaluation. Human Reproduction *26*, 480-490.

Franasiak, J.M., Forman, E.J., Hong, K.H., Werner, M.D., Upham, K.M., Treff, N.R., and Scott, R.T. (2014a). Aneuploidy across individual chromosomes at the embryonic level in trophectoderm biopsies: changes with patient age and chromosome structure. Journal of Assisted Reproduction and Genetics *31*, 1501-1509.

Franasiak, J.M., Forman, E.J., Hong, K.H., Werner, M.D., Upham, K.M., Treff, N.R., and Scott, R.T., Jr. (2014b). The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. Fertility and sterility *101*, 656-663 e651.

Gu, T.P., Guo, F., Yang, H., Wu, H.P., Xu, G.F., Liu, W., Xie, Z.G., Shi, L., He, X., Jin, S.G.*, et al.* (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606-610.

Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J.*, et al.* (2014). The DNA methylation landscape of human early embryos. Nature *511*, 606-610.

Ha, G., Roth, A., Lai, D., Bashashati, A., Ding, J., Goya, R., Giuliany, R., Rosner, J., Oloumi, A., Shumansky, K.*, et al.* (2012). Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Research *22*, 1995-2007.

Handyside, A.H., Lesko, J.G., Tarin, J.J., Winston, R.M., and Hughes, M.R. (1992). Birth of a normal girl after in vitro fertilization and preimplantation diagnostic testing for cystic fibrosis. The New England journal of medicine *327*, 905-909.

Hansen, K.D., Langmead, B., and Irizarry, R.A. (2012). BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. Genome Biology *13*, 1-10.

Hansen, K.D., Timp, W., Bravo, H.C., Sabunciyan, S., Langmead, B., McDonald, O.G., Wen, B., Wu, H., Liu, Y., Diep, D.*, et al.* (2011). Increased methylation variation in epigenetic domains across cancer types. Nat Genet *43*, 768-775.

Hardarson, T., Van Landuyt, L., and Jones, G. (2012). The blastocyst. Human Reproduction.

Heitmann, R.J., Hill, M.J., Richter, K.S., DeCherney, A.H., and Widra, E.A. (2013). The simplified SART embryo scoring system is highly correlated to implantation and live birth in single blastocyst transfers. J Assist Reprod Genet *30*, 563-567.

Hill, M.J., Richter, K.S., Heitmann, R.J., Graham, J.R., Tucker, M.J., DeCherney, A.H., Browne, P.E., and Levens, E.D. (2013). Trophectoderm grade predicts outcomes of single-blastocyst transfers. Fertil Steril *99*, 1283-1289 e1281.

Jiang, L., Zhang, J., Wang, J.-J., Wang, L., Zhang, L., Li, G., Yang, X., Ma, X., Sun, X., Cai, J.*, et al.* (2013). Sperm, but Not Oocyte, DNA Methylome Is Inherited by Zebrafish Early Embryos. Cell *153*, 773-784.

Jirtle, R.L., and Skinner, M.K. (2007). Environmental epigenomics and disease susceptibility. Nat Rev Genet *8*, 253-262.

Kaelin, W.G., Jr., and McKnight, S.L. (2013). Influence of metabolism on epigenetics and disease. Cell *153*, 56-69.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571-1572.

Lee, H.J., Lowdon, R.F., Maricque, B., Zhang, B., Stevens, M., Li, D., Johnson, S.L., and Wang, T. (2015).

Developmental enhancers revealed by extensive DNA methylome maps of zebrafish early embryos. Nature communications *6*, 6315.

Lehnen, H., Zechner, U., and Haaf, T. (2013). Epigenetics of gestational diabetes mellitus and offspring health: the time for action is in early stages of life. Molecular Human Reproduction *19*, 415-422.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Macklon, N.S., Geraedts, J.P., and Fauser, B.C. (2002). Conception to ongoing pregnancy: the 'black box' of early pregnancy loss. Human reproduction update *8*, 333-343.

Oda, M., Glass, J.L., Thompson, R.F., Mo, Y., Olivier, E.N., Figueroa, M.E., Selzer, R.R., Richmond, T.A., Zhang, X., Dannenberg, L*., et al.* (2009). High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. Nucleic Acids Research *37*, 3829-3839.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell *99*, 247-257.

Pujadas, E., and Feinberg, A.P. (2012). Regulated noise in the epigenetic landscape of development and disease. Cell *148*, 1123-1131.

Schubeler, D. (2015). Function and information content of DNA methylation. Nature *517*, 321-326.

Scott Jr, R.T., Upham, K.M., Forman, E.J., Hong, K.H., Scott, K.L., Taylor, D., Tao, X., and Treff, N.R. (2013). Blastocyst biopsy with comprehensive chromosome screening and fresh embryo transfer significantly increases in vitro fertilization implantation and delivery rates: a randomized controlled trial. Fertility and sterility *100*, 697-703.

Shen, H., and Laird, P.W. (2013). Interplay between the cancer genome and epigenome. Cell *153*, 38-55.

Smith, Z.D., Chan, M.M., Humm, K.C., Karnik, R., Mekhoubad, S., Regev, A., Eggan, K., and Meissner, A. (2014). DNA methylation dynamics of the human preimplantation embryo. Nature *511*, 611-615.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature.

Tee, L., Lim, D.H.K., Dias, R.P., Baudement, M.-O., Slater, A.A., Kirby, G., Hancocks, T., Stewart, H., Hardy, C., Macdonald, F*., et al.* (2013). Epimutation profiling in Beckwith-Wiedemann syndrome: relationship with assisted reproductive technology. Clinical Epigenetics *5*, 23-23.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Wang, L., Zhang, J., Duan, J., Gao, X., Zhu, W., Lu, X., Yang, L., Li, G., Ci, W., Li, W*., et al.* (2014). Programming and inheritance of parental DNA methylomes in mammals. Cell *157*, 979-991.

Wells, D., and Delhanty, J.D. (2001). Preimplantation genetic diagnosis: applications for molecular medicine. Trends in molecular medicine *7*, 23-30.

Women's, N.C.C.f., and Health, C.s. (2012). Ectopic pregnancy and miscarriage: diagnosis and initial management in early pregnancy of ectopic pregnancy and miscarriage.

Figures and legends

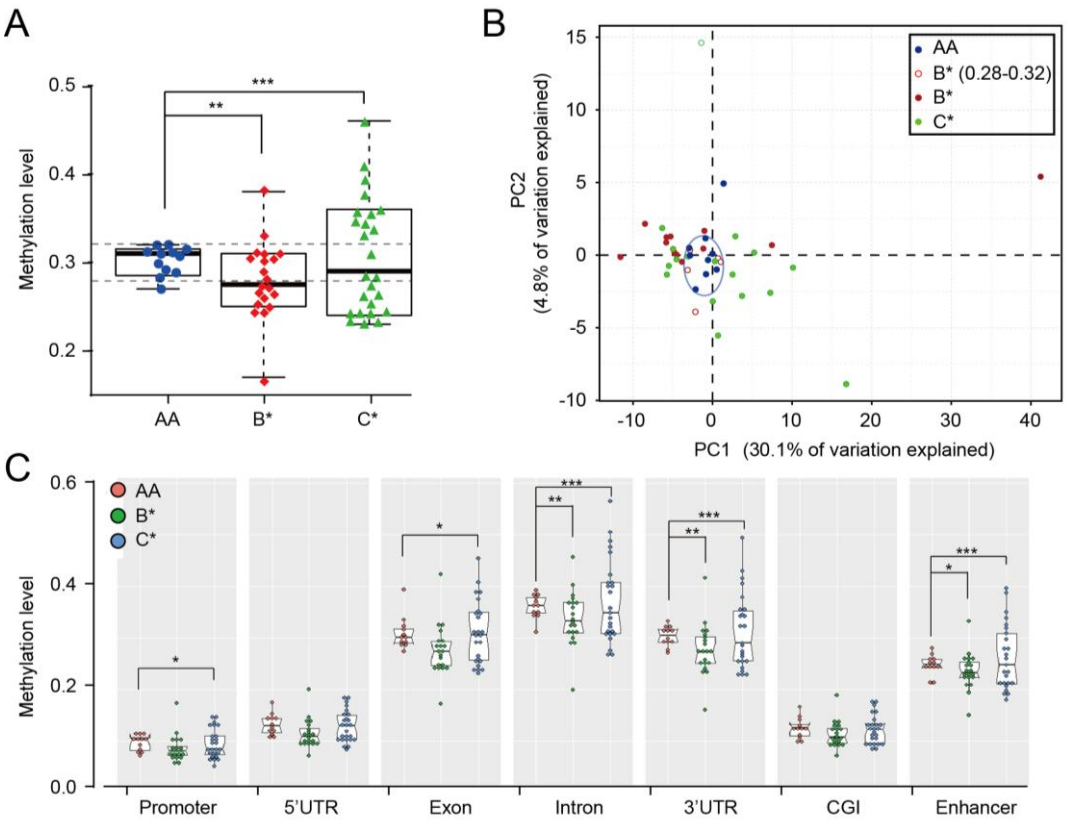


Fig. 1. Global methylome instability in human early embryos.

(**A**) The methylation level of high-quality (AA), middle-quality (B*) and low-quality (C*) blastocysts. 1-4AA and 1-2CB blastocysts were derived from the same parents, as were the 1-6AA and 2-2CC blastocysts. The P value of the homogeneity of variance between AA and B* is 0.004 (**), P value between AA and C* is 7×10^{-5} (***).

(**B**) PCA analyses according to the CpG methylome status of each blastocyst. PC1 represents 30.1% of variation and PC2 explains 4.8% of variation.

(**C**) Average methylation levels of various functional elements in different blastocysts. The P values of the homogeneity of variance between AA and B*or C*: *P<0.05; **P<0.01; ***P<0.001.

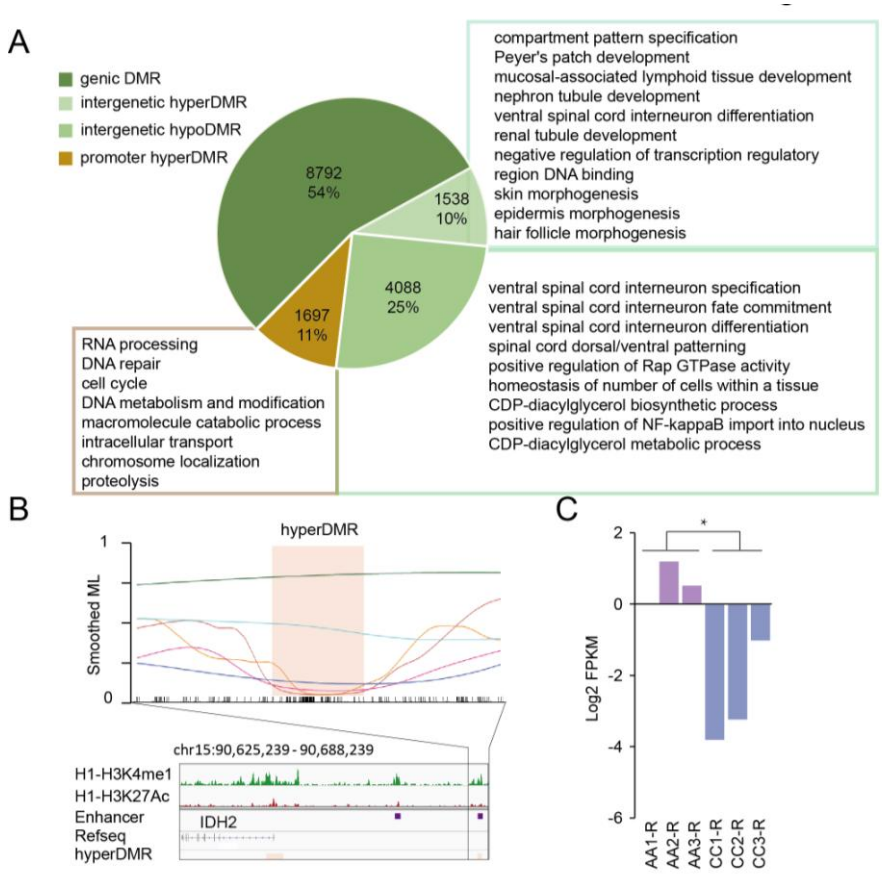


Fig. 2. Functional analyses of altered methylation reprogramming.

(**A**) Functional enrichment of differentially methylated regions between high-quality and low-quality blastocysts. Genes with promoters located in DMRs were used for DAVID GO enrichment analyses. The cis-regulatory functions of intergenic DMRs were analyzed with GREAT tools. GO biological pathways with p-value < 0.05 were considered as the significant difference.

(**B**) The enhancer near IDH2 is differentially methylated. The smoothed methylation levels of high-quality and low-quality blastocysts are showed with extension ±5kb. The smoothed ML analysis was according to previous method (Hansen et al., 2011). Red, orange and purple indicate 1-4AA, 2-4AA and 1-6AA respectively. Blue, cyan, and blue violet indicate 9-3CC, 2-2CC and 8-3CC respectively. The DMR is indicated with pink shading. Short black bars indicate the location of CpG sites.

(**C**) The relative mRNA expression level of IDH2 in AA blastocysts and CC blastocysts. Three RNA-seq replications were performed for each group. P-value=0.03 (*).

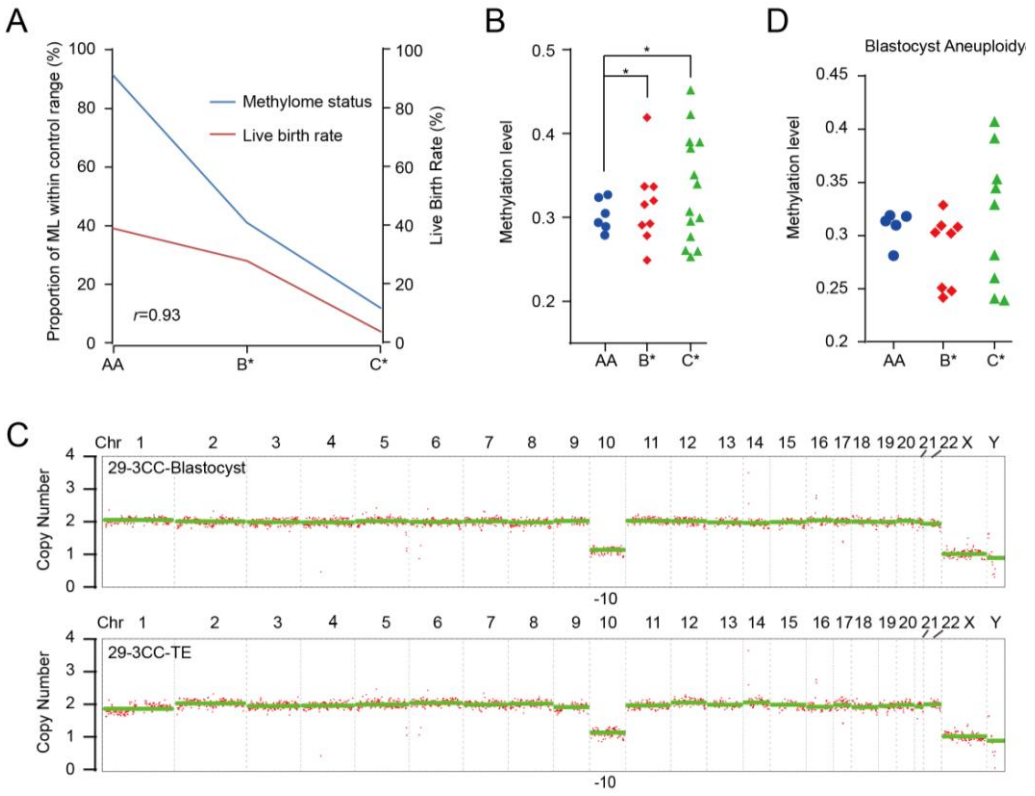


Fig. 3. DNA methylome examination in human early embryos.

(**A**) The blue line indicates the proportion of blastocysts with methylation levels falling within 0.30±0.02. The red line indicates the live birth rate of the blastocysts with different grade. The result was from the patients using fresh embryos. The live birth rate of BC and CB is used to replace the live birth rate of BC、CB and CC blastocysts. The live birth rate of BC and CB should be higher than that of BC、CB and CC altogether.

(**B**) The methylation level of TE biopsied from high-quality (AA), middle-quality (B*) and low-quality (C*) blastocysts. * indicates $p \leqslant 0.05$.

(**C**) The graphical illustration of chromosome copy number was detected from our DNA methylome data.

(**D**) The methylation level of blastocysts with aneuploidy.

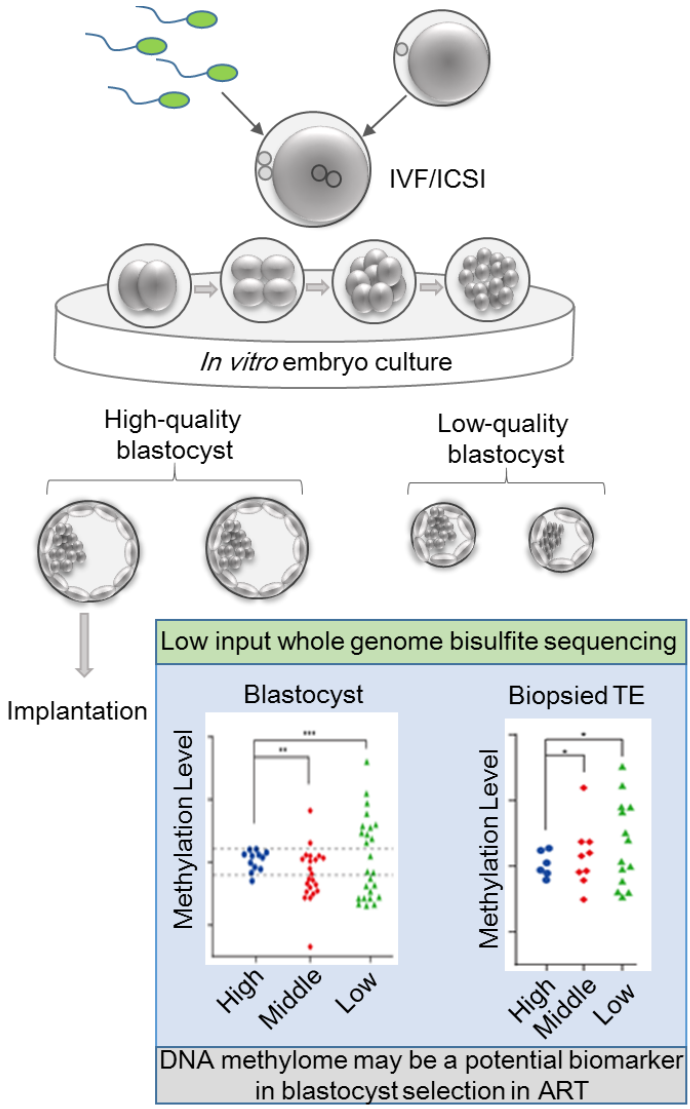


Fig. 4. DNA methylome instability and examination in human early embryos. During ART practice, significant proportions of human early embryos defectively develop and fail in producing the birth alive. We analyzed the DNA methylomes of individual blastocysts and trophectoderm samples with different morphological grades during ART. Our data reveals that the global methylation levels of high-quality blastocysts are similar, while the methylation levels of low-quality blastocysts are divergent and away from those of high-quality blastocysts. We can use the methylation data of cells biopsied from trophectoderm to predict the blastocyst methylation level as well as to detect the aneuploidy of the blastocysts. Our data suggest DNA methylome may be a potential biomarker in blastocyst selection in ART.

Table 1 Aneuploidy results for the embryos by using DNA methylome data.

	Sample	Aneuploid	Paired Sample	Aneuploid
AA	2-4AA	-19	6-5AA	-1p
	25-3AA	+5q, -9q	6-5AA-TE	-1p
			17-5AA	-22
			17-5AA-TE	-22
			19-4AA	-16
			19-4AA-TE	-16
B*	14-4BB	-13	11-5BB	-22
	54-5BB	-14	11-5BB-TE	-22
	32-6BA	+22	15-4BA	-20, -19
	13-5BB	-15, +17	15-4BA-TE	-20, -19
	31-5BB	-21, -22		
	18-5BB	-14		
C*	9-3CC	+7, X	29-3CC	-10
	2-2CC	-21	29-3CC-TE	-10
	36-3CC	+6	30-3CC	-12
	41-3CC	+16	30-3CC-TE	-12
			35-4CC	+16
			35-4CC-TE	+16
			50-4CC	XXY
			50-4CC-TE	XXY
			51-4CC	+4
			51-4CC-TE	+4

Supplemental Data

Supplement to: Guoqiang Li, Yang Yu, Yong Fan, et al. Globally abnormal DNA methylome in human blastocyst

Supplementary Figures 1-3

Supplementary Figures

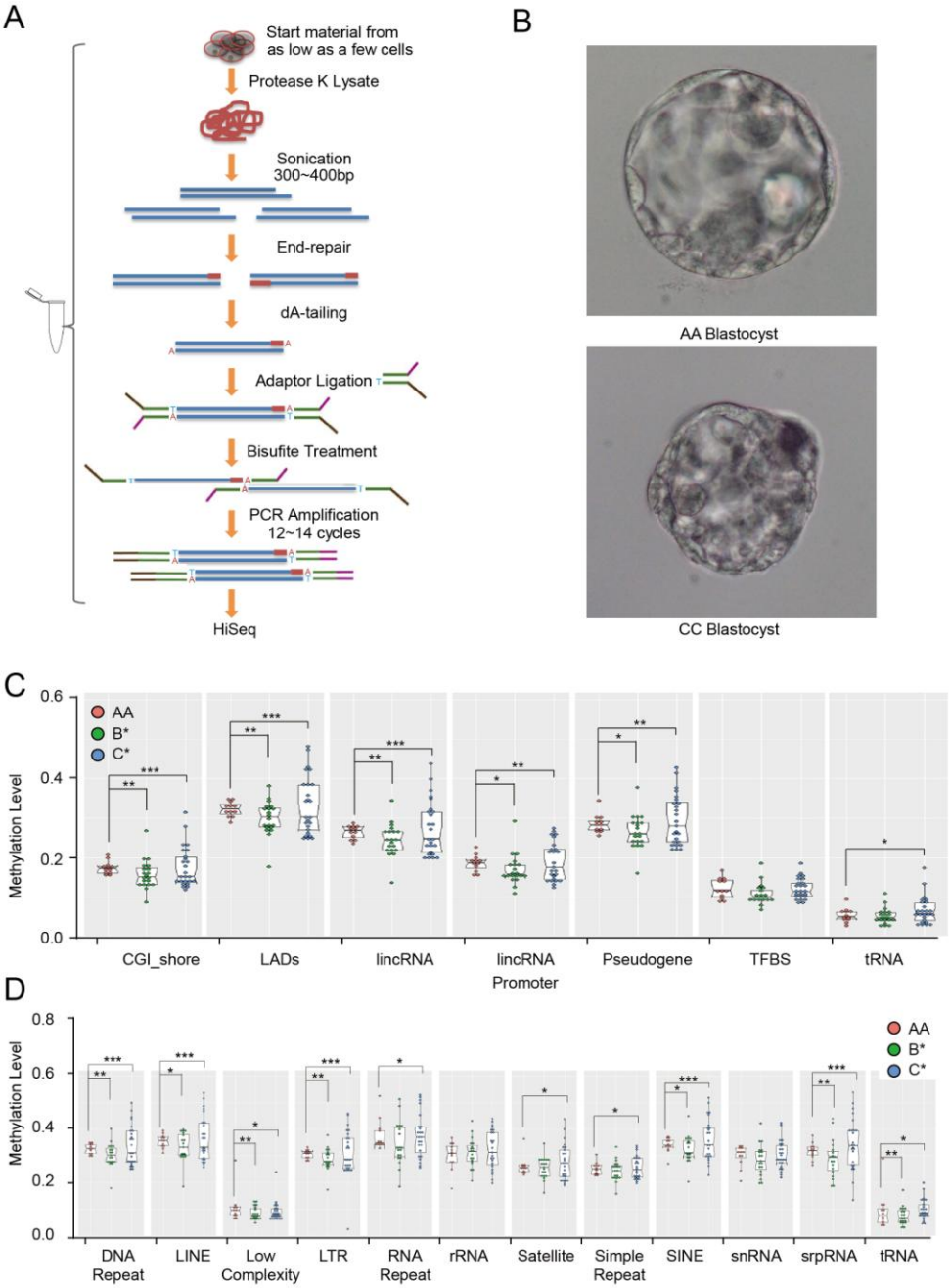


Figure S1. Epigenomic instability of blastocysts. (A) Illustration of the low-input MethylC-seq method for genome-wide methylome profiling. **(B)** Images of representative high (AA) and low grade (CC) blastocysts. (**C)** Average methylation levels of various functional elements in different blastocysts. LADs: Lamina-associated domains; TFBS: transcription factor binding sites. (**D)** Average methylation levels of various repeat elements in different blastocysts. DNA: DNA repeat elements. The P values of the homogeneity of variance between AA and B*or C*: *P<0.05; **P<0.01; ***P<0.001.

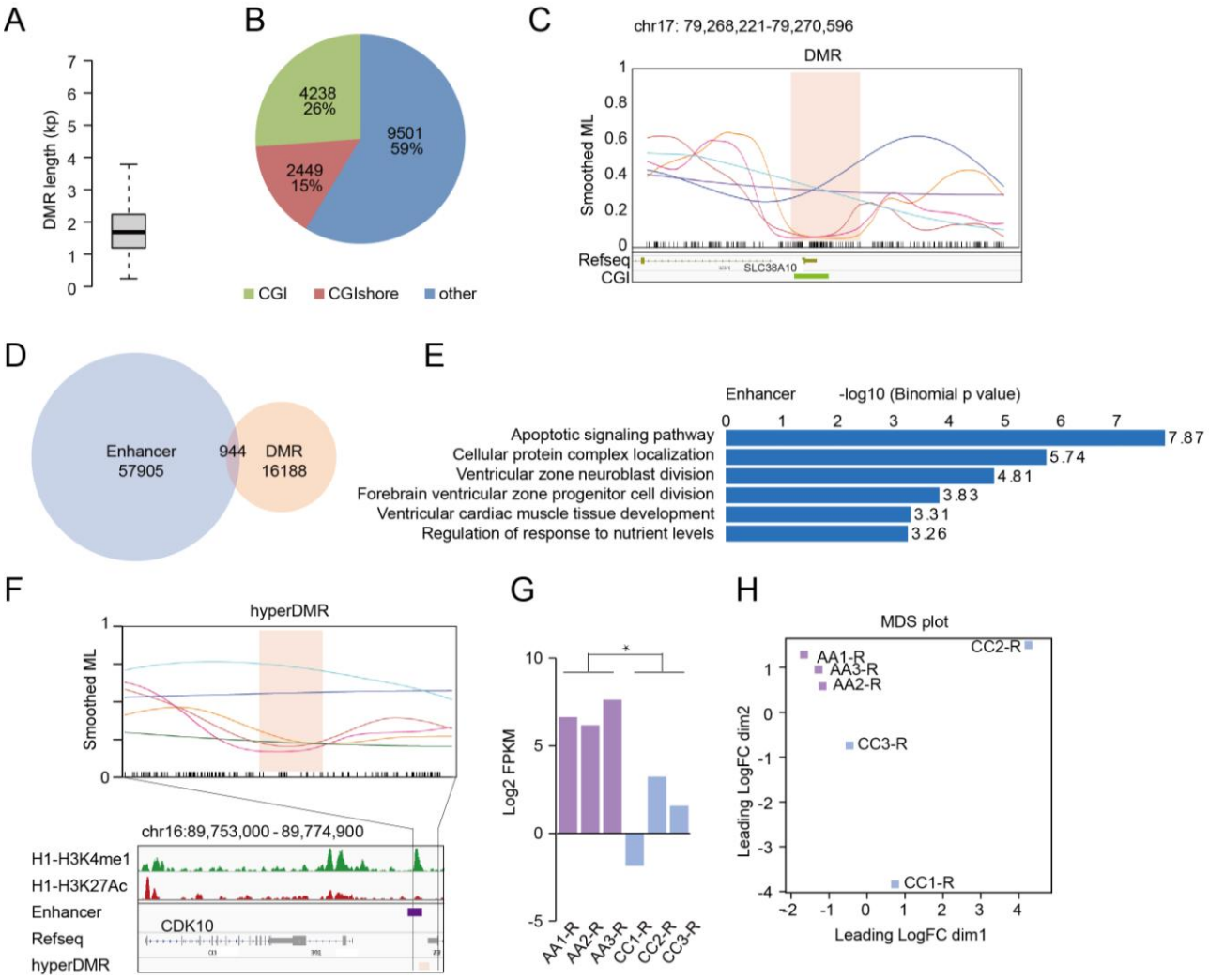


Figure S2. Functional analyses of altered methylation reprogramming. (**A)** The length distribution of DMRs. (**B**) The proportion of DMRs which locate in the regions of CGI or CGI shore. **(C)** Representative snapshot of a DMR (chromosome 17: 79,268,221-79,270,596) between high-quality and low-quality blastocysts. Smoothed methylation levels (sML) of high-quality and low-quality blastocysts are showed with extension ±5kb. Red, orange and purple indicate 1-4AA, 2-4AA and 1-6AA respectively. Blue, cyan, and blue violet indicate 9-3CC, 2-2CC and 8-3CC respectively. The DMR is indicated with pink shading. Short black bars indicate the location of CpG sites. (**D)** The overlap between enhancers and DMRs. (**E**) The GO enrichment of genes associated with differentially methylated enhancers. Terms with p-value < 0.05 were considered as significant enrichment. (**D)** The enhancer adjacent to gene CDK10 is differentially methylated. **(G)** The relative mRNA expression level of CDK10 in AA blastocysts and CC blastocysts. P-value = 0.05 (*). **(H)** PCA analyses according to the expression profile of each blastocyst.

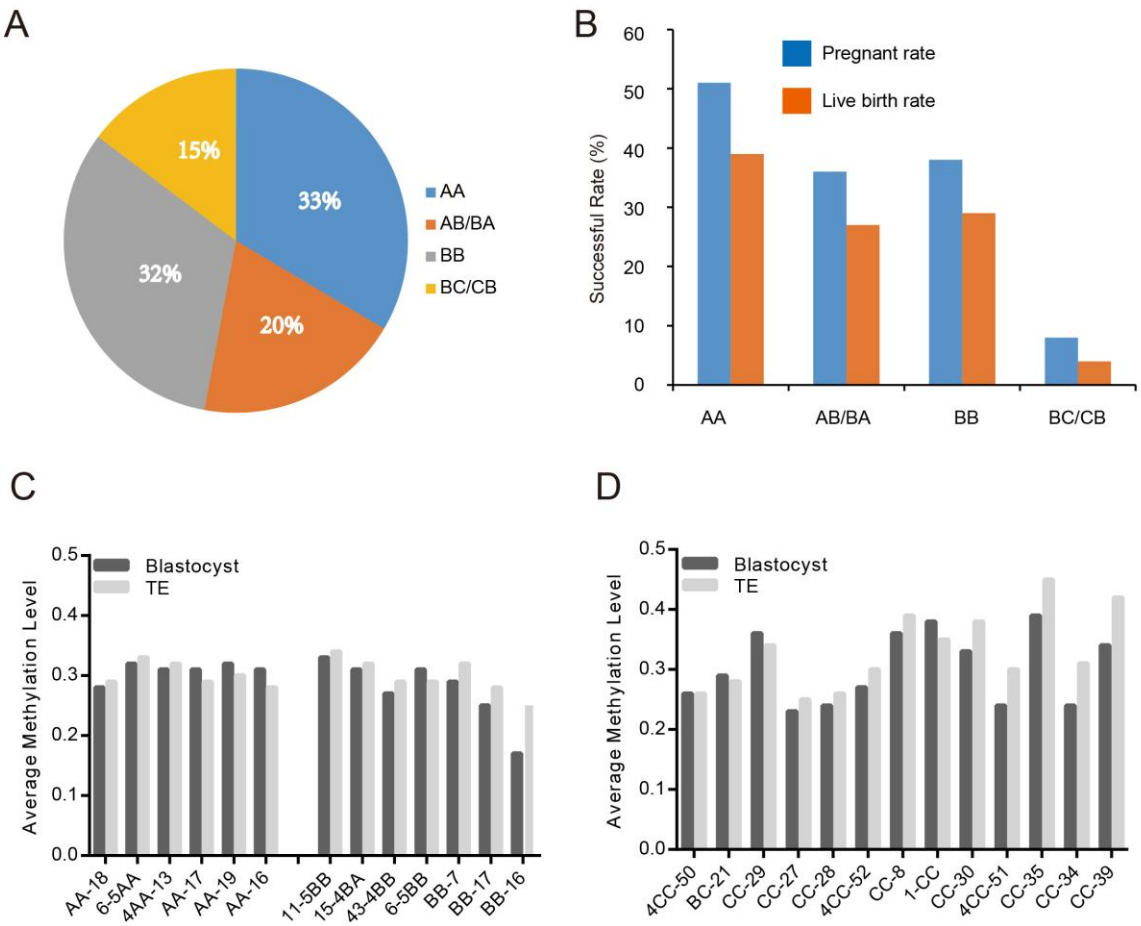


Figure S3. ART data for the blastocysts of different grades. (A) The proportion of different grade blastocysts used for fresh elective single embryo transfer (eSet) in ART practices. **(B)** The pregnant rate and live birth rate of the fresh blastocysts of different grades. **(C)** The methylation levels of the high-quality (AA) and middle-quality (B*) blastocysts with the paired TE samples. **(D)** The methylation levels of the low-quality (C*) blastocysts with the paired TE samples.

Supplemental Table S1

ID Blastocyst	ML	Cov (%)	BSCR (%)	Raw Reads Number	Mapping efficiency	ID Trophectoderm
1-4AA	0.3	71.92	99.01	230,974,819	45%	
2-4AA	0.32	78.42	98.78	154,101,136	42%	
1-6AA	0.27	67.08	98.81	98,820,731	41%	
13-6AA	0.29	47.58	98.84	109,323,604	38%	
14-5AA	0.31	48.02	98.58	106,066,575	29%	
25-3AA	0.28	59.11	99.07	55,418,836	39%	
6-5AA	0.32	13.72	98.62	74,698,378	35%	**6-5AA-TE**
13-4AA	0.31	22.59	98.99	102,987,785	59%	**13-4AA-TE**
16-5AA	0.31	12.17	99.14	104,518,819	60%	**16-5AA-TE**
17-5AA	0.31	17.91	99.15	124,718,872	71%	**17-5AA-TE**
18-4AA	0.28	17.97	98.96	108,208,374	64%	**18-4AA-T**
19-4AA	0.32	33.45	98.92	110,172,482	62%	**19-4AA-TE**
3-5BB	0.31	78.58	99.09	109,661,142	47%	
4-6BB	0.28	72.8	99.05	95,584,355	55%	
5-4BB	0.27	63.61	98.83	113,433,620	50%	
6-4BB	0.38	79.96	98.72	104,340,112	37%	
7-5BB	0.31	67.09	98.91	80,506,828	40%	
8-4BA	0.26	71.19	98.83	106,722,328	36%	
14-4BB	0.25	39.3	98.87	65,685,117	45%	
11-6BB	0.27	29.34	98.56	129,821,322	27%	
54-5BB	0.25	45.41	98.77	72,268,564	42%	
32-6BA	0.24	33.6	98.73	71,334,885	44%	
13-5BB	0.3	12.15	99.08	62,043,346	42%	
31-5BB	0.3	17.94	99.13	61,220,642	42%	
15-4BB	0.24	52.28	98.15	116,247,514	59%	
6-5BB	0.31	11.12	99.13	55,413,791	44%	**6-5BB-TE**
7-4BB	0.29	31	98.44	65,433,130	50%	**7-4BB-TE**
16-5BB	0.17	62.96	98.82	117,490,683	59%	**16-5BB-TE**
43-4BB	0.27	12.88	98.95	60,419,233	35%	**43-4BB-TE**
17-5BB	0.25	33.73	98.59	96,353,788	53%	**17-5BB-TE**
11-5BB	0.33	16.49	98.82	61,435,203	43%	**11-5BB-TE**
15-4BA	0.31	16.6	98.3	67,337,479	50%	**15-4BA-TE**
						13-4BB-TE
						18-5BB-TE
1-2CB	0.23	81.12	98.9	203,429,999	37%	
10-4CC	0.23	53.41	98.85	36,404,385	46%	
7-3CC	0.31	62.84	98.48	63,294,297	56%	

Supplemental Table S2

ID	ML(1x)	1xCov (%)	ML(3x)	3xCov (%)	ML(5x)	5xCov (%)
1-4AA	0.3	71.92	0.30	18.56	0.30	2.52
2-4AA	0.32	78.42	0.33	27.98	0.33	5.48
1-6AA	0.27	67.08	0.27	15.71	0.27	1.85
13-6AA	0.29	47.58	0.29	4.26	0.28	0.26
14-5AA	0.31	48.02	0.30	4.41	0.29	0.25
25-3AA	0.28	59.11	0.28	11.19	0.28	1.12
6-5AA	0.32	13.72	0.32	0.24	0.31	0.04
13-4AA	0.31	22.59	0.29	1.08	0.29	0.18
16-5AA	0.31	12.17	0.31	0.55	0.30	0.13
17-5AA	0.31	17.91	0.27	3.11	0.27	0.43
18-4AA	0.28	17.97	0.27	2.53	0.26	0.34
19-4AA	0.32	33.45	0.31	2.31	0.31	0.26
3-5BB	0.31	78.58	0.31	36.45	0.32	10.32
4-6BB	0.28	72.8	0.28	27.58	0.28	6.01
5-4BB	0.27	63.61	0.27	13.33	0.27	1.53
6-4BB	0.38	79.96	0.39	36.25	0.40	9.99
7-5BB	0.31	67.09	0.31	18.44	0.32	2.56
8-4BA	0.26	71.19	0.00	28.03	0.00	6.66
14-4BB	0.25	39.3	0.24	8.31	0.24	0.70
11-6BB	0.27	29.34	0.26	1.40	0.28	0.14
54-5BB	0.25	45.41	0.25	3.74	0.25	0.19
32-6BA	0.24	33.6	0.24	1.41	0.24	0.08
13-5BB	0.3	12.15	0.30	0.18	0.30	0.04
31-5BB	0.3	17.94	0.31	0.33	0.32	0.04
15-4BB	0.24	52.28	0.24	9.92	0.24	1.35
6-5BB	0.31	11.12	0.31	0.16	0.29	0.03
7-4BB	0.29	31	0.28	1.32	0.27	0.07
16-5BB	0.17	62.96	0.16	14.04	0.17	1.93
43-4BB	0.27	12.88	0.27	0.18	0.25	0.03
17-5BB	0.25	33.73	0.25	4.24	0.25	0.45
11-5BB	0.33	16.49	0.33	0.30	0.32	0.04
15-4BA	0.31	16.6	0.32	0.33	0.33	0.05
1-2CB	0.23	81.12	0.23	32.12	0.23	6.98
10-4CC	0.23	53.41	0.23	6.19	0.23	0.36
7-3CC	0.31	62.84	0.31	9.82	0.31	0.72
9-3CC	0.35	36.39	0.35	1.62	0.33	0.06
11-3CC	0.36	70.49	0.36	15.94	0.36	1.71
2-2CC	0.41	47.21	0.41	4.28	0.40	0.34

Note: The
blastocyst
trophectod
are good qu
poor qualit
grade of tr
methylatio
over total C

Supplemental Table S3

Sample	ML
1-4AA	0.3
2-4AA	0.32
1-6AA	0.27
13-6AA	0.29
14-5AA	0.31
25-3AA	0.28
6-5AA	0.32
13-4AA	0.31
16-5AA	0.31
17-5AA	0.31
18-4AA	0.28
19-4AA	0.32
3-5BB	0.31
4-6BB	0.28
5-4BB	0.27
6-4BB	0.38
7-5BB	0.31
8-4BA	0.26
14-4BB	0.25
11-6BB	0.27
54-5BB	0.25
32-6BA	0.24
13-5BB	0.3
31-5BB	0.3
15-4BB	0.24
6-5BB	0.31
7-4BB	0.29
16-5BB	0.17
43-4BB	0.27
17-5BB	0.25
11-5BB	0.33
15-4BA	0.31
1-2CB	0.23
10-4CC	0.23
7-3CC	0.31
9-3CC	0.35
11-3CC	0.36
2-2CC	0.41
8-3CC	0.46
13-CC	0.34
36-3CC	0.28
24-4CC	0.26
33-4CC	0.25
41-3CC	0.24
14-4CC	0.36
27-4CC	0.23

Supplemental Table S4

chr	start	end	idxStart	idxEnd	cluster	n	width
chr10	130006264	130010971	1047822	1047976	108747	155	4708
chr10	38689760	38692893	354306	354434	31616	129	3134
chr10	89263390	89266280	690846	690987	73621	142	2891
chr10	89575694	89579830	692978	693094	73931	117	4137
chr10	31607236	31610945	299206	299380	25030	175	3710
chr10	81741229	81743590	639258	639399	66328	142	2362
chr10	124766119	124769839	990890	991020	105372	131	3721
chr10	123686273	123688959	979555	979698	104605	144	2687
chr10	60027197	60029497	462578	462733	46892	156	2301
chr10	50745861	50748738	418443	418570	37527	128	2878
chr10	102025308	102028501	792617	792734	85577	118	3194
chr10	128075183	128078142	1029546	1029698	107450	153	2960
chr10	27442467	27446021	260421	260580	21649	160	3555
chr10	71811927	71815219	542865	543022	58508	158	3293
chr10	116851584	116853993	917919	918031	98936	113	2410
chr10	102746444	102748148	801609	801706	86083	98	1705
chr10	32216686	32219752	303390	303525	25628	136	3067
chr10	70715248	70717592	530688	530788	57703	101	2345
chr10	131908078	131910532	1070980	1071092	109957	113	2455
chr10	52383531	52385486	427614	427757	38662	144	1956
chr10	104628913	104631794	826776	826884	87360	109	2882
chr10	26221971	26224453	250093	250217	20554	125	2483
chr10	98345176	98347143	755727	755842	82589	116	1968
chr10	101491598	101493987	787280	787374	85142	95	2390
chr10	13042597	13044136	141016	141096	9292	81	1540
chr10	17270507	17273311	183845	184005	12414	161	2805
chr10	99185796	99187868	766114	766233	83152	120	2073
chr10	120513657	120515760	946890	947027	102212	138	2104
chr10	119805476	119807709	941402	941501	101569	100	2234
chr10	11384718	11387867	118643	118753	8384	111	3150
chr10	103577320	103579601	812534	812638	86667	105	2282
chr10	105991132	105993690	842354	842445	88352	92	2559
chr10	74385037	74387815	574085	574191	60102	107	2779
chr10	114205993	114207788	893409	893496	96822	88	1796
chr10	7829628	7831887	91558	91648	5157	91	2260
chr10	14919410	14922702	162293	162399	10495	107	3293
chr10	131933509	131936389	1071404	1071497	109971	94	2881
chr10	74854430	74857292	577947	578049	60560	103	2863
chr10	25304119	25306834	244181	244280	19647	100	2716
chr10	118440704	118444157	928074	928154	100444	81	3454
chr10	122738386	122741801	969989	970094	103932	106	3416
chr10	100991117	100994647	782101	782239	84689	139	3531
chr10	28591121	28592787	269664	269760	22691	97	1667
chr10	103453584	103457116	810738	810853	86589	116	3533
chr10	101189611	101191565	783935	784010	84876	76	1955
chr10	26985553	26987741	256224	256326	21248	103	2189