

RESEARCH

Open Access



# Alternative splicing signatures in preimplantation embryo development

Geng G. Tian<sup>1</sup>, Jing Li<sup>1,3\*</sup> and Ji Wu<sup>1,2,4\*</sup>

## Abstract

**Background:** Preimplantation embryo development is a highly ordered sequence of processes and it requires a precise temporal and spatial control of gene expression. Alternative splicing (AS) is a crucial process that changes the genomic instructions into functional proteins, playing a critical role in the regulation of gene expression. Therefore, studies of AS can significantly improve our understanding of transcription and splicing events in preimplantation embryo development.

**Results:** To study signatures of AS in embryo development, we firstly identified the critical stage for gene transcription during the preimplantation embryo development. By analyzing single cells RNA-seq (scRNA-seq) data, we found that the two-cell stage is a critical stage for gene transcription in preimplantation embryo development. Further study showed that AS was widespread in preimplantation embryo development, especially at the two-cell stage. In combination with high-throughput chromosome conformation (Hi-C) data, we demonstrated that AS genes were highly enriched in TAD boundaries, while they had no relationship with the A/B compartment and TAD.

**Conclusion:** Our results provide new insight into the relationship among AS, gene transcription and chromatin structure in preimplantation embryo development.

## Introduction

Embryogenesis is a sequential series of dynamic processes including cell division and growth, and the elaboration of differentiation programs leading to cell fate specification. Preimplantation embryo development plays a crucial role in fetal development and requires a precise temporal and spatial control of gene expression. Specifically, profiling of the gene expression in preimplantation embryo development can show when and how cell fate decisions are made. Initially, RNA-seq could not be applied to the study of preimplantation embryo development because of its technological limitations. However, through the development of RNA-seq technology for a low input of cells,

such as Cel-seq [1], Smart-seq [2], and RAGE-seq [3], our understanding of the gene expression of small cell populations has been greatly extended. Recently, single-cell RNA-seq (scRNA-seq) techniques including chromium 10× [4] and Drop-seq [5] have been developed, which can detect the gene expression in just a single cell; these approaches have significantly advanced our knowledge of cell systems. These new technologies have also greatly facilitated the study of preimplantation embryo development. For example, Tang's group made major progress in understanding preimplantation embryo development [6, 7] by profiling the landscape of gene expression during this process. At the transcriptome level, one of the most important types of pre-mRNA processing is alternative splicing (AS) events, which can increase the diversity of the transcriptome and proteome [8] and play an important role in the regulation of gene expression in a variety of eukaryotes. Approximately 95% of multi-exon genes in human undergo AS. Misregulation of AS leads to splicing defects causing severe diseases. Therefore, studies of

\*Correspondence: jing.li@sjtu.edu.cn; jiwu@sjtu.edu.cn

<sup>1</sup> Renji Hospital, Key Laboratory for the Genetics of Developmental & Neuropsychiatric Disorders (Ministry of Education), Bio-X Institutes, School of Medicine, Shanghai Jiao Tong University, Shanghai 200240, China

Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

alternative splicing can significantly improve our understanding of transcription and splicing events in embryo development. Although previous studies reported that AS was specific to different tissues and cell types [9], it remains unclear when and how AS occurs in preimplantation embryo development.

Concurrent with this progress in structure variation, Hi-C was developed to clarify chromatin organization [10]. To better understand the relationship between transcription and chromatin structure in preimplantation embryo development, Xie and Liu's group developed the small-cell Hi-C technology (scHi-C) [11, 12] and revealed that the high-order chromatin structure underwent reprogramming, in which the topologically associating domains (TADs) disappeared in the MII and zygote stages, while TADs started to construct at the two-cell stage and completed at the eight-cell stage. Moreover, they used RNA-seq data to systematically analyze the relationship between gene expression and chromatin organization, but they did not consider genes whose products varied due to AS events. Therefore, it has remained unknown whether AS events are related to the high-order chromatin organization.

## Methods

### RNA-seq data analysis

RNA-seq data were downloaded from the online version of a previous paper [13]. The raw data were trimmed with BBmap (version 38.16) to obtain clean data, which had a quality score higher than 15 ( $Q > 15$ ), followed by removal of the adaptors of those reads. Then, the reads were mapped to the mouse reference genome (mm9) with STAR (version 2.7.1a) [14] software using two-pass mode with the GENCODE annotation file (version Mouse Release M1). Gene expression (FPKM) levels were calculated using Cufflinks (version v2.2.1) [15]. Then, the gene expression was clustered by the c-means method to analyze the time course of gene expression, which was supported by the MFuzz R package [16]. The counts of gene expression were quantified by Rsubread software (version 1.6.2) [17]. DESeq2 package (version 1.10.1) [18] was used to identify the differential gene expression with the following cut-offs:  $FDR < 0.05$  and  $\log_2FC > 2$ .

### Differential splicing and alternative splicing identification

Gene products encoded by different numbers of exons of the source gene were identified using the DEXSeq R package [19] with  $FDR < 0.05$ . Then, rMATS (version v4.0.3beta) was used to identify the different types of AS events with the splice junction file (SJ.out.tab) generated by STAR and two replicate bam files. According to the rMATS results, normalized percent spliced-in (PSI) values were calculated to identify the different AS events

( $FDR < 0.05$ ). The AS genes were visualized with DEXseq and rmats2sashimiplo.

### Protein–protein network

AS gene interactions were input into STRING (version 11.0) [20] to identify the protein–protein network with the highest confidence ( $> 0.9$ ). The network result was then displayed using Gephi software (version 0.9.2) [21].

### Gene ontology (GO) analysis

GO term enrichment analysis was performed using the Gene Ontology website (<https://geneontology.org/>, version 14) [22]. We chose an FDR threshold (Benjamini-corrected p value) of less than 0.05 to represent significance.

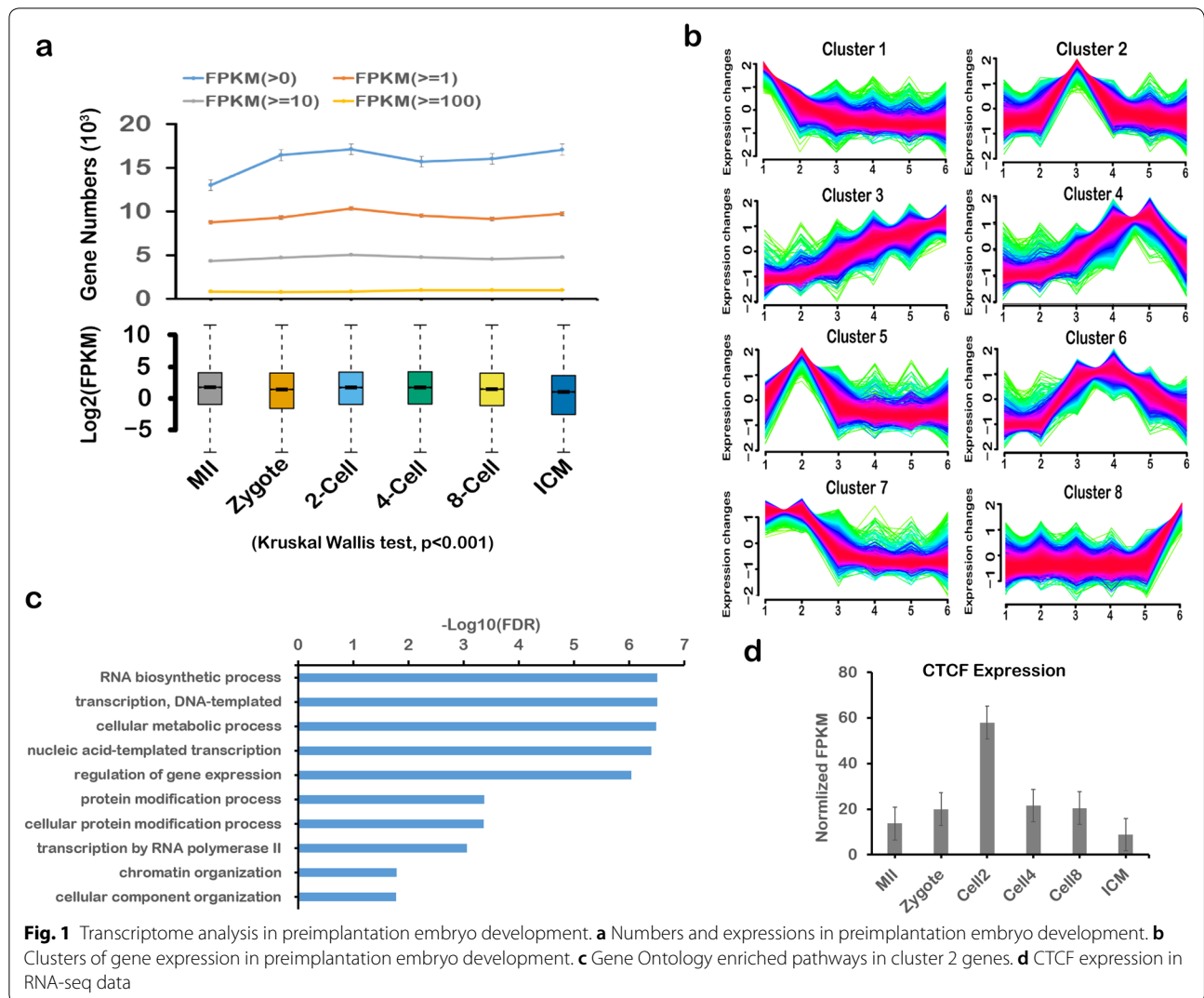
### Hi-C data analysis

Paired-end clean reads of Hi-C were processed as previously described using HiCPro (version 2.7) [11, 23]. Briefly, sequencing reads were aligned to the mouse reference genome (mm9) and removed the re-ligation reads, continuous reads and PCR artifacts. ICM [24] normalization was used to remove the bias in the Hi-C data. HiTC [25] was used to calculate the A/B compartment status in 400 kb resolution with the options: normPerExpected = TRUE, npc = 1, which positive value represented A compartment while negative value represented B compartment. TAD was called with the directional index as previously described [26] in 20 kb resolution. While TAD boundary was defined as the nearby region smaller than 400 kb between two TADs.

## Result

### The two-cell stage is a critical stage for gene transcription in preimplantation embryo development

AS is a process for enabling a messenger RNA (mRNA) to direct synthesis of different protein variants (isoforms) and plays a critical role in the regulation of gene expression. To study signatures of AS in preimplantation embryo development, we firstly identified the critical stage for gene transcription during the embryo development. Gene expression profiling of preimplantation embryo development was performed using RNA-seq data obtained from the study by Xie et al. [27], including MII oocytes, pronuclear stage 5 (PN5) zygotes, two-cell, four-cell, and eight-cell embryos, and inner cell masses (ICMs) from blastocysts. We respectively identified 13,026, 16,345, 17,117, 15,699, 16,012, and 17,080 expressed genes (FPKM  $> 0$ ) in these stages using the STAR software [14], followed by its two-pass mode (Fig. 1a). Meanwhile, according to the gene expression (FPKM), we identified genes differentially expressed in preimplantation embryo development (Kruskal–Wallis test,  $p < 0.001$ ) (Fig. 1a).



The results showed that genes more highly expressed in  $\text{FPKM} \geq 10$  did not change markedly, while genes expression at lower levels in  $\text{FPKM} < 10$  significantly fluctuated, with the two-cell stage having the highest number of expressed genes ( $\text{FPKM} > 1$ ). These findings suggest that the two-cell stage is a critical stage at which most genes with lower expression ( $\text{FPKM} < 10$ ) started to be transcribed during preimplantation embryo development.

To identify those genes specifically expressed in the two-cell stage, we used the fuzzy c-means (FCM) method to perform a time-course analysis of RNA-seq data from preimplantation embryo development. We divided the expressed genes into eight clusters and discovered 1415 genes in cluster 2, representing the genes specifically expressed in the two-cell stage (Fig. 1b). Gene Ontology enrichment showed that these genes were highly related to RNA transcription and gene expression, and

interestingly also to chromatin organization (Fig. 1c). In line with a previous report describing that the two-cell stage is the point at which the construction of topologically associating domains (TADs) is initiated in chromatin organization [11, 12, 26], we found that CTCF (CCCTC-binding factor) was also included in cluster 2, which is a factor that was reported with the formation of TAD and chromatin loops [28], and highly expressed in the two-cell stage (Fig. 1d). Taken together, these results show that the two-cell stage is critical stage for gene transcription during preimplantation embryo development.

#### Dynamic changes of alternative splicing in preimplantation embryo development

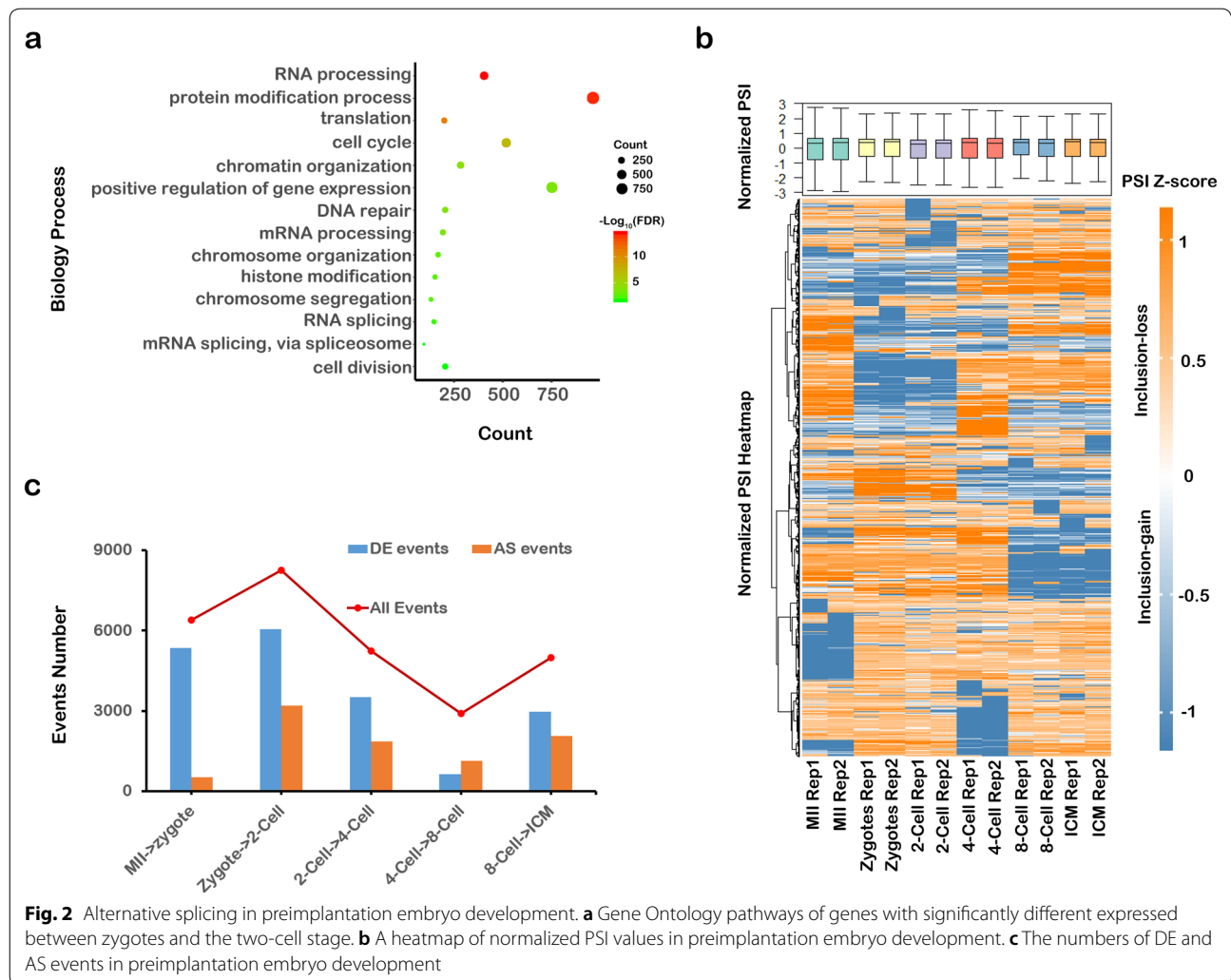
Next, we found that the genes with a significant change of expression in the two-cell stage compared with their expression in zygotes were particularly associated with

RNA splicing (Fig. 2a). It is thus suggested that RNA alternative slicing may play an important role in preimplantation embryo development. To characterize the AS events occurring during preimplantation embryo development, we used percent spliced-in (PSI) to measure the proportion of reads that aligned to splice junctions, which can support the inclusion isoforms [29]. Thousands of AS events were identified according to their PSI, and we divided the genes according to their PSI changes as inclusion-loss and inclusion-gain groups (Fig. 2b). The heatmap showed that PSI dynamically changed in preimplantation embryo development, suggesting that AS events are widespread in preimplantation embryo development. Furthermore, we screened the AS events in preimplantation embryo development and detected 6391 events (MII/zygotes), 8237 events (zygotes/two-cell stage), 5234 events (two-cell/four-cell stage), 2893 events (four-cell/eight-cell stage), and 4982 events (eight-cell stage/ICM). The differentially expressed genes numbered

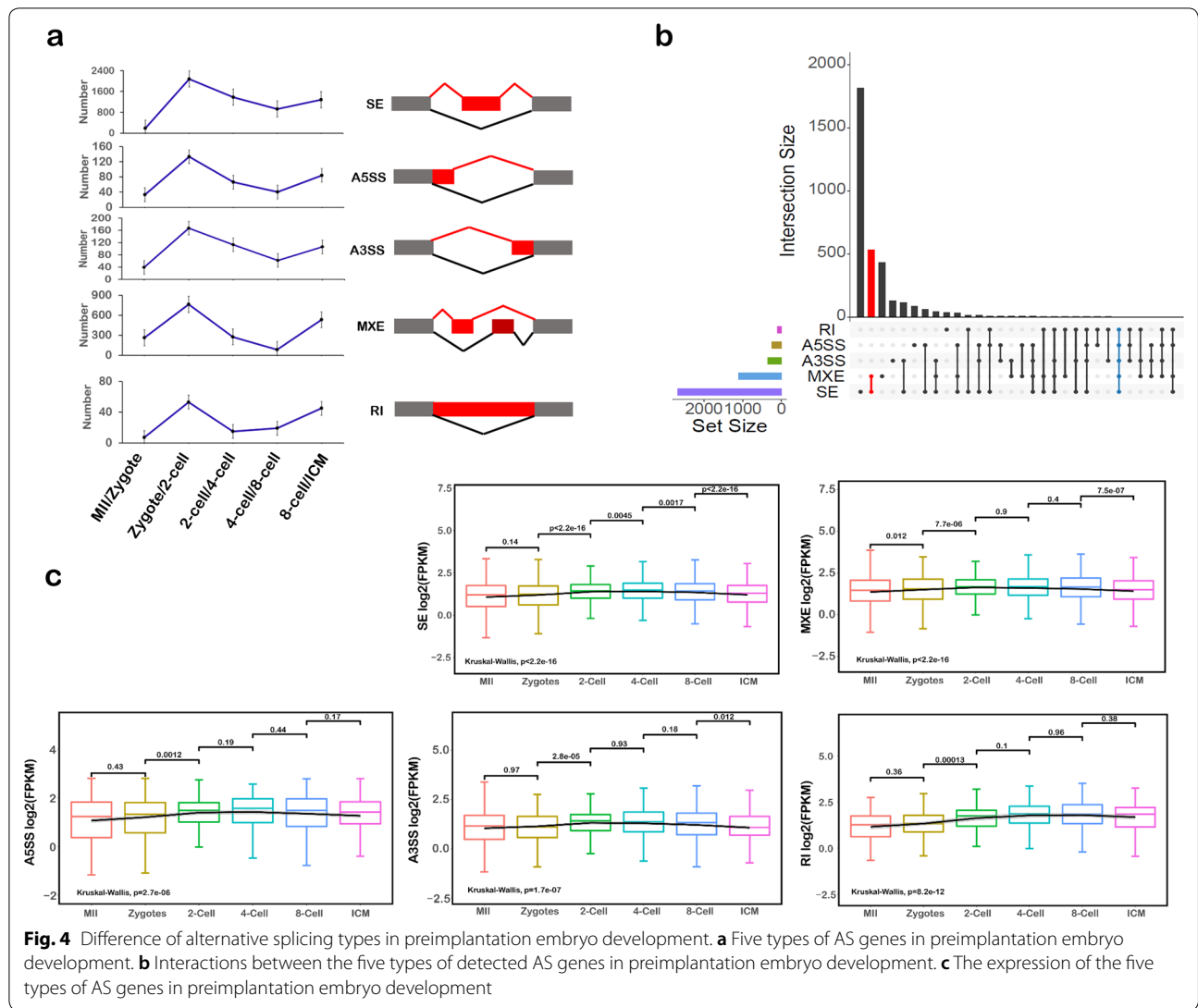
approximately 600 to 5000 depending on the comparison, while genes undergoing AS events numbered approximately 500 to 3000 (FDR<0.05) (Fig. 2c). For both variables, the two-cell stage showed peak values, suggesting that there is an abundance of RNA AS at the two-cell stage.

**Alternative splicing at the two-cell stage**

To further study AS events at the two-cell stage, we identified 695 genes that exhibited AS in the two-cell stage compared with their expression in zygotes using a Venn diagram (Fig. 3a). We then input those genes into String (version 11.0) to create a protein-protein interaction (PPI) network. According to betweenness centrality, a measure representing the frequency of the shortest path in the network, we divided our network into nine modules (Fig. 3b). Among them, we particularly focused on *Smarca1*, a sub-hub gene encoding an SWI/SNF-related matrix-associated actin-dependent







stage), 927 SE events (four-cell/eight-cell stage), and 1283 SE events (eight-cell stage/ICM), and 261 MXE events (MII/zygote), 764 MXE events (zygote/two-cell stage), 274 MXE events (two-cell/four-cell stage), 84 MXE events (four-cell/eight-cell stage), and 533 MXE events (eight-cell stage/ICM). The results show that both types of AS event were particularly abundant at the two-cell stage (Fig. 4a), supporting our previous conclusion that two-cell stage is important for AS events. Meanwhile, by overlap of each type of AS genes, we found two genes, the growth arrest-specific transcript 5 (*Gas5*) and *Psm3*, the former of which is a long noncoding RNA and the latter a coding gene (Fig. 4b). Interestingly, *Gas5* has more than 10 isoforms in the Ensembl database in mouse (visualized in UCSC genome browser). These many variants of *Gas5* suggest that this gene has a key role in preimplantation embryo development. Previous reports on

*Gas5* described that it is related to the self-renewal and pluripotency of stem cells [34, 35], and it can also act as a competing endogenous RNA (ceRNA) to exert its functions. Furthermore, we found that the expression of AS genes was significantly different in the two-cell stage compared with other stages in preimplantation embryo development (Fig. 4c), also supporting our previous findings.

#### Alternative splicing is linked to TAD boundaries at the two-cell stage

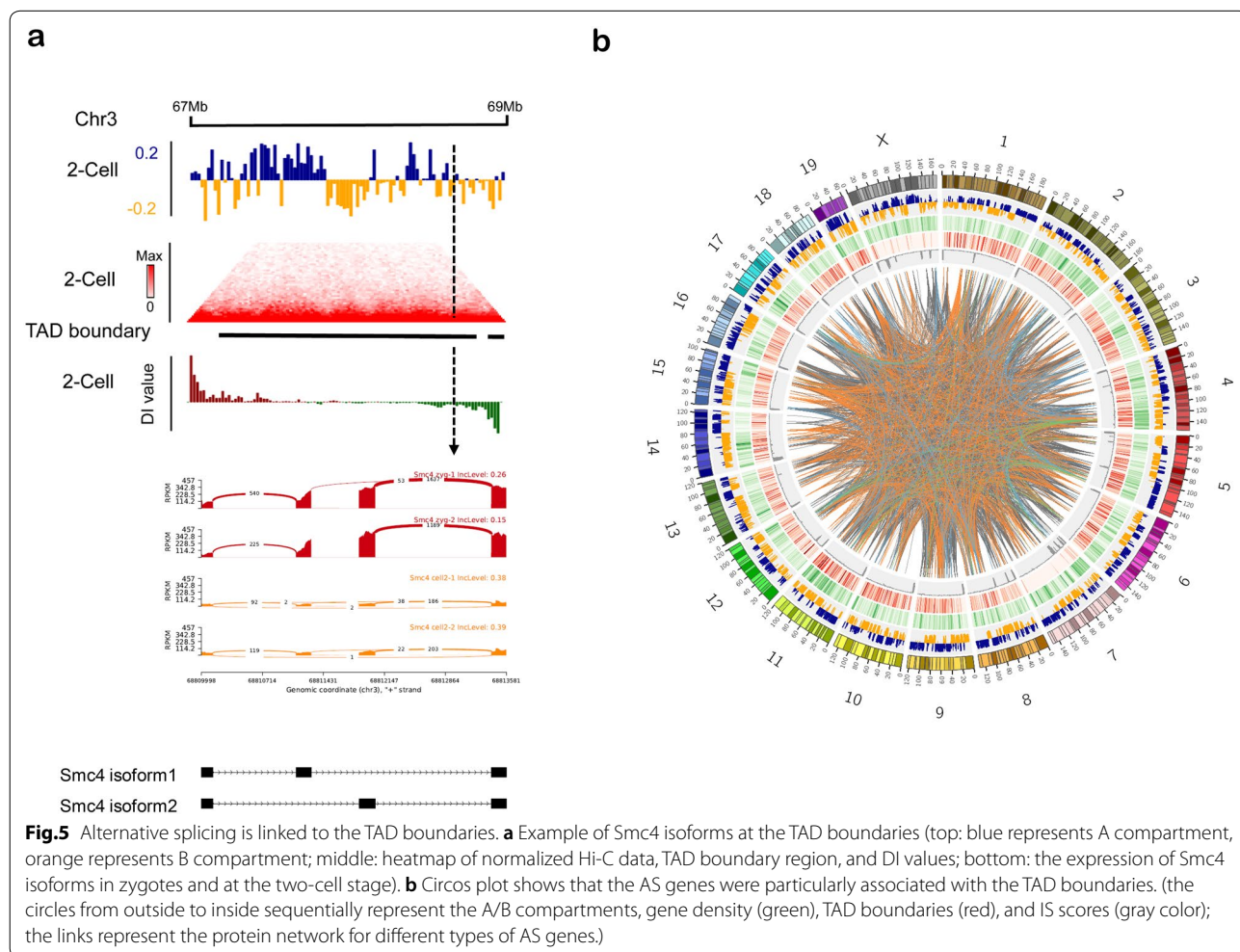
The previous transcriptome data showed that the two-cell stage is a pivotal step in preimplantation embryo development. Meanwhile, the chromatin structure was also shown to be reprogrammed starting at this stage, as it was reported that TADs started to form, were subsequently strengthened, and their construction was

completed at the eight-cell stage [11, 12]. To investigate whether the dynamic changes of chromatin organization were related to AS events, we analyzed the Hi-C data from the two-cell stage. We first measured the relationship between AS events and A/B compartments. By extracting the PC1 score from principal component analysis (PCA) [10], we found that AS events were not related to the A/B compartment status of chromatin structure (Fisher's exact test,  $p = 0.1164$ ). Then, we used the direction index (DI) method to call TADs at the two-cell stage [26]; Fisher's exact test ( $p = 0.4466$ ) showed that AS events were also not related to TADs. Lastly, we defined TAD boundaries as regions of less than 400 kb located between two TADs. We discovered 691 AS genes located in TAD boundaries, among the total of 3628 genes (Fisher's exact test,  $p < 2.2e^{-16}$ ); the results showed that an abundance of AS events occurred in TAD boundaries. Figure 5a shows an example of *Smc4* isoforms in the TAD boundaries (Fig. 5a). *Smc4*, structural maintenance of chromosomes (SMC) 4, belongs to the SMC gene family

and is an important subunit of the cohesin protein complex. This cohesion has been reported to be related to the higher-order chromatin structure [36]. This suggested that AS events may play an important role in chromatin organization. Furthermore, to confirm that AS events were particularly abundant in TAD boundaries, insulation score (IS) was calculated to measure the strength of TAD boundaries; a higher absolute IS score indicated a stronger TAD boundary [37]. We then analyzed the frequency of AS events at the genome-wide scale (Fig. 5b). The circos plot showed that AS events were enriched at the TAD boundaries, but were not related to A/B compartments. Taken together, these results showed a link between AS events and chromatin organization.

### Discussion

Previous research has reported the AS genes expression in embryonic day 8.5, 9.5 and 11.5 in mouse using microarray [38]. They revealed AS events are frequent across developmental stages and tissues, proposing AS



is an important developmental regulator [38]. However, it remains to be study the AS events in preimplantation embryo development because of technology limitation. With the development of scRNA-seq, the transcriptome profile analysis in preimplantation embryo development was successfully preformed [13, 39, 40]. However, they focused on the gene expression levels without distinguishing AS events during preimplantation embryo development.

AS plays a critical role in the regulation of gene expression and protein diversity in a variety of eukaryotes. In this study, we firstly identified the critical stage for gene transcription during the preimplantation embryo development to study signatures of AS in embryo development. By analyzing the time course of gene expression profile in preimplantation embryo development, we found the two-cell stage is a pivotal step in preimplantation embryo development because of the number of expressed genes. Then, we discovered that RNA splicing plays an important role in two-cell stage compared with zygote. Based on calculating PSI, we demonstrated that an abundance of AS events occurred in preimplantation embryonic development, with the two-cell stage having a high frequency of such events, the abundance of which then decreases at the four- and eight-cell stages, followed by a recovery at ICM. Our findings not only show that the two-cell stage is a pivotal step in preimplantation embryo development, but also indicate that AS events play an important role in the early embryo development.

One of the most important features in preimplantation embryo development is chromatin organization. In zygotes, the haploid genomes from paternal and maternal origins start to fuse, which greatly changes the chromatin structure. Previous reports indicated that organization of chromatin, as the structural and functional basis of the genome, can affect the location of DNA, playing important roles in gene transcription, preventing DNA damage, and ensuring DNA duplication and other biological processes [13, 41]. In this study, we found that AS events were particularly abundant at TAD boundaries. This may be because many transcription factors and histone modifications are related to TAD boundaries and the most important RNA polymerase, RNA Pol II, is also enriched in TAD boundaries [26]. At the TAD boundaries, we also identified that Smc4, a cohesin subunit protein, could form chromatin loops with CTCF to regulate the chromatin organization [42].

## Conclusions and perspective

In this study, we first showed that, within the stages of preimplantation embryo development, the two-cell stage had a higher number of expressed genes,

suggesting the importance of this stage within this process. Based on that, we identified a cluster of genes that were highly expressed at the two-cell stage, which may be related to chromatin organization. Then, we discovered AS events in this stage and divided them into five types: A5SS, A3SS, MXE, SE, RI; our data showed that AS events were abundant in preimplantation embryo development, especially at the two-cell stage. According to the profile of AS events in preimplantation embryo development, we found that most AS events were of the SE and MXE types, both of which frequently occurred at the two-cell stage. Lastly, we found that AS events were also related to chromatin structure, with them being particularly common in the TAD boundaries. Interestingly, some of the genes undergoing AS such as CTCF, Smc4, and Smarcb1 are all related to chromatin structure [43]. The findings provide insight into the relationships among gene expression, AS genes and high-order chromatin structure. These could form a self-regulatory loop between gene expression and chromatin organization, in which the transcription of some chromatin structure-related genes is initiated to promote the formation of TADs at the two-cell stage, and these well-constructed TADs then further promote the transcription of other cell development-related genes.

## Abbreviations

scRNA-seq: Single cells RNA-seq; AS: Alternative splicing; TADs: Topologically associating domains; PSI: Percent spliced-in; ICM: Inner cell masses; CTD: c-Terminal domain; SE: Skipped exon; A5SS: Alternative 5' splice site; A3SS: Alternative 3' splice site; MXE: Mutually exclusive exons; RI: Retained intron; ceRNA: Competing endogenous RNA; PCA: Principal component analysis.

## Acknowledgments

This work was supported by National key research and development program of China (2018YFC1003501), National Nature Science Foundation of China (81720108017, 31871329), and Shanghai Jiao Tong University Medicine-Engineering Fund (YG2018ZD12).

## Authors' contributions

GGT conceived, designed and analyzed the data; GGT wrote the paper. JW and JL supervised the experiment work and devised this study. All authors read and approved the final manuscript.

## Funding

This work was supported by National key research and development program of China (2018YFC1003501), National Nature Science Foundation of China (81720108017, 31871329), and Shanghai Jiao Tong University Medicine-Engineering Fund (YG2018ZD12).

## Availability of data and materials

The data and equipment used are presented in the manuscript.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

All authors agree to the publication of this manuscript.

## Competing interests

The authors declare that they have no competing interests.



**Author details**

<sup>1</sup> Renji Hospital, Key Laboratory for the Genetics of Developmental & Neuropsychiatric Disorders (Ministry of Education), Bio-X Institutes, School of Medicine, Shanghai Jiao Tong University, Shanghai 200240, China. <sup>2</sup> Key Laboratory of Fertility Preservation and Maintenance of Ministry of Education, Ningxia Medical University, Yinchuan 750004, China. <sup>3</sup> Department of Bioinformatics and Biostatistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China. <sup>4</sup> Shanghai Key Laboratory of Reproductive Medicine, Shanghai 200025, China.

Received: 6 February 2020 Accepted: 29 February 2020

Published online: 10 March 2020

**References**

- Hashimshony T, Wagner F, Sher N, Yanai I. CEL-seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*. 2012;2(3):666–73.
- Ramskold D, Luo S, Wang YC, Li R, Deng Q, Faridani OR, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*. 2012;30(8):777–82.
- Singh M, Al-Eryani G, Carswell S, Ferguson JM, Blackburn J, Barton K, et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat Commun*. 2019;10(1):3120.
- Zheng GX, Lau BT, Schnall-Levin M, Jarosz M, Bell JM, Hindson CM, et al. Haplotyping germline and cancer genomes with high-throughput linked-read sequencing. *Nat Biotechnol*. 2016;34(3):303.
- Kitzman JO. Haplotypes drop by drop. *Nat Biotechnol*. 2016;34(3):296.
- Guo F, Li L, Li J, Wu X, Hu B, Zhu P, et al. Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res*. 2017;27(8):967–88.
- Fan X, Zhang X, Wu X, Guo H, Hu Y, Tang F, et al. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol*. 2015;16:148.
- Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genetics*. 2008;40(12):1413.
- Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol*. 2017;18(7):437.
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289–93.
- Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, et al. Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature*. 2017;547(7662):232–5.
- Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, et al. 3D Chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell*. 2017;170(2):367–81 e20.
- Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature*. 2016;537(7621):553–7.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
- Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7(3):562–78.
- Kumar L, Futschik ME. Mfuzz: a software package for soft clustering of microarray data. *Bioinformatics*. 2007;21(1):5–7.
- Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res*. 2019;47(8):e47.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
- Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res*. 2012;22(10):2008–17.
- Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acid Res*. 2019;47(D1):D607–D613.
- Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. In: third international AAAI conference on weblogs and social media. 2009.
- Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47(D1):D419–D426.
- Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 2015;16:259.
- Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods*. 2012;9(10):999–1003.
- Servant N, Lajoie BR, Nora EP, Giorgetti L, Chen CJ, Heard E, et al. HiTC: exploration of high-throughput ‘C’ experiments. *Bioinformatics*. 2012;28(21):2843–4.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 2012;485(7398):376–80.
- Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y, et al. The landscape of accessible chromatin in mammalian preimplantation embryos. *Nature*. 2016;534(7609):652–7.
- Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebbersohn A, Abdennur N, et al. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*. 2017;169(5):930–44 e22.
- Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods*. 2010;7(12):1009–155.
- Valencia AM, Collings CK, Dao HT, Pierre RS, Cheng Y-C, Huang J, et al. Recurrent SMARCB1 mutations reveal a nucleosome acidic patch interaction site that potentiates mSWI/SNF complex chromatin remodeling. *Cell*. 2019;179(6):1342–56 e23.
- Sen P, Luo J, Hada A, Hailu SG, Dechassa ML, Persinger J, et al. Loss of Snf5 induces formation of an aberrant SWI/SNF complex. *Cell Rep*. 2017;18(9):2135–47.
- Pyeon D, Price L, Park IW. Comparative molecular genetic analysis of simian and human HIV-1 integrase interactor INI1/SMARCB1/SNF5. *Arch Virol*. 2015;160(12):3085–91.
- Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci USA*. 2014;111(51):E5593–5601.
- Tu J, Tian G, Cheung H-H, Wei W, Lee TL. Gas5 is an essential lncRNA regulator for self-renewal and pluripotency of mouse embryonic stem cells and induced pluripotent stem cells. *Stem Cell Res Ther*. 2018;9(1):71.
- Wang J, Gong X, Tian GG, Hou C, Zhu X, Pei X, et al. Long noncoding RNA growth arrest-specific 5 promotes proliferation and survival of female germline stem cells in vitro. *Gene*. 2018;653:14–21.
- Yang R, Kerschner JL, Gosalia N, Neems D, Gorsic LK, Safi A, et al. Differential contribution of cis-regulatory elements to higher order chromatin structure and expression of the CFTR locus. *Nucleic Acids Res*. 2016;44(7):3082–94.
- Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, et al. Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature*. 2015;523(7559):240–4.
- Revil T, Gaffney D, Dias C, Majewski J, Jerome-Majewska LA. Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics*. 2010;11(1):399.
- Xue Z, Huang K, Cai C, Cai L, Jiang C-Y, Feng Y, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature*. 2013;500(7464):593–7.
- Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566(7745):496–502.
- Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature*. 2015;518(7539):331–6.
- Racko D, Benedetti F, Dorier J, Stasiak A. Transcription-induced supercoiling as the driving force of chromatin loop extrusion during

formation of TADs in interphase chromosomes. *Nucleic Acids Res.* 2018;46(4):1648–60.

43. Arzate-Mejia RG, Recillas-Targa F, Corces VG. Developing in 3D: the role of CTCF in cell differentiation. *Development.* 2018;145(6):137729.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

