



Effect for Human Genomic Variation During the BMP4-Induced Conversion From Pluripotent Stem Cells to Trophoblast

Hai-tao Li^{††}, Yajun Liu^{1,2,3†}, Hongde Liu^{1*} and Xiao Sun^{1*}

¹ State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, China, ² The Second Affiliated Hospital of Zhengzhou University, Zhengzhou, China, ³ Academy of Medical Sciences of Zhengzhou University Translational Medicine Platform, Zhengzhou University, Zhengzhou, China

OPEN ACCESS

Edited by:

Jianzhong Su,
Wenzhou Medical University, China

Reviewed by:

Wei Jiang,
Nanjing University of Aeronautics
and Astronautics, China
Ruiping Wang,
The University of Texas MD Anderson
Cancer Center, United States

*Correspondence:

Hongde Liu
liuhongde@seu.edu.cn
Xiao Sun
xsun@seu.edu.cn

^{††} These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Genetics

Received: 07 December 2019

Accepted: 26 February 2020

Published: 07 March 2020

Citation:

Li H, Liu Y, Liu H and Sun X (2020)
Effect for Human Genomic Variation
During the BMP4-Induced Conversion
From Pluripotent Stem Cells to
Trophoblast. *Front. Genet.* 11:230.
doi: 10.3389/fgene.2020.00230

The role of genomic variation in differentiation is currently not well understood. Here, the genomic variations were determined with the whole-genome sequencing for three pairs of pluripotent stem cell lines and their corresponding BMP4-induced trophoblast cell lines. We identified ~3,500 single nucleotide variations and ~4,500 indels by comparing the genome sequenced data between the stem cell lines and the matched BMP4-induced trophoblast cell lines and annotated them by integrating the epigenomic and transcriptomic datasets. Relatively, introns enrich more variations. We found ~45% (42 genes) of the differentially expressed genes in trophoblasts that associate genomic variations. Six variations, located at transcription factor binding sites where H3K4me3 and H3K27ac are enriched in both H1 and H1_BMP4, were identified. The epigenetic status around the genomic variations in H1 was similar to that in H1_BMP4. This means that the variation-associated gene's expression change can not be attributed to epigenetic alteration. The genes associated with the six variations were upregulated in differentiation. We inferred that during the differentiation, an increased in the expression level of the MEF2C gene is due to a genomic variation in chromosomes 5: 88179358 A > G, which is at a binding site of TFs KLF16, NR2C2, and ZNF740 to MEF2C. Allele G shows a higher affinity to the TFs in the induced cells. The increased expression of MEF2C leads to an increased expression of TF MEF2C's target genes, subsequently affecting the differentiation. Although genomic variation should not be a dominant factor in differentiation, we believe that genomic variation could indeed play a role in the differentiation from stem cells into trophoblast.

Keywords: genomic variations, pluripotent stem cells, trophoblast, whole genome sequencing, epigenomic and transcriptomic data

INTRODUCTION

Stem cell differentiation involves a complex but poorly understood biological process. Genetic and epigenetic factors have in the past been intensively studied for this process. Some key transcription factors (TFs) affecting differentiation have been identified, such as POU5F1, SOX2, and KLF4 (Nichols et al., 1998; Takahashi and Yamanaka, 2006; Goolam et al., 2016). Previously, our lab

identified TFs for DNase I hypersensitive sites based on public available chromatin accessibility data of human embryonic stem cells (ESCs) and BMP4-induced trophoblast cell lines (Liu et al., 2017). The chromatin structure, including nucleosome positions, DNA methylation, and histone modifications also changed during the differentiation (Su et al., 2012; Xie et al., 2013; de Boni et al., 2018). However, genomic variation, another important factor, was rarely considered in the differentiation-related study.

The cells forming the outer layer of a blastocyst in early development are referred to as trophoblasts. These cells significantly contribute to the placenta and membrane, and also provides nutrients to the embryo. This layer of trophoblasts is collectively named “the trophoblast.” The trophoblast is the first cell type differentiated from the zygote during the first stage of pregnancy. Human ESCs provide a very useful model for studying the early development of human embryos and trophoblasts. However, the genomic variation that occurs in the trophoblast differentiation is not well studied *in vivo*. Since Xu et al. proved that bone morphogenetic protein 4 (BMP4) could induce human ESCs, efficiently differentiating to trophoblast lineage, multiple research institutes have adopted this system as a model to study trophoblast lineage specification *in vitro* (Xu et al., 2002). Roberts’s et al. considered this model reliability by analyzing gene expression profiling through RNA sequencing (RNA-Seq) technology (Yabe et al., 2016; Liu et al., 2017).

Recently, studies have demonstrated how genetic variation affects gene expression (Banovich et al., 2018; Delahaye et al., 2018). Kilpinen et al. (2013) suggest that genetic variation appears to be the primary driver of gene expression variation. Furthermore, DeBoever et al. (2017) show that the vast majority of genetic variations associated with gene expression levels are located in the regulatory regions of human induced pluripotent stem cells (iPSCs).

An integrative analysis using combinations of genomic, epigenomic, and transcriptomic data will provide a basis for biomarker discovery and help to provide insight of disease etiology. Genomic sequence variation, epigenetic factor, and gene expression are interdependent and jointly contribute to the normal functioning or dysfunction of tissue (Delahaye et al., 2018). For example, the sequencing variations can alter the TF binding strength to regulate gene expression directly (Karczewski et al., 2011; Madsen et al., 2018; Johnston et al., 2019). We are interested in the role of the genome variations in differentiation from human ESCs to trophoblasts.

In this study, we analyzed the genomic DNA sequences for three paired human ESCs and iPSC (H1, H9, and MRucR) (Sheridan et al., 2019). We demonstrated that some of the genomic variations can affect the differentiation by altering TFs’ binding affinity.

MATERIALS AND METHODS

Whole Genome Sequence of Cell Lines in This Study

Genomic DNA sequences were determined for three stem cells and their corresponding BMP4-induced cell lines

(**Supplementary Table S1**), in which two types of stem cells were human ESC lines (H1 and H9), and third an iPSC cell line, MRucR. The cell lines that differentiated into the trophoblast were named H1_BMP4, H9_BMP4, and MRucR_BMP4, respectively. The three cell lines and the matched BMP4 induced cell lines was obtained from the Roberts’s laboratory, University of Missouri. For more details of these three pairs of cell lines see Sheridan et al. (2019). MRucR iPSC was established from umbilical cords of babies born to mothers who experienced an early-onset form of preeclampsia during their pregnancies. The BMP4-inducing experiment was performed as previously described (Sheridan et al., 2019). Briefly, the trophoblast stem cells were exposed to BMP4 in combination with signaling inhibitors of ACTIVIN-A (A83-01) and FGF2 (PD173074) (BAP treatment) (Sheridan et al., 2019).

Library Preparation and Sequencing

The quality of isolated genomic DNA was verified using these three methods: (1) DNA purity and concentration were identified by NanoPhotometer® spectrophotometer (IMPLEN, CA, United States) (OD260/OD280). The Optical Density (OD) value of the qualified sample ranged between 1.8 – 2.0. (2) DNA degradation, and suspected RNA/Protein contamination were verified by electrophoresis on 1% agarose gels. (3) The concentration and purity of DNA samples were further quantified precisely by the Qubit DNA Assay Kit in Qubit®2.0 Fluorometer (Life Technologies, CA, United States). A total amount of 1 µg DNA per sample was required for library generation.

Paired-end DNA libraries were prepared according to the manufacturer’s instructions (Illumina Truseq Library Construction). First, 1.0 µg Genomic DNA was sheared into an average size of 350 base pair (bp) fragments by Covaris S220 sonicator. Second, the ends of the gDNA fragments were repaired; 3’ ends were adenylated. Both ends of the gDNA fragments were ligated at the 3’ ends with paired-end adaptors (Illumina) with a single ‘T’ base overhang, and purified using AMPure SPRI beads from Agencourt. The size distribution and concentration of the libraries were then determined using Agilent 2100 Bioanalyzer and were qualified by real-time PCR (2 nM), respectively. Lastly, DNA libraries were sequenced on Illumina HiSeq X according to the manufacturer’s instructions for paired-end 150 bp reads.

Whole Genome Sequencing Data Analysis

The raw image files obtained from the Hi-Seq platform were processed with the Illumina pipeline for base calling and stored as FASTQ format (Raw data). Quality control (QC) was as follows: (1) to filter reads with adapter contamination (>10 bp aligned to the adapter allowing ≤ 10% mismatches). (2) to discard the reads containing more than 10% uncertain nucleotides. (3) to discard the paired reads when a single read has more than 50% low quality nucleotides (Phred quality < 5).

After quality control, the sequenced reads were mapped to the GRCh37 assembly of the human genome by Burrows-Wheeler Aligner (BWA) software using default settings (Li

and Durbin, 2009). Subsequently, we used Samtools (Li et al., 2009) and Picard¹ with default settings to sort reads, remove duplicated reads, and to generate the final bam file. If one or more pair read(s) had multiple mapping positions, the best one was selected. If there were multiple best mapping positions, we randomly picked one.

Variation Calling and Functional Annotation

The analysis flowchart is shown in **Figure 1A**. The genomic variations were determined with the whole-genome sequencing for three pairs of pluripotent stem cell lines and their corresponding BMP4-induced trophoblast cell lines and annotated them by integrating the epigenomic and transcriptomic datasets. We only focused on the direct effect on the differentiation of genomic variation which is located in those genomic regions where epigenetic markers remain comparable between H1 and H1_BMP4.

Genomic variations were identified by comparing genomic DNA sequences between stem cell lines and matched BMP4-induced trophoblast cell lines using the following steps. Single nucleotide variation (SNV) and small somatic insertions and deletions (indels) were identified using the Strelka2 (Kim et al., 2018) tool. All 'PASS' calls identified by Strelka2 were retained for downstream analyses. BreakDancer (Chen et al., 2009) was applied to detect structural variation (SV). The SVs that received minimal confidence scores (90 for insertions, inversions, deletions, and translocations) were selected for downstream analyses.

The ANNOVAR tool was used to produce statistical analyses of the SNVs/indels (Wang et al., 2010). The variation position, variation type, conservative prediction, and other information were obtained at this step through a variety of databases, including DbSNP, 1000 Genome, and the reference sequence (Sherry et al., 2001; Pruitt et al., 2007; The 1000 Genomes Project Consortium, 2010). The variations were then assigned to genes by associating them with the nearest transcription start sites (TSSs) using the BEDOPS toolkit (Neph et al., 2012) and Gencode v21 human annotation (Harrow et al., 2012).

The regions include exon, intron, intergenic region, UTR5, UTR3, and the upstream (variant overlaps 1-kb region upstream of transcription start site)/downstream (variant overlaps 1-kb region downstream of transcription end site) regions.

The annotation detail of these three cell lines is listed in **Supplementary Excel File 1**. The version of the bioinformatics tools used is listed in **Supplementary Table S2**.

Public Datasets Analyzed in This Study

The goal of the work is to assess the effect of the genomic variations during stem cell differentiation. We therefore need to exclude the effect of the epigenomic variation (**Figure 1A**). We thus combined DNA methylation, histone modifications, chromatin accessibility, and transcriptomic datasets for the

H1 and H1_BMP4 (differentiated with BAP treatment) cell lines. All the ChIP-Seq and DNase-Seq data in H1 and H1_BMP4 were generated by the ENCODE Consortium² and retrieved from the GEO database according to their accession number (**Supplementary Table S3**) (Lister et al., 2011; Roadmap Epigenomics Consortium et al., 2015; Davis et al., 2018). The sequencing data were mapped to the human reference genome GRCh37/hg19 by BWA. For ChIP-Seq data, we performed peak calling by using the MACS2 tool with default settings (Zhang et al., 2008). Differential peaks between the two cell lines were also identified with MACS2. We are interested in the genomic variations where there is no significant change between the two matched cell lines in the epigenomic data, especially for histone marks H3K27ac and H3K4me3 since the two represent the active state for enhancers and promoters.

The methylation state at CpG sites in whole genome bisulfite sequencing (WGBS) data were mapped to GRCh38/hg38 (**Supplementary Table S3**) (Roadmap Epigenomics Consortium et al., 2015). We used the liftOver tool to transform the genome coordinates from hg38 to hg19 (Dreszer et al., 2012).

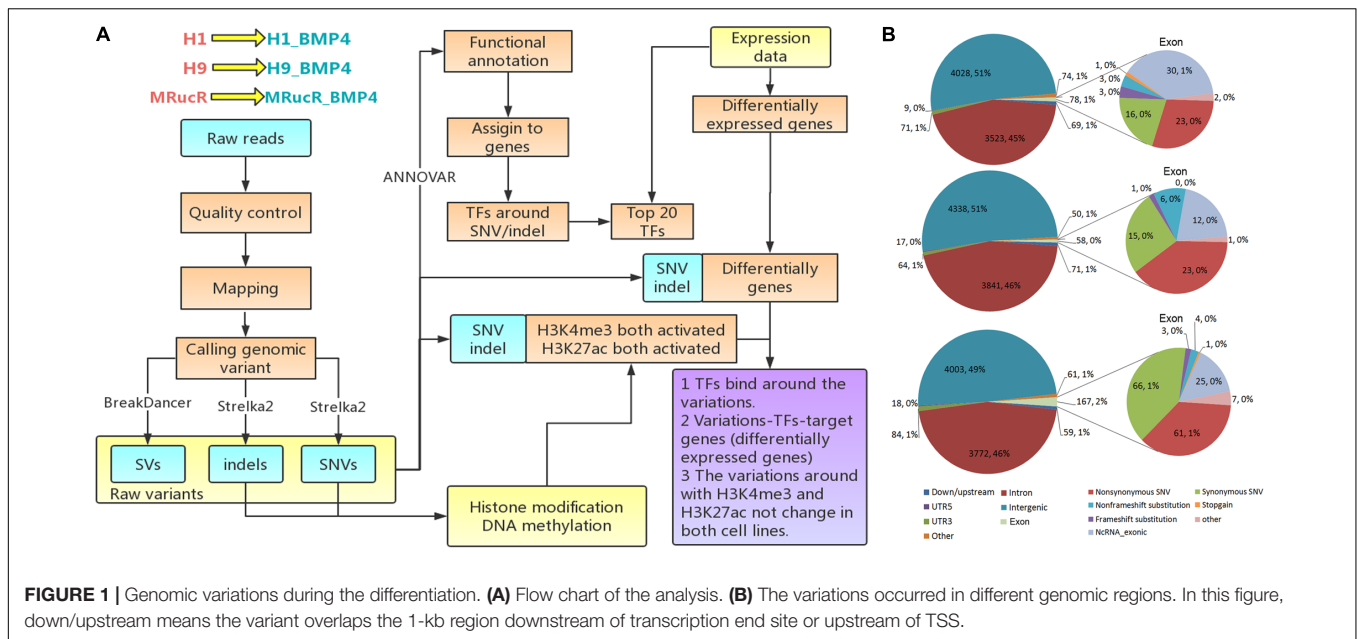
The gene expression data (Xie et al., 2013) were retrieved from the ENCODE project (**Supplementary Table S3**) with the following identifiers: ENCFF245NXP, ENCFF809MAC, ENCFF051SBM, and ENCFF787HWI. RNA abundance was represented as the logarithm of the transcripts per million (TPM) provided by the RSEM program (Li and Dewey, 2011). The data was used in two aspects. One to identify the differentially expressed genes, so as to check which genes are influenced by the genomic variations. The other is to find the transcription factors that truly have a function in cells. This was done by checking the expression level of the gene that encodes the transcription factor (**Figure 1A**). Differentially expressed mRNAs were identified with limma (Ritchie et al., 2015). Genes with Fold Change > 2 or <1/2 and false discovery rate (FDR) *p*-value < 0.05 were identified as significantly differentially expressed genes between ESC H1 and differentiated to trophoblast H1_BMP4.

Prediction for Transcription Factor Binding Sites (TFBSs) Around the Genomic Variations

To estimate the change of the TF binding affinity due to the genomic variation, the 150-bp DNA sequences were extracted around the genomic variations and inputted into a bioinformatics tool, HOMER (Heinz et al., 2010), to calculate the affinity (**Figure 1A**). Each of the sequences includes one kind of genomic variation. The affinity of a TF to a specific DNA sequence can be estimated by comparing the DNA sequence and the motif of the TF. The motif is the most favorable binding DNA pattern of the TF and can be represented with Position Weight Matrix (PWM). The comparison will be a score on the motif (motif score). A high score means a high affinity between the DNA sequence and the TF. We listed the TFs with a motif score ≥ 10 in the stem cell and its BMP4-induced cells and found the difference of the TFs between the cells.

¹<http://broadinstitute.github.io/picard>

²<https://www.encodeproject.org/>



RESULTS

The Distribution of Genomic Variations During the Differentiation

We performed the high-throughput DNA sequencing for three pluripotent stem cell lines (H1, H9, and MRucR) and differentiated trophoblasts (H1_BMP4, H9_BMP4, and MRucR_BMP4). After quality control (Supplementary Figure S1), genomic variations were identified by comparing the DNA sequence between the stem cell lines and matched BMP4-induced trophoblast cell lines. Reads mapping rate was more than 98%, and the sequencing depth was beyond 31X (Supplementary Tables S4, S5). We used Strelka2 to identify SNV and indels. SV was identified with BreakDancer (Chen et al., 2009; Kim et al., 2018). The Flow chart of our analysis is shown in Figure 1A. We identified approximately 3,500 SNVs and 4,500 indels in the ESC and iPSC lines, respectively (Table 1A). The count of SV was relatively rare (~230) (Table 1A).

The frequency of SNVs was counted. According to our statistics, the single-base substitution patterns were similar in the three pluripotent stem cells (Supplementary Figure S2a). The highest frequency of SNV was the conversion of cytosine to thymine (C > T), and the lowest frequency of SNV was the conversion of guanine to cytosine (G > C). Among the SVs (Table 1B), the proportion of translocation, including inter- and intra-, was high (about 77%~87%). The percentage of insertion inversion and deletion was 12% ~ 22%. We did not detect the structure deletion events here. SV had similar distributions in these three pairs of cell lines (Figure 1B).

We then annotated SNV and indels with ANNOVAR (Wang et al., 2010). The SNV and indels were counted in different genomic regions (Supplementary Excel File 1). To calculate fold enrichments, ratios of the variation counts in a certain category of genomic regions were divided by the proportions of this category

in the whole genome length. The variations were significantly enriched in the introns (Figure 2).

Distribution profiles of SNV/indels in the 2500-bp genomic regions around the TSSs seemly occurred in a periodic pattern in the cell lines (Supplementary Figure S2b). SNP frequency spectra show striking periodicities across nucleosomal regions, and SNPs have a preference for nucleosomes (Langley et al., 2014). We therefore speculated that there is an association between the patterns and nucleosome distribution in this region.

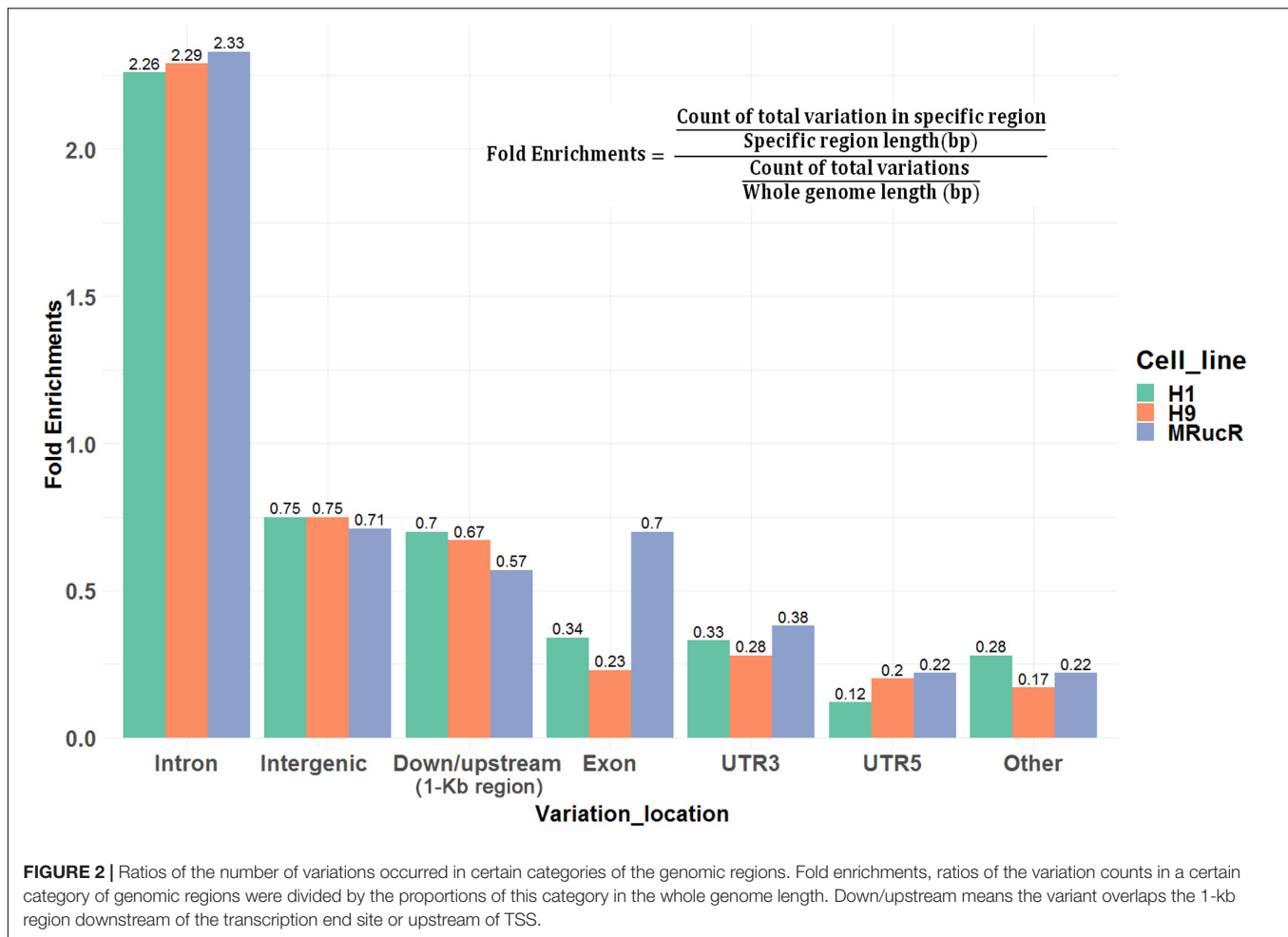
Transcription Factors That Bind Around Genomic Variations Sites

A genomic variation may directly alter the affinity of TF's binding as this variation occurs exactly at TFBSs, thus influencing the

TABLE 1 | Number of genomic variants (A) and structural variants (B) in which stem cells differentiate into trophoblast cells.

(A)					
Matched hESC-trophoblast cell lines	Indels	SNV	SV		
H1	4071	3781	233		
H9	4703	3736	235		
MRucR	4729	3435	225		
(B)					
Matched hESC-trophoblast cell lines	INS	INV	ITX	CTX	DEL
H1	31	2	92	108	0
H9	19	10	96	110	0
MRucR	44	6	84	91	0

Indels, small somatic insertions and deletions; SNV, single nucleotide variant; SV, structural variation; INS, insertion; INV, inversion; ITX, intra-chromosomal translocation; CTX, inter-chromosomal translocation; DEL, deletions.

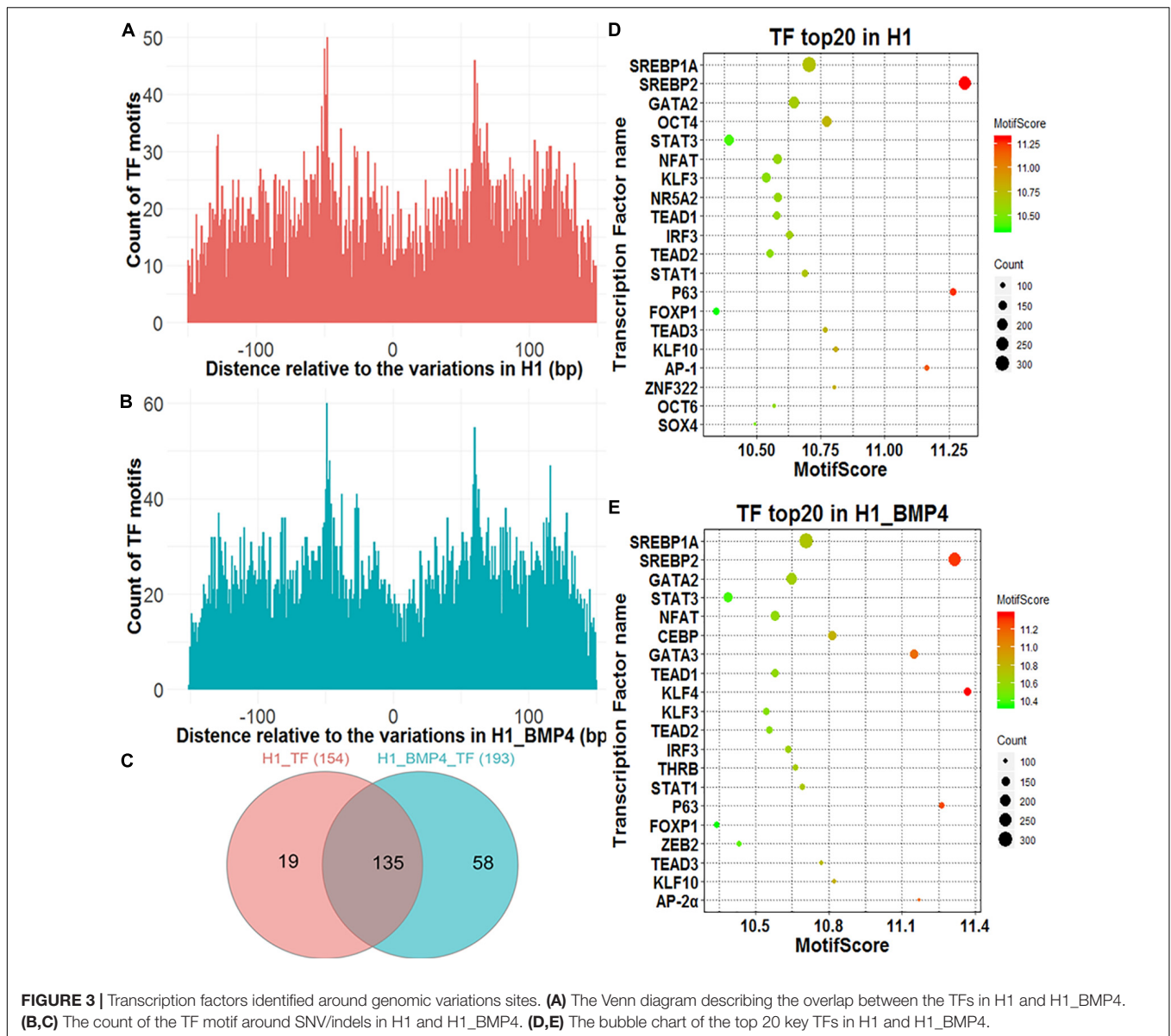


transcription levels of the downstream target genes of the TFs. We assessed the effect of the variations on the alteration of TF binding with HOMER (Heinz et al., 2010). The 150-bp DNA sequences around the variation sites were inputted into HOMER to identify motifs of the TFs and compared the affinity of the TFs variation between the stem cell and matched BMP4-induced trophoblast cell lines. In **Figures 3A,B**, although counts of TF motifs around SNV/indels positions decreased, the number was not zero. There were some motifs that were far away from the SNV/indels. We intend to find out what TFs bind at the SNV/indels-harbored sites.

A TF with its motif in one genomic region does not mean the TF indeed binds at the region in a cell because the TF maybe is not expressed in the cell. We therefore need to check if the TF is abundant enough in the cell, namely, to check if the expression level of the gene encoding the TF is high enough. To do this, a necessary step is to exclude those TFs with a low abundance. The expression of TF-encoding genes was shown for H1 and H1_BMP4 (**Supplementary Figure S3**). Here, the cutoff of $\text{TPM} \geq 3$ was used to select the TFs that indeed exist in the cell. We identified 154 and 193 TFs in H1 and H1_BMP4 cells, respectively, with an overlap of 135 TFs (**Figure 3C**). The top 20 key TFs were chosen by setting $\text{TPM} \geq 3$ and motif

score ≥ 10 (**Figures 3D,E** and **Supplementary Tables S6, S7**). In the top 20 TFs of H1 and H1_BMP4, some widely known to be involved in placental development were found, such as GATA3 (Kubaczka et al., 2015), POU5F1 (Wang et al., 2012), and KLF4 (Abad et al., 2013). A target gene of SREBP1A is the transcriptional repressor BHLHB2, which also promotes the differentiation of stem cells to trophoblast giant cells (Lecomte et al., 2010). ZEB2 has recently been identified to play critical roles in the regulation of the epithelial-mesenchymal transition and trophoblast differentiation (DaSilva-Arnold et al., 2019).

We identified the TFs that differ in affinity around the genomic variation sites between H1 and H1_BMP4. The TFs were chosen under conditions of $\text{TPM} \geq 3$, motif score ≥ 10 , and motif score count difference ($\text{abs} \geq 10$). We found three TFs (Zfp281, OCTs, and KLF3), whose affinities near the genomic variations differ between H1 and H1_BMP4 cell lines. The three kinds of TFs have a higher motif score count (count of motif score ≥ 10) in H1 than in H1_BMP4. Interestingly, these TFs are widely known to be involved in trophoblast differentiation. For example, Zfp281 (Krüppel-like zinc finger transcription factor), a zinc finger transcription factor, which shapes the transcriptome of trophoblast cells and regulates early placental development, has also been investigated in a recent article (Ishiuchi et al., 2019).



Moreover, the other TFs were widely known to be important in placental development (**Supplementary Table S8**).

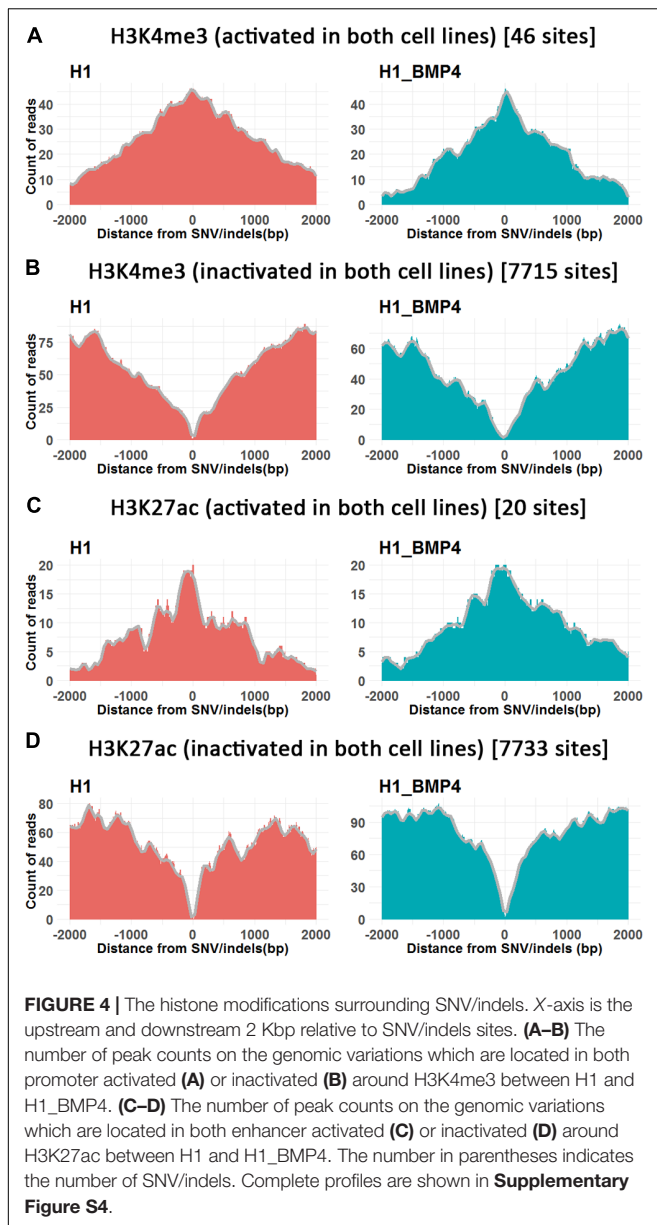
In short, we found that there were indeed some genomic variations at TFBSs.

Correlation Between Genome Variations and Epigenomics

Genome variation was reported to be correlated to epigenetic alteration. Li et al. demonstrated that about 2/3 of eQTLs were due to variations that altered chromatin accessibility or histone marking (Li et al., 2016). We studied the correlation between genetic variations and chromatin alteration in H1 and H1_BMP4 cell lines. The sequencing data of H3K4me3, H3K27ac, H3K27me3, and DNA methylation data were retrieved from ENCODE (**Supplementary Table S5**).

In this analysis, the genomic variations were divided into four categories according to whether the variation site was at the peaks of the histone modifications and DNA methylation in H1 and H1_BMP4 (Lister et al., 2011; Roadmap Epigenomics Consortium et al., 2015). We counted the peaks of the histone modifications and DNA methylation within the 4K-bp genomic region around SNV/indel sites (**Figure 4** and **Supplementary Figure S4**).

The enriching regions (peaks) of DNA methylation and H3K27me3 were considered as “inhibited” regions. Modified regions (peaks) by H3K4me3 and H3K27ac indicate “promoter activated” and “enhancer inactivated” regions, respectively. By comparing the epigenomic states of the regions around the genomic variations between the stem cells and matched BMP4-induced trophoblast cell lines, 66 variations that do not associate an epigenomic alteration were identified (**Table 2**, and



Supplementary Excel File 2). Next, we evaluated the effects of the genomic variation around which the epigenetic modifications do not alter between H1 and H1_BMP4.

The Direct Effect of Genomic Variation on Differentiation

We are interested in the gene expression change that is only caused by the genomic variation instead of the epigenetic variation. We therefore only focused on the genomic variations which are located in those genomic regions where epigenetic markers remain comparable between H1 and H1_BMP4.

We identified 5,688 differentially expressed genes between H1 and H1_BMP4 (FDR p -value ≤ 0.05 and fold change ≥ 2 or $\leq 1/2$) (Figure 5A). In the genes, there are 3,763 up-regulated and 1,925 down-regulated genes. Since each genomic variation was assigned

TABLE 2 | Number of the genomic variations that occur in regions associating different switches of the epigenomic modifications between H1 and H1_BMP4.

Regions	Modification types	State switch from H1 to H1_BMP4	The number of genomic variations
–	H3K27me3 and DNA methylation	In -> T	0
–	H3K27me3 and DNA methylation	T -> In	0
Enhancer	H3K27ac	I -> A	18
Enhancer	H3K27ac	A -> I	12
Promoter	H3K4me3	I -> A	30
Promoter	H3K4me3	A -> I	30
Promoter	H3K4me3	A -> A	46
Enhancer	H3K27ac	A -> A	20

In, inhibited; T, transcribed; I, inactivated; A, activated.

to a gene, we could identify the genomic variations that associate differentially expressed genes.

One-hundred-and-twenty-nine differential expression genes were determined using stricter criteria (p -value $\leq 1E-5$ and fold change ≥ 4 or $\leq 1/4$). We found a total of 26 genes with variations in the differentiation process out of the 129 differential expression genes (Figure 5B). The functional annotation tool of KEGG was applied to the 26 genes (Figure 5C). As expected, two pathways implicated in regulating human ESC differentiation was found with an adjusted p -value < 0.05 , namely the “signaling pathways regulating pluripotency of stem cells” and the “TGF-beta signaling pathway,” which represent the differentiation pathways. For instance, inhibition of the TGF-beta signaling pathway could be sufficient for the derivation and long term expansion of human trophoblast cells and the placenta (Xu et al., 2017; Knöfler et al., 2019).

In the induction by BMP4, BMP ligands are bound to the BMP4 complex, then transphosphorylate the intracellular signaling proteins, Smad1/5/8. The phospho-Smad1/5/8 interacts with Smad4, and the complex translocates into the nucleus where it interacts with transcription cofactors and regulates expression of target genes in a cell type-specific manner (Shimasaki et al., 2004). Based on these biological processes, 235 target genes of Smad 1/4/5/6 were retrieved from the TF targets in the Harmonizome database (Rouillard et al., 2016). Among them, 94 are differentially expressed between H1 and H1_BMP4 (FDR p -value ≤ 0.05 and fold change ≥ 2 or $\leq 1/2$). There are 72 up-regulated (30.64%) and 22 down-regulated genes (9.36%). In the 94 BMP4-induced genes, 42 (44.68%) contain the genomic variations (Figure 5D), suggesting a tight association between the variation and BMP4 induced differentiation.

In particular, by excluding the variations associating the epigenetic modification alteration, six genomic variations, which associate differentially expressed genes, were identified (Table 3, and Supplementary Excel File 3). Five of the six variations were located in the promoter regions of genes MEF2C, TNFAIP8,

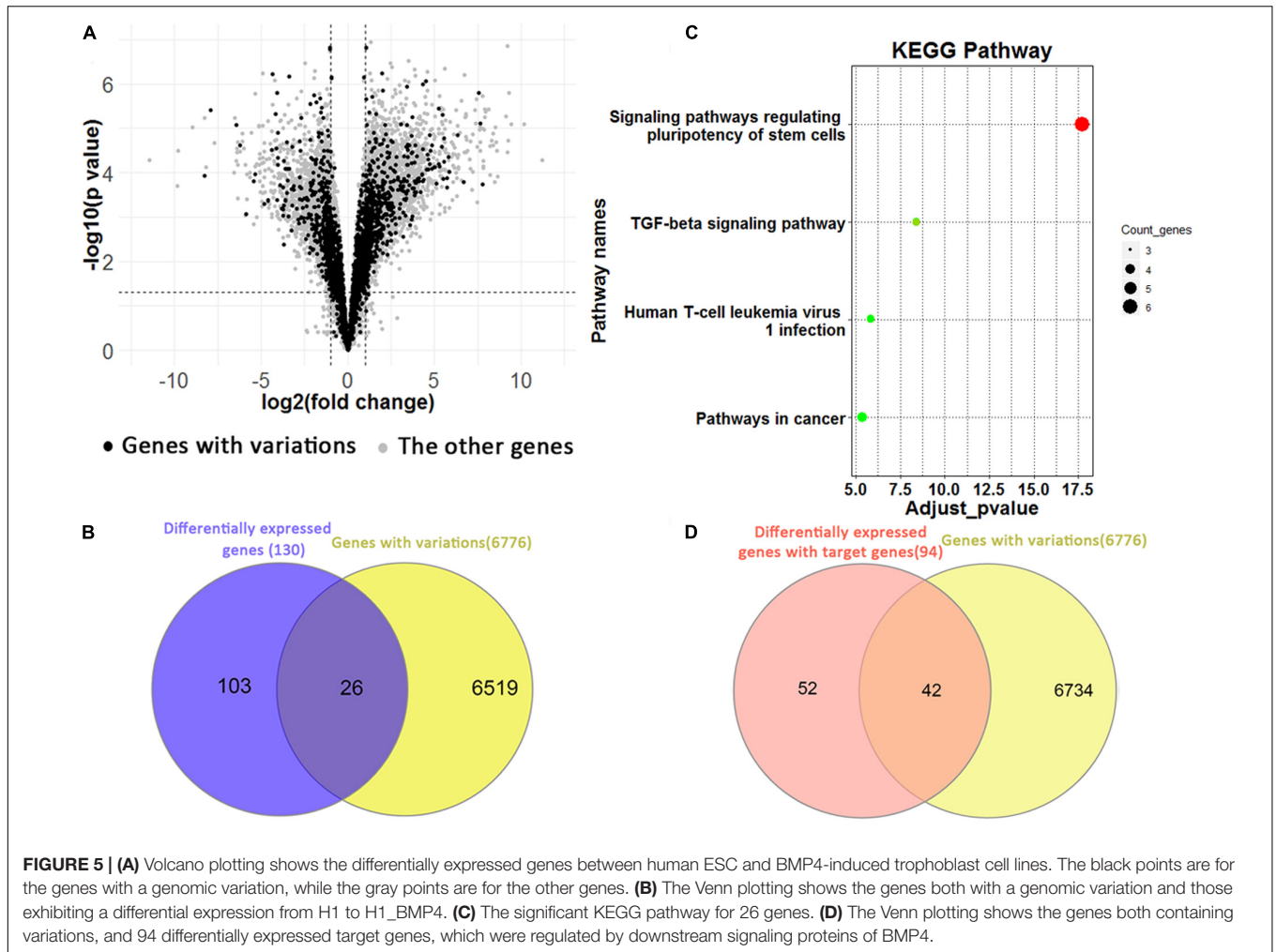


FIGURE 5 | (A) Volcano plotting shows the differentially expressed genes between human ESC and BMP4-induced trophoblast cell lines. The black points are for the genes with a genomic variation, while the gray points are for the other genes. **(B)** The Venn plotting shows the genes both with a genomic variation and those exhibiting a differential expression from H1 to H1_BMP4. **(C)** The significant KEGG pathway for 26 genes. **(D)** The Venn plotting shows the genes both containing variations, and 94 differentially expressed target genes, which were regulated by downstream signaling proteins of BMP4.

TABLE 3 | Information of six genomic variations.

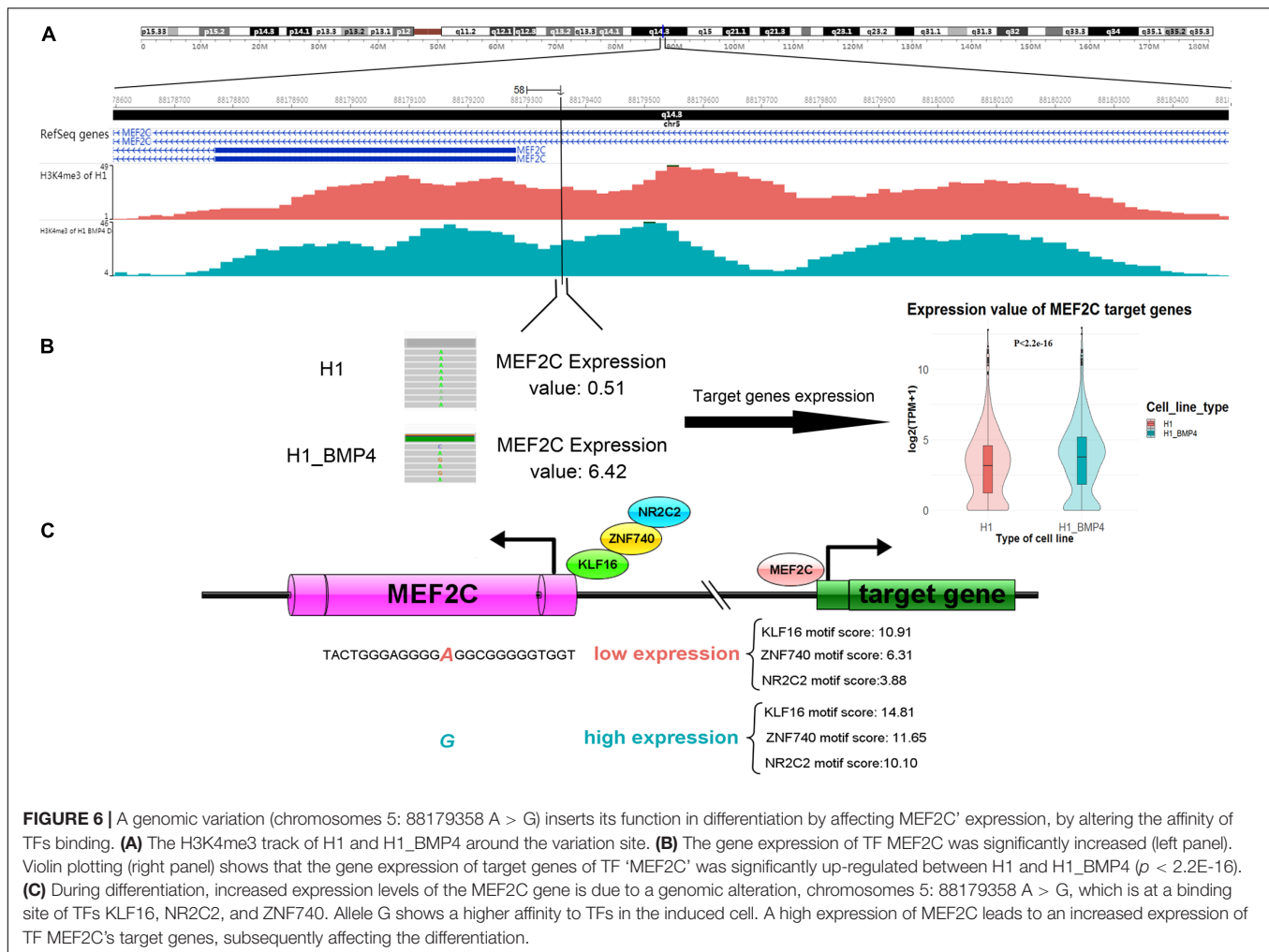
Chr	Start	Ref	Alt	Location	Gene.refGene	Modification types
chr5	118605129	ATG	A	Intronic	TNFAIP8	H3K4me3
chr5	88179358	A	G	ncRNA_intronic	MEF2C	H3K4me3
chr17	62339443	G	T	Intronic	TEX2	H3K4me3
chr7	41740828	G	A	ncRNA_intronic	INHBA-AS1	H3K4me3
chr1	20811135	T	C	Intronic	CAMK2N1	H3K4me3
chr7	94023911	A	G	UTR5	COL1A2	H3K27ac

Column Location lists the genomic region of the variation. ncRNA_intronic means non-coding RNA genes. Gene.refGene is the gene associating the variation. Modification types indicate the H3K4me3 or H3K27ac around the variation in both H1 and H1_BMP4.

TABLE 4 | The TF motif scores of MEF2C.

TF_Name	Start	End	Strand	H1 Expression(TPM)	H1_BMP4 Expression (TPM)	H1_BMP4_motif score	H1_motif score	motif score_diff
KLF16	11	21	-	14	15.52	14.81	10.90	3.90
ZNF740	7	16	-	5.4	7.29	12.63	7.09	5.54
ZNF740	6	15	-	5.4	7.29	10.66	5.53	5.13
NR2C2	10	24	+	9.27	12.99	10.09	3.87	6.22

The TFs in this table were chosen as follows: (1) the motif scores of the TFs larger than 10 in H1 or H1_BMP4; (2) TPM value of the TF larger than 3 in both H1 and H1_BMP4; (3) the fold change of the expressed TF genes between H1 and H1_BMP4 less than 2.



TEX2, INHBA-AS1, and CAMK2N1. The remaining one was at the enhancer of gene COL1A2.

We highlighted a core function of the six genes (MEF2C, TNFAIP8, TEX2, INHBA-AS1, CAMK2N1, and COL1A2) in differentiation and placenta development. TNFAIP8 plays a role in immune homeostasis, inflammatory responses, tumor genesis, and development. TNFAIP8 is also highly expressed in most normal human tissues especially for immunity-related tissues like the placenta (Zhang et al., 2018). INHBA encodes a member of the TGF- β (transforming growth factor-beta) superfamily of proteins which has been proven to promote the differentiation of human embryonic stem cells into trophoblasts (Puc at, 2007). COL1A2 encodes one of the chains for type I collagen, the fibrillar collagen found in most connective tissues, and it is an early stage marker of osteoblast differentiation (Parisuthiman et al., 2005). MEF2C plays a pivotal role in myogenesis, neural crest, and craniofacial development, and may have an influence on maintaining the differentiated state of muscle cells (Zweier et al., 2010). In a recent study, dysregulation of MEF2 expression or signaling in early pregnancy may be associated with placenta-related pregnancy disorders, including preeclampsia (Li et al., 2017). In total,

the six genes both associate a genomic variation in their regulatory regions and show differential expression from H1 to H1_BMP4, suggesting that the genomic variations associated with differentiation.

Since the six genomic variations are in non-coding regions, we assessed the effect of the variation in altering the binding affinity of TFs. We applied JASPAR2018³ to calculate the motif score of 12-bp DNA sequences around the six SNVs/indels to the TFs that bind at those variation-harbored regions (Khan et al., 2018). We listed the motif scores in **Supplementary Excel File 3**. We found that the allele G of a variation (chr5: 88179358 A > G), which is at the enhancer of gene MEF2C, can increase the motif scores ≥ 10 in the induced cells (**Table 4**). It should be considered, however, that the MEF2C expression level is ~ 4 fold higher in H1_BMP4 than in H1. The results mean that the genomic variation accounts for the increase of MEF2C expression by increasing TFs' affinity at MEF2C's promoter. Moreover, the H2K27ac does not show a significant change between the two cell lines. We

³<http://jaspar.genereg.net/>

confirmed that the MEF2C expression increase is caused by the genomic variation.

Importantly, the MEF2C gene itself encodes a TF protein. More recently, MEF2 was proven to regulate human trophoblast differentiation and invasion (Li et al., 2017). We retrieved the list of 954 MEF2C target genes reported in ENCODE TF targets (Rouillard et al., 2016). We then compared the TF MEF2C target genes' expression between H1 and H1_BMP4. The result showed that they are significantly up-regulated in the transition from human ESC and the trophoblast (paired sample *t*-test, $p < 2.2E-16$) (Figure 6).

Functional annotations by Enrichr (Kuleshov et al., 2016) on KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene ontology) suggest the target genes of TF MEF2C were enriched in the "NF-kappa B signaling pathway" (Pollheimer and Knöfler, 2005), "Primary immunodeficiency," "Systemic lupus erythematosus," "positive regulation of NF-kappaB transcription factor activity," "positive regulation of immune response," and so on (adjust *p*-value ≤ 0.05 , Supplementary Figure S5). Literature suggests that the placenta is an important immune-related tissue (Erkers et al., 2017). We therefore considered that the target genes of TF MEF2C have a significant impact on the cell differentiation process.

Briefly, the gene expression of MEF2C was significantly increased due to the high-affinity binding sites of TFs like KLF16, NR2C2, and ZNF740, a variation site located in MEF2C (chromosomes 5: 88179358 A > G). Further, the expression levels of these three TFs were not significantly different between H1 and H1_BMP4, suggesting that MEF2C's expression increase is not caused by a rise in the abundance of the TFs of MEF2C, but due to the genomic variation (Figure 6). Interestingly, literature indicates that simultaneous depletion of KLF2, KLF4, and KLF5 leads to differentiation of the embryonic stem cell, and it has been postulated that other members of the KLF family, such as KLF16, may play similar roles in embryonic stem cells (Jiang et al., 2008; Andreoli et al., 2010). It has been established that NR2C2 (TR4) plays a critical role in embryonic development and differentiation.

NR2C2 is expressed in blastocysts and embryonic stem cells and can act as transcriptional activators in hESC (Shyr et al., 2009; O'Geen et al., 2010). Altogether, this indicates that due to the up-regulation of MEF2C, its target genes were up-regulated. Importantly, the upregulation of MEF2C is caused by the genomic variation (chr5: 88179358 A > G), which alters the affinity of MEF2C's TFs.

DISCUSSION

Understanding the genetic underpinnings of complex traits remains a major challenge in human genetics. In this study, we obtained paired genomic DNA sequences of human ESC and the trophoblast from three cell lines (H1, H9, and MRucR) through whole genome sequencing, and integrated the epigenomic (DNA

methylation, histone modifications and chromatin accessibility) and transcriptomic datasets to investigate the impact of the genome variations in human ESC differentiation to the trophoblast. We found that the SNVs and indels generally tend to be located in the intron regions rather than in the other regions.

We focused on the gene expression variation caused by genomic variation rather than the epigenetic variation. Six genomic variations were identified. One of them, located in MEF2C (chromosomes 5: 88179358 A > G), is a TF. This SNV increased the TF binding strength to regulate gene expression directly. Thereby, leading to an increase in the expression of downstream target genes affecting the differentiation of human ESC into the trophoblast. It suggested that the variations in the non-coding region played an important role in the differentiation process.

The inducer, BMP4, is the most significant factor to differentiate to the trophoblast. BMP4 is able to inhibit the Activin/Nodal signaling pathway and activate the BMP signaling pathway, which is required for human ESCs to differentiate into trophoblasts (Xu et al., 2002). We showed that ~45% (42 genes) of the differentially expressed BMP4-induced genes associated with genomic variations. Although genomic variations are not possible to be the only dominant factor in differentiation, some genomic variations indeed have an effect on differentiation. There are two limitations in this work. The first is that the conclusion is confined to only three pairs of cell lines. A similar analysis should be carried out in more cell lines of iPSC. The second limitation is that comprehensive biochemical experiments are still needed to validate the conclusion. CRISPR technology in cultured cells could be employed to prove the role of genomic variation in the differentiation process (Zhou et al., 2014, p. 2).

DATA AVAILABILITY STATEMENT

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Wang et al., 2017) at the BIG Data Center (Zhang et al., 2019), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession number CRA002190, which is publicly accessible at <https://bigd.big.ac.cn/gsa>.

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design, analyzed and discussed the results. XS conceived the project and completed the core program. HaL, HoL, and YL performed the computational analysis. HaL and HoL wrote the manuscript.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (Nos. 81830053, 61972084)

and the Key Research & Development Program of Jiangsu Province (BE2016002-3).

ACKNOWLEDGMENTS

We thank Professor Toshihiko Ezashi (University of Missouri) and Professor R. Michael Roberts (University of Missouri) for providing genomic materials obtained from three human pluripotent stem cells (H1, H9, and MRucR). We would also like

to thank Professor Toshihiko Ezashi and Yue Hou for critical reading and helpful discussions for the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00230/full#supplementary-material>

REFERENCES

- Abad, M., Mosteiro, L., Pantoja, C., Cañamero, M., Rayon, T., Ors, I., et al. (2013). Reprogramming *in vivo* produces teratomas and iPSC cells with totipotency features. *Nature* 502, 340–345. doi: 10.1038/nature12586
- Andreoli, V., Gehrau, R. C., and Bocco, J. L. (2010). Biology of Krüppel-like factor 6 transcriptional regulator in cell life and death. *IUBMB Life* 62, 896–905. doi: 10.1002/iub.396
- Banovich, N. E., Li, Y. L., Raj, A., Ward, M. C., Greenside, P., Calderon, D., et al. (2018). Impact of regulatory variation across human iPSCs and differentiated cells. *Genome Res.* 28, 122–131. doi: 10.1101/gr.224436.117
- Chen, K., Wallis, J. W., McLellan, M. D., Larson, D. E., Kalicki, J. M., Pohl, C. S., et al. (2009). BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat. Methods* 6, 677–681. doi: 10.1038/nmeth.1363
- DaSilva-Arnold, S. C., Kuo, C. Y., Davra, V., Remache, Y., Kim, P. C. W., Fisher, J. P., et al. (2019). ZEB2, a master regulator of the epithelial-mesenchymal transition, mediates trophoblast differentiation. *Mol. Hum. Reprod.* 25, 61–75. doi: 10.1093/molehr/gay053
- Davis, C. A., Hitz, B. C., Sloan, C. A., Chan, E. T., Davidson, J. M., Gabdank, I., et al. (2018). The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.* 46, D794–D801. doi: 10.1093/nar/gkx1081
- de Boni, L., Gasparoni, G., Haubenreich, C., Tierling, S., Schmitt, I., Peitz, M., et al. (2018). DNA methylation alterations in iPSC- and hESC-derived neurons: potential implications for neurological disease modeling. *Clin. Epigenet.* 10:13. doi: 10.1186/s13148-018-0440-440
- DeBoever, C., Li, H., Jakubosky, D., Benaglio, P., Reyna, J., Olson, K. M., et al. (2017). Large-scale profiling reveals the influence of genetic variation on gene expression in human induced pluripotent stem cells. *Cell Stem Cell* 20, 533.e7–546.e7. doi: 10.1016/j.stem.2017.03.009
- Delahaye, F., Do, C., Kong, Y., Ashkar, R., Salas, M., Tycko, B., et al. (2018). Genetic variants influence on the placenta regulatory landscape. *PLoS Genet.* 14:e1007785. doi: 10.1371/journal.pgen.1007785
- Dreszer, T. R., Karolchik, D., Zweig, A. S., Hinrichs, A. S., Raney, B. J., Kuhn, R. M., et al. (2012). The UCSC Genome Browser database: extensions and updates 2011. *Nucleic Acids Res.* 40, D918–D923. doi: 10.1093/nar/gkr1055
- Erkers, T., Stikvoort, A., and Uhlin, M. (2017). Lymphocytes in placental tissues: immune regulation and translational possibilities for immunotherapy. *Stem Cells Int.* 2017:5738371. doi: 10.1155/2017/5738371
- Goolam, M., Scialdone, A., Graham, S. J. L., Macaulay, I. C., Jedrusik, A., Hupalowska, A., et al. (2016). Heterogeneity in Oct4 and Sox2 targets biases cell fate in 4-cell mouse embryos. *Cell* 165, 61–74. doi: 10.1016/j.cell.2016.01.047
- Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., et al. (2012). GENCODE: the reference human genome annotation for the ENCODE project. *Genome Res.* 22, 1760–1774. doi: 10.1101/gr.135350.111
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple combinations of lineage-determining transcription factors Prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589. doi: 10.1016/j.molcel.2010.05.004
- Ishichi, T., Ohishi, H., Sato, T., Kamimura, S., Yorino, M., Abe, S., et al. (2019). Zfp281 shapes the transcriptome of trophoblast stem cells and is essential for placental development. *Cell Rep.* 27, 1742–1754. doi: 10.1016/j.celrep.2019.04.028
- Jiang, J., Chan, Y.-S., Loh, Y.-H., Cai, J., Tong, G.-Q., Lim, C.-A., et al. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat. Cell Biol.* 10, 353–360. doi: 10.1038/ncb1698
- Johnston, A. D., Simões-Pires, C. A., Thompson, T. V., Suzuki, M., and Grealley, J. M. (2019). Functional genetic variants can mediate their regulatory effects through alteration of transcription factor binding. *Nat. Commun.* 10, 1–16. doi: 10.1038/s41467-019-11412-11415
- Karczewski, K. J., Tatonetti, N. P., Landt, S. G., Yang, X., Slifer, T., Altman, R. B., et al. (2011). Cooperative transcription factor associations discovered using regulatory variation. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13353–13358. doi: 10.1073/pnas.1103105108
- Khan, A., Fornes, O., Stigliani, A., Gheorghie, M., Castro-Mondragon, J. A., van der Lee, R., et al. (2018). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* 46, D260–D266. doi: 10.1093/nar/gkx1126
- Kilpinen, H., Waszak, S. M., Gschwind, A. R., Raghav, S. K., Witwicki, R. M., Orioli, A., et al. (2013). Coordinated effects of sequence variation on DNA binding, Chromatin structure, and transcription. *Science* 342, 744–747. doi: 10.1126/science.1242463
- Kim, S., Scheffler, K., Halpern, A. L., Bekritsky, M. A., Noh, E., Källberg, M., et al. (2018). Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* 15, 591–594. doi: 10.1038/s41592-018-0051-x
- Knöfler, M., Haider, S., Saleh, L., Pollheimer, J., Gamage, T. K. J. B., and James, J. (2019). Human placenta and trophoblast development: key molecular mechanisms and model systems. *Cell. Mol. Life Sci.* 76, 3479–3496. doi: 10.1007/s00018-019-03104-3106
- Kubaczka, C., Senner, C. E., Cierlitz, M., Araúzo-Bravo, M. J., Kuckenberger, P., Peitz, M., et al. (2015). Direct induction of trophoblast stem cells from murine fibroblasts. *Cell Stem Cell* 17, 557–568. doi: 10.1016/j.stem.2015.08.005
- Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44, W90–W97. doi: 10.1093/nar/gkx1126
- Langley, S. A., Karpen, G. H., and Langley, C. H. (2014). Nucleosomes shape DNA polymorphism and divergence. *PLoS Genet.* 10:e1004457. doi: 10.1371/journal.pgen.1004457
- Lecomte, V., Meugnier, E., Euthine, V., Durand, C., Freyssen, D., Nemoz, G., et al. (2010). A new role for sterol regulatory element binding protein 1 transcription factors in the regulation of muscle mass and muscle cell differentiation. *Mol. Cell Biol.* 30, 1182–1198. doi: 10.1128/MCB.00690-699
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, L., Gong, X., and Rubin, L. P. (2017). Expression of MEF2 transcription factors in human placenta and involvement in trophoblast invasion and differentiation. *FASEB J.* 31:692.8. doi: 10.1096/fasebj.31.1_supplement.692.8

- Li, Y. I., van de Geijn, B., Raj, A., Knowles, D. A., Petti, A. A., Golan, D., et al. (2016). RNA splicing is a primary link between genetic variation and disease. *Science* 352, 600–604. doi: 10.1126/science.aad9417
- Lister, R., Pelizzola, M., Kida, Y. S., Hawkins, R. D., Nery, J. R., Hon, G., et al. (2011). Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471, 68–73. doi: 10.1038/nature09798
- Liu, Y., Ding, D., Liu, H., and Sun, X. (2017). The accessible chromatin landscape during conversion of human embryonic stem cells to trophoblast by bone morphogenetic protein 4. *Biol. Reprod.* 96, 1267–1278. doi: 10.1093/biolre/iox028
- Madsen, J. G. S., Rauch, A., Van Hauwaert, E. L., Schmidt, S. F., Winnefeld, M., and Mandrup, S. (2018). Integrated analysis of motif activity and gene expression changes of transcription factors. *Genome Res.* 28, 243–255. doi: 10.1101/gr.227231.117
- Neph, S., Kuehn, M. S., Reynolds, A. P., Haugen, E., Thurman, R. E., Johnson, A. K., et al. (2012). BEDOPS: high-performance genomic feature operations. *Bioinformatics* 28, 1919–1920. doi: 10.1093/bioinformatics/bts277
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., et al. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379–391. doi: 10.1016/S0092-8674(00)81769-81769
- O'Geen, H., Lin, Y.-H., Xu, X., Echipare, L., Komashko, V. M., He, D., et al. (2010). Genome-wide binding of the orphan nuclear receptor TR4 suggests its general role in fundamental biological processes. *BMC Genomics* 11:689. doi: 10.1186/1471-2164-11-689
- Parisuthiman, D., Mochida, Y., Duarte, W. R., and Yamauchi, M. (2005). Biglycan modulates osteoblast differentiation and matrix mineralization. *J. Bone Miner. Res.* 20, 1878–1886. doi: 10.1359/JBMR.050612
- Pollheimer, J., and Knöfler, M. (2005). Signalling pathways regulating the invasive differentiation of human trophoblasts: a review. *Placenta* 26, S21–S30. doi: 10.1016/j.placenta.2004.11.013
- Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 35, D61–D65. doi: 10.1093/nar/gkl842
- Pucéat, M. (2007). TGF β in the differentiation of embryonic stem cells. *Cardiovasc. Res.* 74, 256–261. doi: 10.1016/j.cardiores.2006.12.012
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Roadmap Epigenomics Consortium Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518, 317–330. doi: 10.1038/nature14248
- Rouillard, A. D., Gundersen, G. W., Fernandez, N. F., Wang, Z., Monteiro, C. D., McDermott, M. G., et al. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* 2016:baw100. doi: 10.1093/database/baw100
- Sheridan, M. A., Yang, Y., Jain, A., Lyons, A. S., Yang, P., Brahmasani, S. R., et al. (2019). Early onset preeclampsia in a model for human placental trophoblast. *Proc. Natl. Acad. Sci. U.S.A.* 116, 4336–4345. doi: 10.1073/pnas.1816150116
- Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., et al. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 29, 308–311. doi: 10.1093/nar/29.1.308
- Shimasaki, S., Moore, R. K., Otsuka, F., and Erickson, G. F. (2004). The bone morphogenetic protein system in mammalian reproduction. *Endocr. Rev.* 25, 72–101. doi: 10.1210/er.2003-0007
- Shyr, C.-R., Kang, H.-Y., Tsai, M.-Y., Liu, N.-C., Ku, P.-Y., Huang, K.-E., et al. (2009). Roles of testicular orphan nuclear receptors 2 and 4 in early embryonic development and embryonic stem cells. *Endocrinology* 150, 2454–2462. doi: 10.1210/en.2008-1165
- Su, J., Shao, X., Liu, H., Liu, S., Wu, Q., and Zhang, Y. (2012). Genome-wide dynamic changes of DNA methylation of repetitive elements in human embryonic stem cells and fetal fibroblasts. *Genomics* 99, 10–17. doi: 10.1016/j.ygeno.2011.10.004
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676. doi: 10.1016/j.cell.2006.07.024
- The 1000 Genomes Project Consortium, (2010). A map of human genome variation from population scale sequencing. *Nature* 467, 1061–1073. doi: 10.1038/nature09534
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38:e164. doi: 10.1093/nar/gkq603
- Wang, Y., Song, F., Zhu, J., Zhang, S., Yang, Y., Chen, T., et al. (2017). GSA: genome sequence archive*. *Genomics Proteom. Bioinform.* 15, 14–18. doi: 10.1016/j.gpb.2017.01.001
- Wang, Z., Oron, E., Nelson, B., Razis, S., and Ivanova, N. (2012). Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* 10, 440–454. doi: 10.1016/j.stem.2012.02.016
- Xie, W., Schultz, M. D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., et al. (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* 153, 1134–1148. doi: 10.1016/j.cell.2013.04.022
- Xu, J. G., Zhu, S. Y., Heng, B. C., Dissanayaka, W. L., and Zhang, C. F. (2017). TGF- β 1-induced differentiation of SHED into functional smooth muscle cells. *Stem Cell Res. Ther.* 8:10. doi: 10.1186/s13287-016-0459-450
- Xu, R.-H., Chen, X., Li, D. S., Li, R., Addicks, G. C., Glennon, C., et al. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat. Biotechnol.* 20:1261. doi: 10.1038/nbt761
- Yabe, S., Alexenko, A. P., Amita, M., Yang, Y., Schust, D. J., Sadovsky, Y., et al. (2016). Comparison of syncytiotrophoblast generated from human embryonic stem cells and from term placentas. *PNAS* 113, E2598–E2607. doi: 10.1073/pnas.1601630113
- Zhang, L., Liu, R., Luan, Y., and Yao, Y. (2018). Tumor necrosis factor- α induced protein 8: pathophysiology, clinical significance, and regulatory mechanism. *Int. J. Biol. Sci.* 14, 398–405. doi: 10.7150/ijbs.23268
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. doi: 10.1186/gb-2008-9-9-r137
- Zhang, Z., Zhao, W., Xiao, J., Bao, Y., Wang, F., Hao, L., et al. (2019). Database resources of the BIG data center in 2019. *Nucleic Acids Res.* 47, D8–D14. doi: 10.1093/nar/gky993
- Zhou, H. Y., Katsman, Y., Dhaliwal, N. K., Davidson, S., Macpherson, N. N., Sakthidevi, M., et al. (2014). A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. *Genes Dev.* 28, 2699–2711. doi: 10.1101/gad.248526.114
- Zweier, M., Gregor, A., Zweier, C., Engels, H., Sticht, H., Wohlleber, E., et al. (2010). Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum. Mutat.* 31, 722–733. doi: 10.1002/humu.21253

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li, Liu, Liu and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.