Biology of Reproduction, 2017, 97(1), 5–17
doi:10.1093/biolre/iox071
Research Article
Advance Access Publication Date: 4 July 2017

OXFORD

Research Article

Bovine lineage specification revealed by single-cell gene expression analysis from zygote to blastocyst†

Qingqing Wei, Liang Zhong, Shaopeng Zhang, Haiyuan Mu, Jinzhu Xiang, Liang Yue, Yunping Dai and Jianyong Han*

State Key Laboratory for Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing, China

***Correspondence:** State Key Laboratory for Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing 100083, China. Tel: +010-62734382; E-mail: hanjy@cau.edu.cn

†**Grant Support:** This work was supported by National Key Research and Development Program of China (2016YFA0100202), National Natural Science Foundation of China (31571497), Beijing Natural Science Foundation (6152004), and Chinese Universities Scientific Fund (2017QC163).

Received 27 April 2017; Revised 14 June 2017; Accepted 30 June 2017

Abstract

Preimplantation embryos undergo zygotic genome activation and lineage specification resulting in three distinct cell types in the late blastocyst. The molecular mechanisms underlying this progress are largely unknown in bovines. Here, we sought to analyze an extensive set of regulators at the single-cell level to define the events involved in the development of the bovine blastocyst. Using a quantitative microfluidics approach in single cells, we analyzed mRNA levels of 96 genes known to function in early embryonic development and maintenance of stem cell pluripotency in parallel in 384 individual cells from bovine preimplantation embryos. The developmental transitions can be distinguished by distinctive gene expression profiles and we identified *NOTCH1*, expressed in early developmental stages, while T-box 3 (*TBX3*) and fibroblast growth factor receptor 4 (*FGFR4*), expressed in late developmental stages. Three lineages can be segregated in bovine expanded blastocysts based on the expression patterns of lineage-specific genes such as disabled homolog 2 (*DAB2*), caudal type homeobox 2 (*CDX2*), ATPase H+/K+ transporting non-gastric alpha2 subunit (*ATP12A*), keratin 8 (*KRT8*), and transcription factor AP-2 alpha (*TFAP2A*) for trophectoderm; GATA binding protein 6 (*GATA6*) and goosecoid homeobox (*GSC*) for primitive endoderm; and Nanog homeobox (*NANOG*), teratocarcinoma-derived growth factor 1 (*TDGF1*), and PR/SET domain 14 (*PRDM14*) for epiblast. Moreover, some lineage-specific genes were coexpressed in blastomeres from the morula. The commitment to trophectoderm and inner cell mass lineages in bovines occurs later than in the mouse, and *KRT8* might be an earlier marker for bovine trophectoderm cells. We determined that *TDGF1* and *PRDM14* might play pivotal roles in the primitive endoderm and epiblast specification of bovine blastocysts. Our results shed light on early cell fate determination in bovine preimplantation embryos and offer theoretical support for deriving bovine embryonic stem cells.

Summary Sentence

Gene expression analysis of single blastomeres from zygote to blastocyst sheds light on the early cell fate determination in bovine preimplantation embryos and offers theoretical support for deriving bovine embryonic stem cells.

Key words: bovine, preimplantation embryo, single-cell gene expression, lineage specification.

Introduction

Mammalian preimplantation embryonic development (PED), from zygote to blastocyst, involves a tightly regulated series of events comprising zygotic genome activation (ZGA) and lineage specification. ZGA that occurs after fertilization is the first major transition in which maternal transcripts are specifically degraded and are replaced by zygotic transcripts produced by the new diploid nucleus containing both maternal and paternal genes. It is essential for continued progression of embryonic development, but its onset varies across species occurring in mice at the 2-cell stage [1], in pigs at the 4-cell stage [2, 3], and in bovines and humans at the 8-cell stage [4, 5].

Another major event during PED is lineage specification. Following ZGA, embryos undergo two consecutive lineage segregations resulting in three different types of cells in the late blastocyst (LB), namely, the trophectoderm (TE), the primitive endoderm (PE), and the epiblast (EPI). The first lineage specification occurs at the morula stage, when outer cells segregate from the inner cells and differentiate into the TE. The second lineage specification occurs when the inner cell mass (ICM) differentiates into the PE and EPI at the blastocyst stage. Finally, the TE gives rise to extraembryonic tissues, whereas the PE and EPI subsequently develop into the extraembryonic yolk sac and the embryo proper, respectively.

While the progress of PED is highly conserved among mammals and characterized by the same morphologic stages, bovine and mouse embryos show marked differences in the amount of time spent at each stage and gene expression and function. For example, bovine POU class 5 homeobox 1 (*POU5F1*) mRNA and protein persist for some time in both the TE and PE before gradually becoming EPI exclusive after hatching [6–9]. The bovine *NANOG* protein first appears in a subset of ICM cells before becoming EPI restricted and mutually exclusive in LBs [10]. The *CDX2*-deficient embryos failed to form blastocoels, hatch and implant in the mouse, whereas knockdown of *CDX2* in cattle resulted in normal embryonic development and blastocyst formation, and the deficient embryos were even able to develop for 15 days after transfer into recipient cows [11–14]. Therefore, it is urgently necessary to study the unique gene expression pattern in bovine preimplantation embryos to better understand the mechanisms of embryonic development specific to bovines.

Pluripotent embryonic stem cells (ESCs) are isolated from the ICM of blastocysts. ESCs can be able to differentiate in vitro and in vivo into all three germ layers, including germ cells [15, 16]. Therefore, they have great potential for the regenerative medicine and animal breeding. Over the past several decades, stable ESCs were successfully established from mice, rhesus monkeys, humans, and rats [17–20]. Many efforts have been made to derive bovine embryonic stem cells (bESCs) to facilitate precise genetic manipulation [21, 22]. However, no study has reported the generation of pluripotent bESCs capable of germ-line transmission, mostly because culture conditions were unsuitable to maintain pluripotency, and there is inadequate information regarding the unique molecular mechanisms of bovine early embryonic development [23]. This emphasizes the need to more comprehensively define the molecular mechanism of early embryonic development in bovines.

With the development of transcriptomic analyses, such as microarrays and next-generation sequencing, numerous studies regarding the transcriptomic dynamics of preimplantation bovine embryos have emerged [24–27]. However, all measurements were performed on populations of cells, which limits in-depth analysis of cell differentiation and the regulatory function underlying the transcriptome. Deciphering gene expression in embryos at the single-cell level is a crucial step towards understanding early developmental processes. Therefore, we sought to analyze an extensive set of regulators in combination at the single-cell level to better define the events involved in the development of the bovine blastocyst. Considering the sample size, assayed gene number, and sensitivity in detecting quantitative differences at the mRNA level, high-throughput single-cell quantitative PCR (qPCR) represents a favorable option [28–30]. We analyzed mRNA levels of 96 genes known to drive cellular fate or function in early embryonic development and the maintenance of stem cell pluripotency in parallel in 384 individual cells from the zygote through to the expanded LB, thus providing unprecedented insight into the earliest cell fate decisions of the developing bovine embryo.

Materials and methods

Bovine in vitro fertilized embryo collection and culture

Cumulus oocyte complexes were retrieved from 2–8-mm follicles of bovine ovaries, which were collected at a local abattoir. Oocytes with an intact cumulus oophorus were selected and matured in M199 medium (Gibco, Paisley, UK), supplemented with 10 μg/ml follicle-stimulating hormone (FSH), 1 μg/ml luteinizing hormone (LH), 1 μg/ml estradiol (E_2), 10% (v/v) fatal bovine serum (FBS) (Sigma-Aldrich, MO, USA), and 1% (v/v) penicillin-streptomycin (Gibco, Paisley, UK) for 24 h. Maturated oocytes were transferred to fertilization medium supplemented with 1.8 IU/ml heparin, 20 μM D-penicillamine, 10 μM hypotaurine, and 1 μM epinephrine (all Sigma-Aldrich, MO, USA). Subsequently, frozen-thawed semen from a fertile bull was centrifuged over a discontinued 45%–95% percoll gradient, and sperm were added to the matured oocytes at a final concentration of 1×10^6/ml. After 10 h of incubation, putative zygotes were mechanically denuded by pipetting in TL-Hepes medium and then cultured for the appropriate time in synthetic oviduct fluid medium under a 5% CO_2 atmosphere at 38.5°C.

Individual blastomeres isolation

Embryos at different developmental stages were collected at the following times postfertilization: zygote (1C) at 8 h, 2-cell (2C) on day 1, 4-cell (4C) on day 2, 8-cell (8C) on day 3, 16-cell (16C) on day 4, 32-cell (32C) on day 5, early blastocyst (EB) on day 6, and LB on day 8. The zona pellucida of all embryos was removed by treatment with 0.5% pronase (Sigma-Aldrich, MO, USA). Two- to eight-cell denuded embryos were incubated in M2 medium (Sigma-Aldrich, MO, USA) without calcium and magnesium and disaggregated by repeated mouth pipetting. The denuded morula and blastocyst were incubated with 0.25% trypsin (Invitro-

gen, Carlsbad, USA) for 5–10 min at 37°C and transferred to M2 medium without calcium and magnesium. Single blastomeres were isolated by gentle, repeated mouth pipetting with a finely pulled glass tip. After disaggregation, all of the blastomeres were removed from the manipulation drops, washed three times with Dulbecco phosphate-buffered saline (Invitrogen, Carlsbad, USA) containing 0.1% bovine serum albumin (Sigma-Aldrich, MO, USA), and placed into reverse transcription-polymerase chain reaction (RT-PCR) master mix for lysis, sequence-specific reverse transcription, and preamplification.

Multiplexed primer design for single-cell analysis

Messenger RNA sequences for each selected gene were retrieved from National Center for Biotechnology Information, and only the common regions were used for genes with different transcripts. Gene-specific primers were designed using Primer3 and ensuring that each primer within the designed pool had a similar melting temperature (Tm) and a maximum complimentary sequence of 7 bp to all the other primers. All primers, previously tested using cDNA of bovine blastocysts for amplification efficiency and specificity, are listed in Supplementary Table S1.

One-tube single-cell sequence-specific preamplification

A total of 96 primer pairs were pooled to a final concentration of 100 nM for each primer. Individual cells isolated from embryos were transferred to microtubes with 5 μl of RT-PCR master mix containing 2.5 μl of CellsDirect reaction mix (Invitrogen, Carlsbad, USA), 0.5 μl of primer pool, 0.1 μl of RT/Taq enzyme (Invitrogen, Carlsbad, USA), and 1.9 μl of nuclease free water in each well. Tubes were immediately frozen on dry ice. After brief centrifugation at 4°C, tubes were immediately placed in a PCR machine to perform cell lyses and sequence-specific reverse transcriptions at 50°C for 1 h. Afterwards, reverse transcriptase inactivation and Taq polymerase activation were achieved by heating to 95°C for 3 min. Subsequently, in the same tube, cDNA went through 20 cycles of sequence-specific amplification by denaturing at 95°C for 15 s, and annealing and elongation at 60°C for 15 min. To avoid evaporation, the resulting products were stored at –80°C.

High-throughput microfluidic single-cell quantitative PCR

The preamplified cDNA products were diluted 5-fold prior to analysis with SsoFast EvaGreen Supermix with Low ROX (Bio-Rad, California, USA) and individual qPCR primers in 96:96 dynamic arrays on a Biomark System (Fluidigm, San Francisco, USA). Threshold crossing (Ct) values were calculated from the system's software (BioMark Real-time PCR Analysis).

Processing and visualization of single-cell data

All raw Ct values obtained from the BioMark System were converted into log2 gene relative expression levels by subtracting the assumed background Ct of 24. Samples with low or absent expression levels of endogenous control genes, suggesting that they correspond to empty wells or bad single cells, were excluded from subsequent analysis.

Principal component analysis (PCA) was performed using the svd command in R. Hierarchical clustering was performed using Euclidean distances, and dendrograms were displayed along the row-scaled heatmaps using the fluidigmSC package. Boxplots and violin plots were generated with the originLab and R software. To simultaneously visualize the gene expression levels, the star command in R was used. In this analysis, the expression levels were converted to a gene-specific scale where the lowest expression level in the data set corresponds to 0 and the highest expression level corresponds to 1. Thus, the radii of the pies in Figures 5B and 6A can be compared across cells at the same time. For constructing Figures 2D, 3D, 4D, 4G, 5B, 5D, 6A and 6B, and Supplementary Figure S2, we used PCA to clarify cells into corresponding cell types. For the LB, the cells with a PC1 score below –10 were classified as TE, cells with a PC1 score $\geq$0 and a PC3 score $\leq$0 were classified as PE, and cells with a PC1 score $\geq$0 and a PC3 score $\geq$10 were classified as EPI. For the EB, cells with a PC1 score $\leq$–10 were classified as TE and cells with a PC1 score $\geq$0 were classified as ICM. The remainder cells fulfilling none of these criteria were considered as uncertain.

Results

We used in vitro fertilized bovine embryos to perform a quantitative analysis of gene expression patterns in single cells using the Fluidigm Biomark System with 96:96 Dynamic Array chips coupled with gene-specific primers. The use of EvaGreen real-time PCR chemistry and melting curve analysis allows for monitoring specific signals. Single blastomeres were isolated by manual disaggregation of bovine embryos harvested at the appropriate times on the basis of morphology (Figure 1A and 1B). We analyzed a total of 384 single cells obtained at all preimplantation developmental stages from zygote to expanded blastocyst (Table 1). We focused on genes known to drive cellular fate or function in early embryonic development and maintenance of stem cell pluripotency in the mouse, bovine, and human. In general, the final 96 genes (Supplementary Table S1), involved in cell differentiation, chromatin modification, signal transduction, metabolism, and pluripotency, were selected for their utility in detecting embryonic single-cell gene expression. We used actin beta (*ACTB*) and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) as control genes. The absolute expression levels of all genes at all stages analyzed are shown in Supplementary Figure S2.

Developmental transitions are defined by gene expression profiles

To determine whether different developmental stages and degrees of cell potency can be defined on the basis of the gene expression profiles of these genes, we first performed unsupervised hierarchical clustering of all blastomeres and all genes (Figure 2A). This revealed that cells cluster primarily according to their developmental stage of origin. Interestingly, some blastomeres of neighboring stages clustered together, indicating that the transcriptional profile of these two stages are still relatively similar. Cells of the morula and blastocyst stages were clearly distinguished from the earlier stages of development, indicating that lineage segregation first occurs at the morula stage. In addition, the zygote, 2-cell- and partial 4-cell-stage cells clustered together, which can be explained by a major maternal–zygotic transition (MZT).

To better investigate the predominant gene expression patterns, we applied PCA to our data set. This mathematical analysis transforms the data to a new coordinate system that reduces as much of the variance as possible into single dimensions, depicted on axes of a graph, highlighting similarities and differences within the data set. For the PCA shown in Figure 2B, the data points are single cells

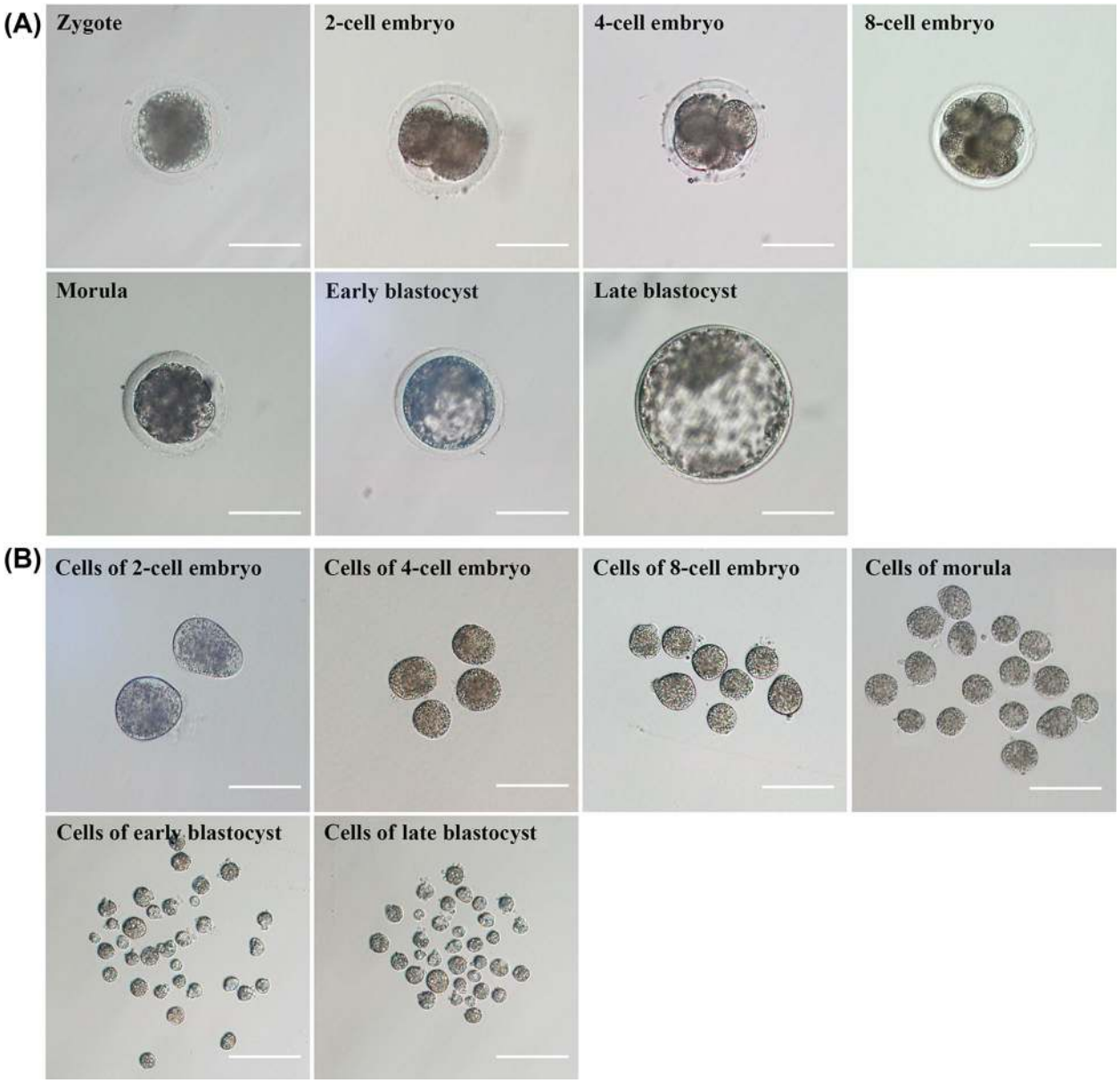

Figure 1. Morphology of bovine preimplantation embryos. (A) Microscopy imaging of bovine preimplantation embryos at zygote, 2-cell, 4-cell, 8-cell, morula, EB, and LB stages. (B) Microscopy imaging of corresponding isolated single blastomeres from embryos at different stages. Note the full recovery of all of the individual blastomeres of 2- and 4-cell embryos. Scale bar, 100 μm.

Table 1. The number of embryos and individual cells analyzed.

384 single cells across all stages		
Developmental stage	Number of embryos	Number of cells
1-cell	11	11
2-cell	8	15
4-cell	7	28
8-cell	7	42
16-cell (early morula)	2	19
~32-cell (compact morula)	3	77
EB	5	96
Expanded LB	6	96

from all preimplantation developmental stages, and the variables are the 96-dimensional gene sets. In the PCA of our data set, the first component (PC1) is represented on the x-axis and explains 42.28% of the observed variance, while the second component (PC2) explains 15.25% of the remaining variance on the y-axis. The PCA revealed that all blastomeres distribute according to their developmental origin. The most significant variation in the data set was due to differences between early and later stages of development, with the 8-cell stage marking this transition (Figure 2B), reflecting the MZT. Notably, a small portion of 8-cell-stage blastomeres clustered together with earlier stages, suggesting no completed MZT in these cells. In contrast, cells of the later developmental stages from 16-cell to LB were not separated by PC1, but by PC2 on the y-axis.

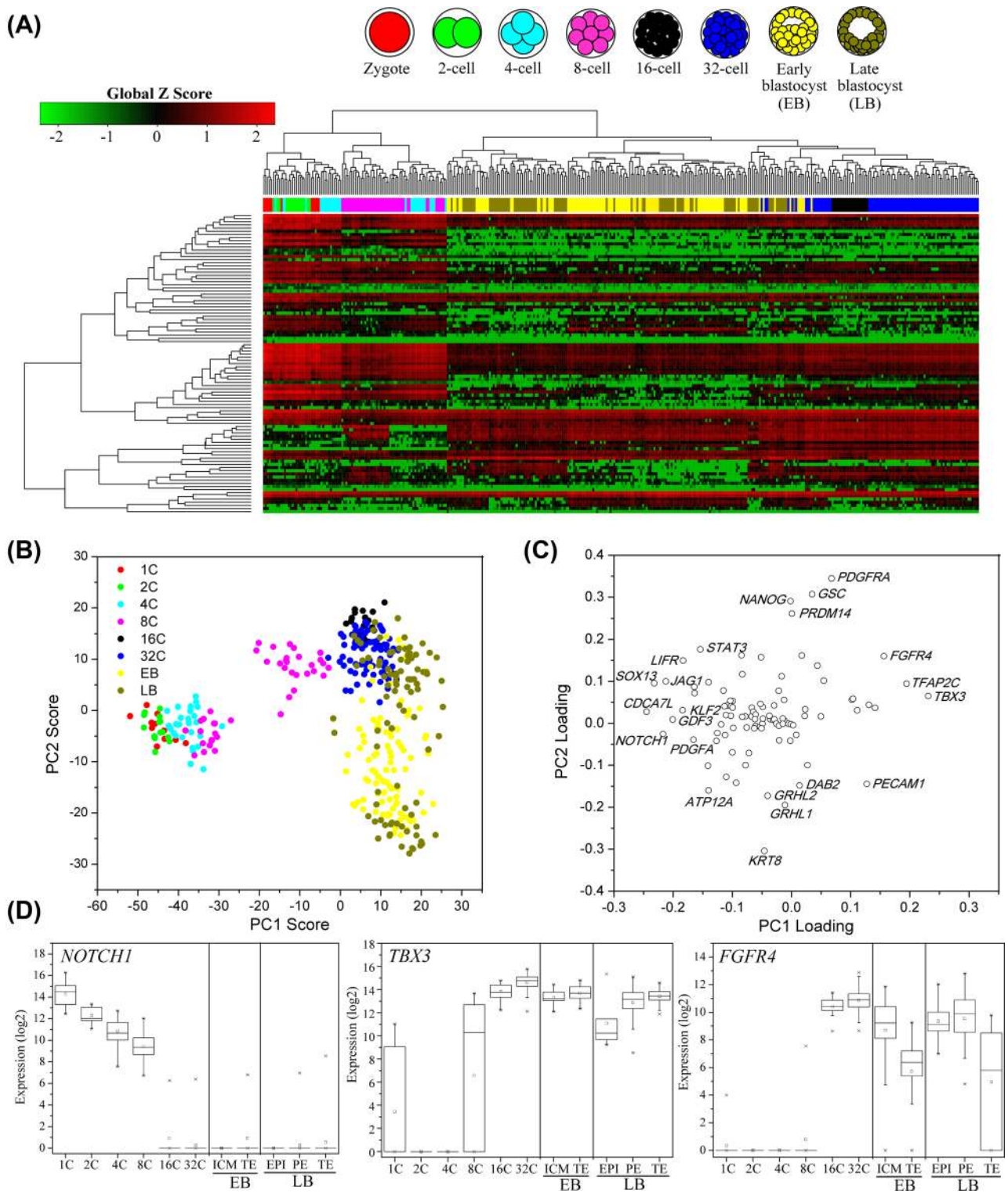

Figure 2. Developmental transitions are defined by gene expression profiles. (A) Hierarchical clustering of single cells derived from embryos at all stages, from zygote to LB, based on the expression levels of 96 genes with combined hierarchical clustering of genes. Each stage is colored according to the scheme above throughout the figures. (B) Principal component (PC) projection of individual cells based on the gene expression profiles. The cells are colored according to embryonic stages as noted. The first component represents the developmental transition from zygote to the 16-cell stage on the x-axis, and the second component represents ICM-to-TE lineage segregation on the y-axis. (C) PC projections of all genes, showing the contribution of each gene to the first two PCs in Figure 2B. For example, a gene with a more positive PC1 loading is more enriched in cells with more positive PC1 scores in Figure 2B. (D) Absolute expression levels of *NOTCH1*, *TBX3*, and *FGFR4* across developmental stages and predicted cell types (early blastocyst (EB) and late blastocyst (LB)). The boxed region represents the middle 50% of expression values, the black bar indicates the median values, and the whiskers indicate the maximum and minimum values. Cells with outlying expression values are depicted as asterisks. A background of Ct = 24 was used to obtain an absolute expression level.

As shown in Figure 2C, the PCA components consisted of contributions from all 96 genes in the data set, and genes with a more positive PC score enriched in cells corresponding to a more positive PC loading. This enabled us to identify genes that are enriched in the early and late developmental stages, namely, cell division cycle associated 7 like (*CDCA7L*), SRY-box 13 (*SOX13*), *NOTCH1*, and jagged 1 (*JAG1*) in the early stages and *TBX3*, transcription factor AP-2 gamma (*TFAP2C*), and *FGFR4* in the late stages. We found that *NOTCH1* is a maternal gene that is highly expressed in the zygote and downregulated up to the 16-cell stage. Furthermore, zygotic expression of *NOTCH1* is not activated at the following preimplantation developmental stages (Figure 2D). *TBX3* and *FGFR4* lie on the extreme right hand of PC1, suggesting that zygotic *TBX3* and *FGFR4* accumulation occurs at later preimplantation developmental stages (Figure 2C).

Three lineages can be segregated in the late blastocyst

The LB consists of three lineages with different development potency, namely, TE, EPI, and PE. We thus asked whether these three different cell types could be identified based on our data sets. First, we performed unsupervised hierarchical clustering and PCA of individual blastomeres from LBs. Hierarchical clustering of the gene expression profiles from 96 single cells obtained from LBs revealed that three cell types exist at this stage (Figure 3A). Twenty-eight cells highly expressed known TE markers such as *CDX2*. Six cells were enriched in EPI-specific genes such as *NANOG*. However, the known PE-restricted marker *GATA6* was simultaneously expressed in TE and PE cells. We thus clarified those cells expressing *GATA6* but not expressing *CDX2* as PE cells. Interestingly, a small set of seven cells obtained from the LB appeared to not closely align with any cell type, indicating that these cells were still undergoing differentiation.

When we applied PCA to the expression data of LBs, we found that PC1 explains 28.3% of the observed variance, while PC2 and PC3 explain 8.14% and 5.1%, respectively. A projection of the data set onto PC1, PC2, and PC3 segregated individual cells into three well-defined clusters (Figure 3B). The TE cluster could be separated from the ICM (containing EPI and PE) by the PC1 on the x-axis, whereas the other two cell types of the ICM could be distinguished from each other by PC2 and PC3 on the y- and z-axes, respectively. We next sought to identify the most contributive genes to clarifying these three cell types in the LB. As shown in Figure 3C and 3D and Supplementary Figure S2, *DAB2*, *CDX2*, *ATP12A*, *KRT8,* and *TFAP2A* are the most specific markers of the TE, whereas, *TDGF1* and *NANOG* are enriched in the EPI. Notably, *TDGF1* adopts the most extreme distribution of the genes analyzed, closer to the top than *NANOG* reflecting it might be a more reliable marker of the EPI in bovines (Figure 3C). As reported in bovines, GATA6 protein is detectable in both PE and TE cells [10]. Supporting this, our results demonstrate that *GATA6* mRNA is not expressed in the EPI but is highly expressed in both PE and TE lineages. Similarly, the other two PE-specific markers in mouse, platelet derived growth factor receptor alpha (*PDGFRA*) and hepatocyte nuclear factor 4 alpha (*HNF4A*), are enriched in the ICM, but are equally expressed in PE and EPI lineages (Supplementary Figure S2). Notably, *GSC*, which is expressed specially in the mouse primitive streak, is enriched in the ICM and PE but is also highly expressed in the EPI lineage (Figure 3D).

The two lineages can be segregated but not completely resolved in the early blastocyst

The EB consists of two lineages with distinct developmental potency, namely, the outer cells TE and inner pluripotent cells ICM. Subsequently, the ICM gives rise to the PE and EPI in the LB. We thus asked whether these two different cell types could be identified based on the expression pattern of gene sets. As described above, we first applied unsupervised hierarchical clustering and PCA to our data of individual cells derived from the EB. Unsupervised hierarchical clustering segregates the EB single cells into two clusters, the TE and ICM (Supplementary Figure S1). The gene contribution plot indicates that the TE cells highly express TE markers, including *KRT8*, *ATP12A*, msh homeobox 2 (*MSX2*), *DAB2*, *TFAP2A*, and *CDX2*, while the ICM cells are enriched in *GSC*, *PDGFRA*, *HNF4A*, signal transducer and activator of transcription 3 (*STAT3*), runt related transcription factor 1 (*RUNX1*), *PRDM14*, leukemia inhibitory factor receptor alpha (*LIFR*), *FGFR4*, and *NANOG* (Figure 4B). Nevertheless, cells do not present as two obvious distinct populations in the PCA score plot (Figure 4A).

To further illustrate the identity of these cells, we projected the gene expression patterns of individual cells from the EB onto the first three PCs calculated for the LB (Figure 4C). This analysis revealed that some single cells from the EB are either TE like or PE like, but several cells resemble no defined cell type, suggesting that the first lineage segregation is still not resolved at this stage. Moreover, the ICM transcription factor *NANOG* was not so obviously enriched in the ICM cells at this stage (Figure 3D), supporting the view in the mouse that the TE cells are defined earlier than the ICM and the markers for the ICM are not completely segregated until one or two cycles later [31].

To determine whether these genes enriched in each lineages show differential expression levels across single cells of the EB, we analyzed our data with violin plots (Figure 4E). As expected, the endogenous control gene *ACTB* shows a unimodal expression. The TE-enriched genes, such as *ATP12A* and *TFAP2A*, clearly show bimodal distributions, reflecting high expression levels in TE cells and low or no expression in ICM cells. The ICM-enriched genes, such as *GSC*, *HNF4A*, *PDGFRA*, *RUNX1*, *NANOG,* and *PRDM14*, also show bimodal expression patterns, fitting with the extreme distribution of the genes in PCA shown in Figure 4B.

Chromatin modifiers crucial for bovine embryonic development

Chromatin modification plays an important role in regulating embryo development. Therefore, we next wondered whether we could identify chromatin modifiers that might play an instructive role in bovine lineage specification. Figures 3C and 4B show the chromatin modifiers that make the greatest contribution to the segregation of three cell types. Notably, among the 13 genes relating to chromatin modification we detected, *PRDM14* is the most enriched gene in the ICM and EPI cells. This result is consistent with a previous report in mice [32]. However, we detected no obvious specific chromatin modifiers in bovine TE or PE cells in contrast to a report describing DNA methyltransferase 3 beta (*DNMT3B*) as the most enriched chromatin modifier in mouse TE cells. Ordering individual cells according to their PC1 score and plotting the expression levels of the selected genes accordingly further confirmed that the expression of *PRDM14* is higher in ICM cells (Figure 4D). Its bimodal distribution in the EB also proved this expression pattern (Figure 4E). Interestingly, the violin plot shows that *PRDM14* is also heterogeneously

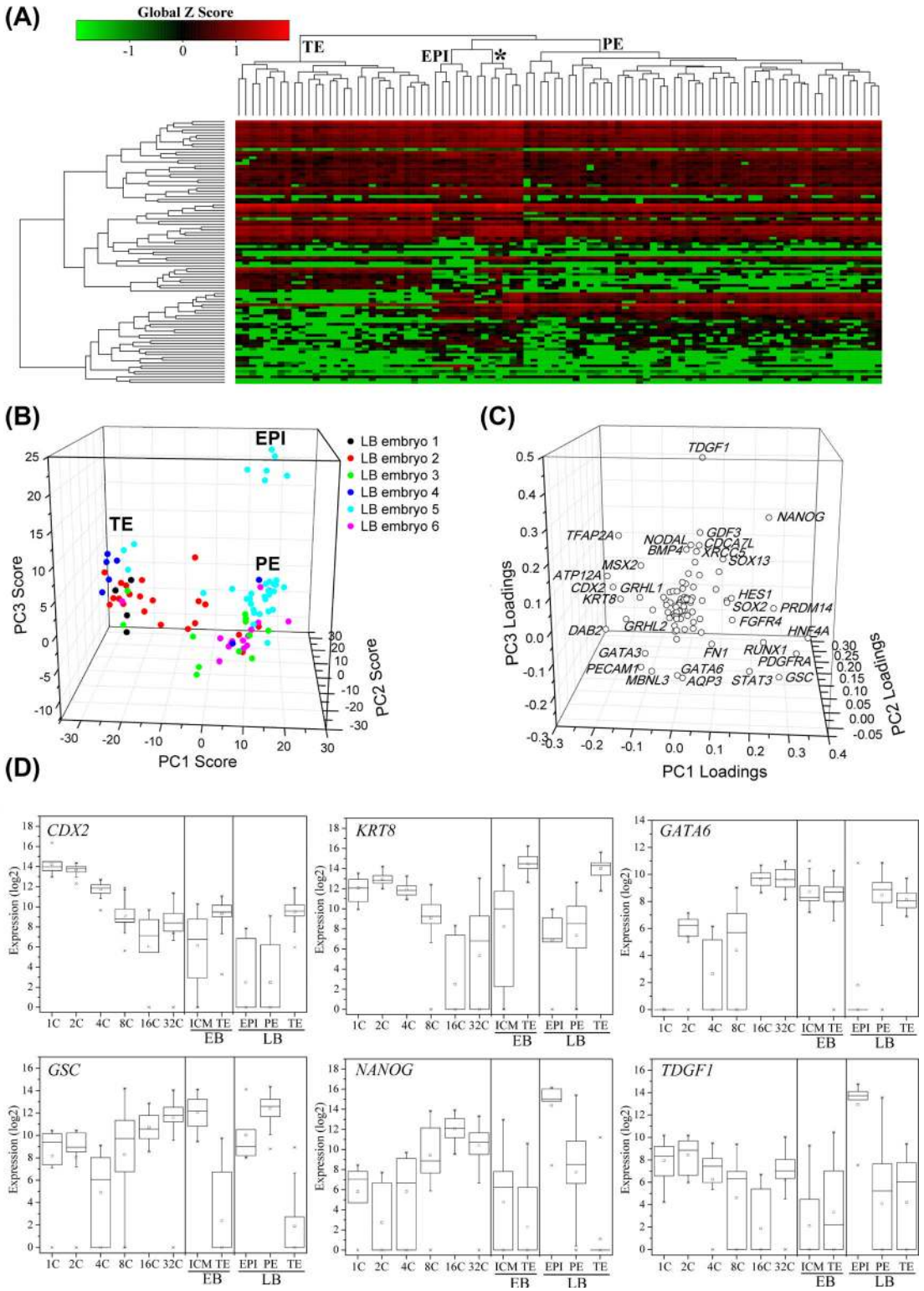

Figure 3. Three cell types identified in the LB. (A) Hierarchical clustering of individual cells collected from late blastocysts. Cells are defined as trophectoderm (TE), epiblast (EPI), and primitive endoderm (PE) based on their expression levels of known markers *CDX2*, *NANOG*, and *GATA6*, respectively. The asterisk (*) marks seven transitional cells with TE and ICM expression characteristics. (B) Principal component (PC) projections of LB single cells. (C) PC projections of all genes, showing the contribution of each gene to the first three PCs. The first PC can be interpreted as discriminating between ICM and TE. The third PC can be interpreted as discriminating between EPI and PE. (D) Absolute expression levels of six lineage-enriched genes across developmental stages and predicted cell types (EB and LB). The boxed region represents the middle 50% of expression values, the black bar indicates the median values, and the whiskers indicate the maximum and minimum values. Cells with outlying expression values are depicted as asterisks. A background of Ct = 24 was used to obtain an absolute expression level.

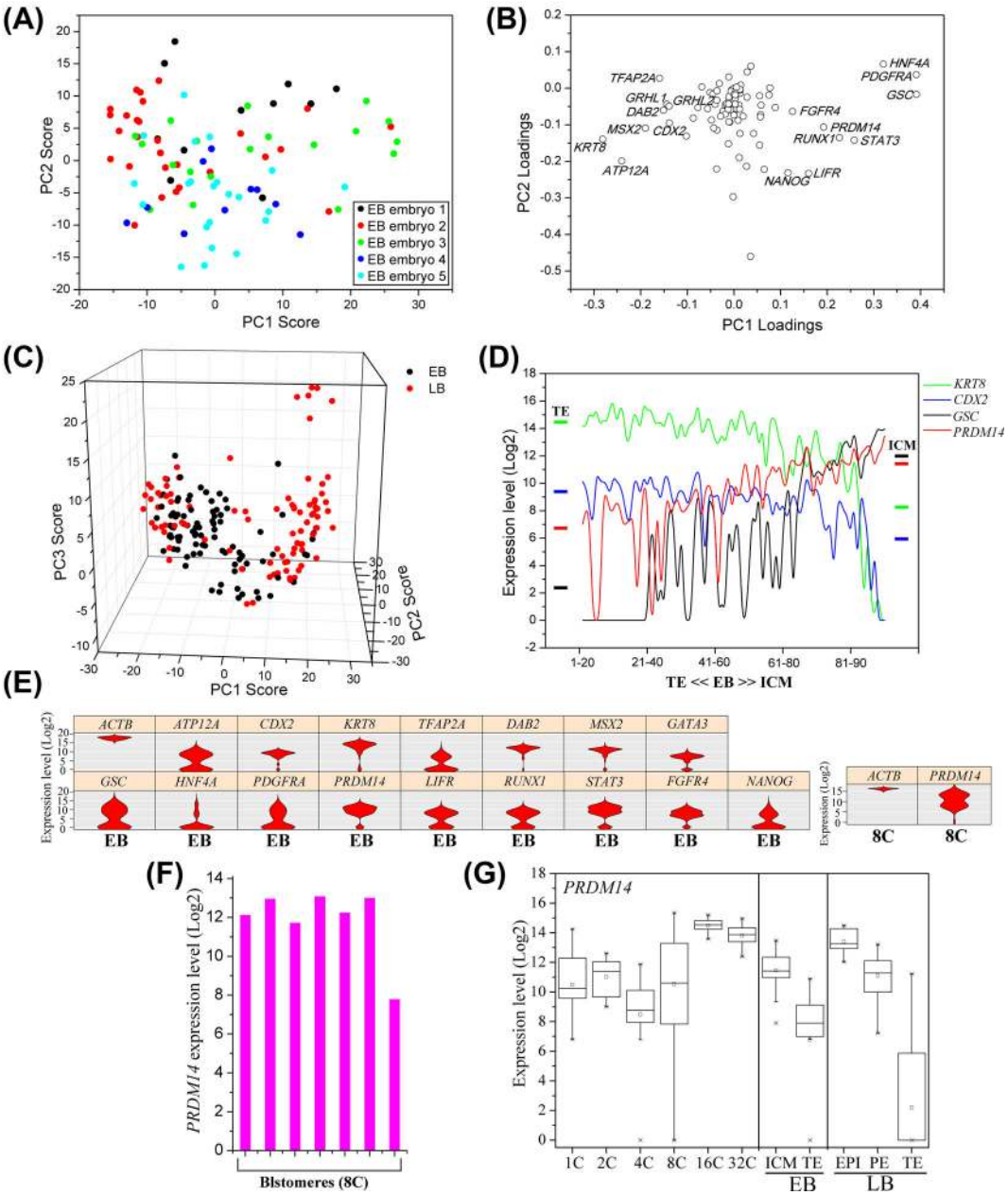

Figure 4. Developmental progression to two distinct cell types in the early blastocyst. (A) PC projections of the single cells from early blastocysts. (B) PC projections of all genes, showing the contribution of each gene to the first two PCs in Figure 4A. (C) Position, based on the expression of 96 genes, of individual cells from early blastocyst projected onto the first three PCs of the LB stage data. (D) Expression profiles of *KRT8, CDX2, GSC,* and *PRDM14* across the population of single cells in the early blastocyst. Blastomeres were ordered along the x-axis according to the PC1 score in (C). The traces represent the average expression value of each gene in moving windows of 20 cells. The colored bars labeled "TE" and "ICM" represent the mean expression levels at the blastocyst stage for TE and ICM, respectively. (E) Violin plots representing expression levels of selected genes (fold change above background) in individual cells at the early blastocyst and 8-cell stages. As expected, the lineage-specific markers show bimodal expression profiles at the early blastocyst. *PRDM14* also shows bimodal expression at the 8-cell stage. (F) Bars represent the absolute *PRDM14* expression level in each cell of a representative 8-cell stage embryo. (G) Absolute expression levels of *PRDM14* across developmental stages and predicted cell types (EB and LB). The boxed region represents the middle 50% of expression values, the black bar indicates the median values, and the whiskers indicate the maximum and minimum values. Cells with outlying expression values are depicted as asterisks. A background of Ct = 24 was used to obtain an absolute expression level.

expressed in 8-cell-stage embryos when zygotic activation occurs (Figure 4E). Moreover, analysis of individual embryos revealed that this heterogeneous expression pattern is also intraembryonic at the 8-cell stage (Figure 4F). Additionally, *PRDM14* is initially expressed at early stages and then strongly upregulated at the 16-cell stage (Figure 4G). These results indicate that *PRDM14* might participate in the zygotic activation event and lineage specification in bovine early embryos.

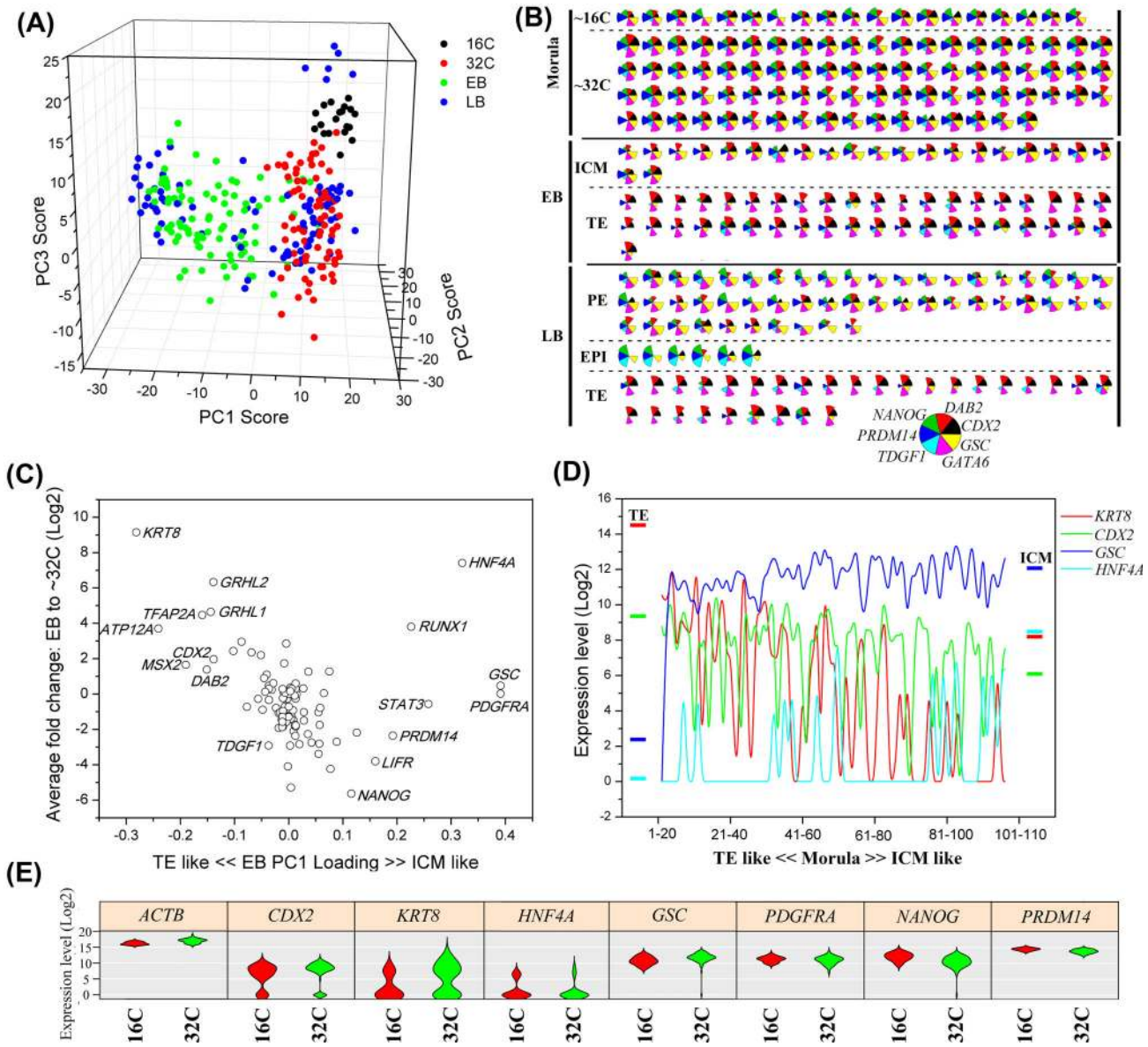

Figure 5. Lineage-specific genes highly and heterogeneously expressed in morula single cells. (A) The gene expression profiles from single cells at the 16-cell, 32-cell, and EBs stages projected onto the first three PCs of the LB data in Figure 3B. (B) Expression levels from 16-cell through to the LB stage for seven lineage-specific genes—*CDX2*/*DAB2* (TE), *NANOG*/*TDGF1*/*PRDM14* (EPI), *GATA6* (PE/TE), and *GSC* (PE/EPI)—are plotted as "slices" of "pies" representing individual cells. The radius of a slice reflects the expression level of the genes. The cells from the EB and LB stages are subdivided into TE and ICM or TE, EPI, and PE based on their projected positions in (A). (C) The average fold-change in expression between single cells from the 32-cell stage and either the ICM or TE cells of the early blastocyst is plotted against ICM/TE specificity. The average fold-change is calculated using either the TE or ICM of early blastocysts average depending on which is higher. The ICM/TE specificity is determined by each gene's contribution to PC1 (Figure 4B). (D) *KRT8*, *CDX2*, and *HNF4A* expression levels vary across the morula stage population. The morula stage cells on the x-axis are sorted according to their PC1 score in (A), so that TE-like cells are on the far left and ICM-like cells are on the far right. The traces represent moving averages of the given genes expression level in overlapping windows of 20 cells. Colored bars marked "TE" and "ICM" show the average expression levels in the early blastocyst TE and ICM, respectively. (E) Violin plots representing expression levels in morula individual cells for a number of genes; note the bimodal distribution of *KRT8*, *CDX2*, and *HNF4A*.

Lineage-specific genes highly and heterogeneously expressed in morula single cells

In the mouse, cellular differentiation is thought to initiate at the compacted and polarized late 8-cell stage. However, the compaction and polarization events occur at the 32-cell stage in bovines [33]. To visualize whether the cellular differentiation occurs at these stages in bovines, we projected the expression patterns of cells from the uncompacted 16-cell stage through to the EB stage onto the first three PCs computed for the LB (Figure 5A). During the 16–32-cell stage transition, the cells moved closer to the TE lineage.

As already described, the three cell types identified in the LB highly expressed specific genes. Interestingly, many of these genes that subsequently exhibited lineage-restricted expression in the LB are coexpressed at high levels in the majority of individual blastomeres at the morula stage. Moreover, their expression levels were comparable to the levels in lineage-specific single cells at the

Table 2. The gene pair correlations within LB ICM.

Inverse correlations		Positive correlations	
Gene pair	Score	Gene pair	Score
GATA6/TDGF1	−0.4867528274	*NANOG/PRDM14*	0.4575578175
GATA6/NANOG	−0.3641331176	*NANOG/TDGF1*	0.4639140282
GATA6/PRDM14	−0.2861280795	*PRDM14/TDGF1*	0.2989349973

blastocyst. Representative markers of the TE (*CDX2, DAB2*), EPI (*NANOG*, *TDGF1*, *PRDM14*), and PE (*GATA6*, *GSC*) are shown in Figure 5B. Zygotic activation occurs at the 8–16-cell stage in bovines. Thus, maternal transcripts might account for the high mRNA transcript levels in the morula blastomeres. However, for some of these genes, it is possible that their mRNAs are reduced in opponent lineages and increased within their own respective lineages during the transition from the morula to the LB.

To determine the probable mechanism of the first lineage segregation, we next identified the earliest significant expression differences during the transition from morula to the EB. We analyzed in detail the single-cell expression data derived from morulae and EBs. Cells of the EB were categorized as ICM or TE according to their PC1 score. Then, the average changes in gene expression that occurred between the 32-cell and relative cell types of blastocyst stages were calculated. *KRT8* and *HNF4A* changed greatly during this transition, with *KRT8* induction probably occurring in primordial TE cells (Figure 5C). However, the expression level of other lineage-specific genes, such as *CDX2* and *GSC*, was actually unchanged. PCA revealed some dispersion of cells along the PC1 axis, suggesting differential expression in single blastomeres of the morula (Figure 5A). Ordering the 16- and 32-cell-stage cells according to their PC1 score revealed that the variation of *KRT8* appears more significant than the TE-restricted marker *CDX2* (Figure 5D). The violin plots show that the distribution in expression of the endogenous control gene *ACTB* and the majority of other genes is unimodal, whereas *KRT8* clearly shows a bimodal distribution at the morula stage (Figure 5E; Supplementary Figure S3). This suggests that *KRT8* might be an early TE marker during the first lineage segregation event in bovine preimplantation embryos.

The second lineage specification from the inner cell mass to the primitive endoderm and epiblast

The inner cells from the EB will subsequently differentiate into the PE and EPI cells. To gain insight into the mechanism of the second lineage specification in bovines, we further analyzed our single-cell data sets from blastocysts to discover the earliest expression differences within the ICM. First, we calculated the correlation coefficient in gene expression for each gene pair with the 22 ICM cells in the EB data set and the 56 ICM cells in the LB data set as determined by their PC1 scores (Table 2). Consistent with the protein staining for *NANOG* and *GATA6* [10], we find that the *NANOG* expression is inversely correlated with *GATA6* expression in the LB ICM. Notably, *TDGF1/GATA6* gene pair is inversely correlated stronger than *NANOG/GATA6* gene pair. However, they are not correlated in EB ICM cells (data not shown). The correlation of *NANOG/GATA6* gene pair in bovines was not as strong as in mice [30], suggesting that *NANOG* mRNA is still not completely restricted to the EPI at the expanded blastocyst stage.

As shown in Figure 6A and 6B, the EPI-enriched genes *NANOG*, *TDGF1,* and *PRDM14* are initially weakly expressed in ICM cells of the EB, but are strongly activated in a subpopulation of ICM cells during the transition to the LB. Meanwhile, the PE specificity for *GATA6* is achieved by downregulating in opposing cell types during the transition from the early to the LB. This indicates that *TDGF1* and *PRDM14* might play pivotal roles in the second lineage specification of bovine embryos.

Discussion

In this study, we applied single-cell analysis to examine the expression patterns of 96 genes from hundreds of individual blastomeres collected from zygote to the LB across multiple preimplantation developmental stages of bovine embryos. Our study provides insight into the generation of the three lineages in bovine blastocysts at single-cell resolution (Supplementary Figure S4).

The MZT serves at least three important purposes including degradation of oocyte-specific transcripts, replacement of maternal transcripts with zygotic ones, and transcriptional upregulation of zygote-specific genes that will function to reset and reprogram gene expression patterns for the continued development of embryo [34]. The chromatin state of cells is crucial for gene expression regulation and development. In cattle, the major burst of gene activation occurs at the 8–16-cell stage. In this study, the transition from 8–16-cell stage involves the upregulation of multiple genes, such as lineage-restricted transcriptional factors including *NANOG* and *GATA6*, and chromatin modifiers including *PRDM14*, tet methylcytosine dioxygenase 1 (*TET1*), and histone deacetylase 1 (*HDAC1*). Meanwhile, DNA methyltransferase 1 (*DNMT1*) and *DNMT3B*, which are responsible for maintaining DNA methylation and de novo methylation, respectively, show downregulation patterns during this transition. Thus, these five chromatin modifiers might play key roles in the zygote genome activation event.

As seen in mouse embryos, not all bovine blastomeres at the morula stage express equivalent mRNA levels for TE markers, such as *CDX2* and *KRT8*, reflecting their previous cleavage history [30, 35]. Nonetheless, the majority of morula blastomeres coexpress high levels of genes restricted in the three lineages. This finding potentially provides the molecular foundation for the known developmental plasticity of bovine morula blastomeres [36, 37]. Additionally, we find the same phenomenon that some of the genes show a decrease of transcript levels in cells of opposing lineages as blastocyst cell types develop from the uncommitted blastomeres of the morula.

It has been proven that there are several differences between mouse and bovine early embryonic development. In common with a previous marker genes expression analysis, the mRNA expression of TE- and ICM-specific genes known in the mouse, such as *CDX2* and *POU5F1*, do not vanish in opposing lineages within the bovine blastocyst [13]. As a key TE transcription factor, *Cdx2* can repress *Pou5f1* transcription through binding to the conserved region four of the distal *Pou5f1* enhancer in mouse TE cells [38]. However, a study by Berg et al. [13] revealed that at the early stage, *CDX2* does not repress *POU5F1* expression in bovines, as the bovine *POU5F1* locus does not contain the cis-acting regulatory region (*TFAP2A/2C*) necessary for extinguishing its transcription in the TE. Moreover, *POU5F1* mRNA expression in the cattle TE may be retained until the PE/EPI segregation. Importantly, the TE cells of the EB retain the ability to contribute to the ICM derivatives in chimeras, indicating that the trophoblast cells may not become committed during bovine preimplantation development at stages developmentally equivalent to the mouse. This is also supported by the observation that the PC projections of individual blastomeres in the bovine EB do not show two obvious populations. The reason for this difference might be

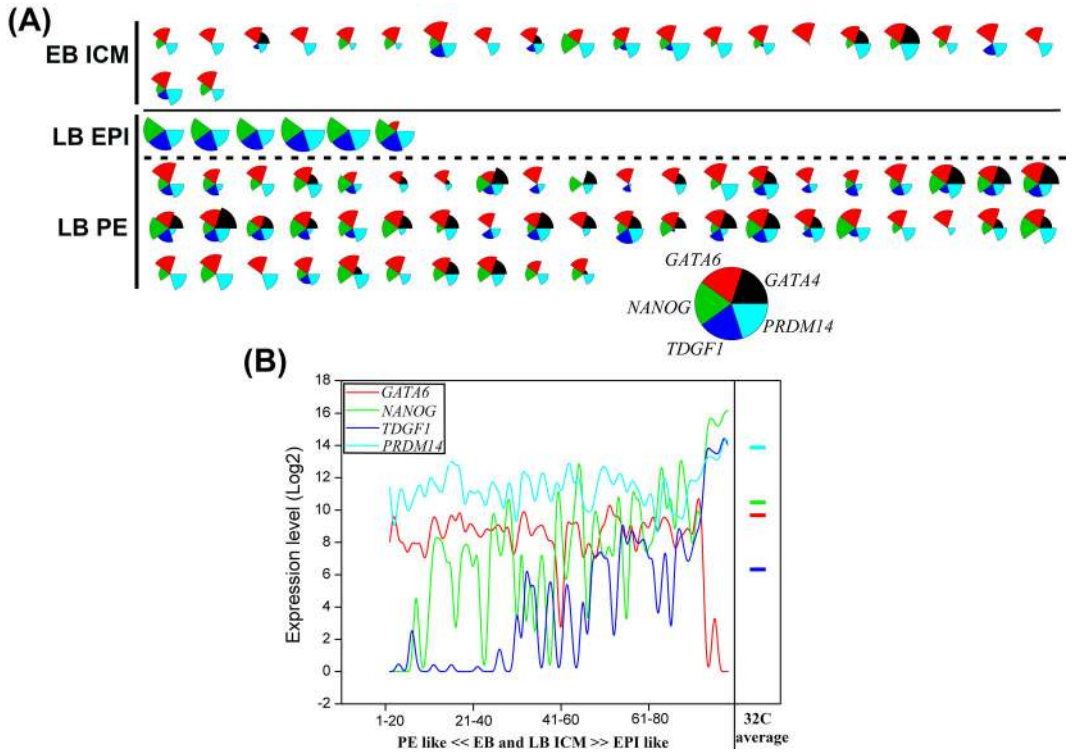

Figure 6. The second lineage specification from the ICM to the PE and EPI. (**A**) Expression levels of EPI- and PE-associated genes shown as pie plots according to the same logic as Figure 5B. (B) Variation of EPI- and PE-associated genes across ICM cells at the EB and LB stages. On the x-axis, cells from both stages are sorted according to their projection score to PC3 (Figure 4C). The traces represent moving averages of the given gene's expression level in overlapping windows of 20 cells. The colored side bars mark the average expression levels of 32-cell stage.

explained by the different developmental modes of mice and bovines. Mouse embryos implant to the uterus 1 day after the blastocyst formation and undergo a flurry of trophoblast proliferation and differentiation into giant cells, ectoplacental cone and extraembryonic ectoderm, before initiating gastrulation. However, bovine blastocysts float in the uterus for approximately 2 weeks during which the embryos undergo further differentiation [39].

Evidence has shown that PRDM14 regulates pluripotency and epigenetic reprogramming by directly binding to the regulatory regions of its target genes through its zinc finger domains. It can repress or activate its targets by multiple mechanisms. On one hand, PRDM14 can recruit polycomb repressive complex 2 (PRC2), a heteromeric multiprotein complex with histone H3 lysine 27 trimethylation (H3K27me3) activity [40], to repress expression of its target genes [41, 42]. On the other hand, PRDM14 appears to activate the expression of some targets by cooperating with other sequence-specific transcriptional regulators such as estrogen-related receptor beta [41]. It has also been shown that PRDM14 can contribute to H3 arginine 26 demethylation (H3R26me2) and DNA demethylation to activate its target genes by interacting with protein arginine methyltransferase 4 and ten-eleven translocation (TET), respectively [43]. A previous report in the mouse revealed that *Prdm14* mRNA exhibits a heterogeneous expression pattern in 4-cell-stage embryos and is subsequently highly enriched in ICM cells of the blastocyst. Forced expression of PRDM14 at the 2-cell stage leads to increased H3R26me2 and can induce a pluripotent ICM fate [32]. Similar to this, we find that *PRDM14* mRNA shows heterogeneous expression pattern in bovine 8-cell-stage embryos, and is highly expressed in ICM and EPI cells of bovine blastocysts. Contrary to the mouse, enhancer of zeste 2 polycomb repressive complex 2 subunit (*EZH2*), which encodes for a component of PRC2 complex, is highly expressed before implantation in the bovine (Supplementary Figure S2). Additionally, *TET1*, which encodes a protein responsible for DNA demethylation, is also highly expressed at late stages (Supplementary Figure S2). We therefore hypothesize that *PRDM14* might promote stem cell fate allocation to the ICM by repressing differentiation genes and activating pluripotent genes through PRC2 and TET, respectively, as is proposed in mouse ESCs. The precise mechanisms require further investigation in bovine PED.

TDGF1, also known as Cripto-1 (CR1), is a cell surface glycosylphosphatidylinositol-linked glycoprotein that can function either in an autocrine or paracrine manner. As a member of the epidermal growth factor/CR1-FRL-1-Cryptic family, CR1 functions as an obligatory coreceptor for transforming growth factor family members, NODAL and growth differentiation factor 1/3, by activating anaplastic lymphoma kinase 4/7 signaling pathways that involve SMADs 2, 3, and 4. In addition, CR1 can activate non-Smad-dependent signaling elements such as phosphatidylinositol 3 kinase, AKT, and mitogen-activated protein kinase. Both of these pathways depend upon the 78 kDa glucose-regulated protein. Finally, CR1 can facilitate signaling through the canonical WNT/and NOTCH pathways by functioning as a chaperone protein for low-density lipoprotein receptor-related protein 5/6 and NOTCH, respectively [44]. *Tdgf1* is essential for early embryonic development and maintains embryonic stem cell pluripotency in the mouse. During mouse embryonic development, *Tdgf1* remains strongly expressed in the ICM and EPI until gastrulation, where it becomes restricted to the primitive streak [45, 46]. Consistent with this, *TDGF1* mRNA is

enriched in EPI cells of bovine LBs. Whether *TDGF1* plays an important role in bovine early embryonic development and maintenance of stem cell pluripotency as occurs in the mouse needs to be further illustration.

Several signaling pathways are crucial for the lineage segregation of preimplantation embryos. Similar to the mouse, stimulation of bovine embryos with fibroblast growth factor 4 (FGF4) resulted in ICMs composed entirely of PE, suggesting that PE formation depends on FGF signaling [10]. In this study, we only examined the expression of *FGFR2* and *FGFR4* due to the low quality of primers for other ligands and receptors of FGF signaling. Contrary to what is claimed by a published report in the mouse [30], *FGFR4* but not fibroblast growth factor receptor 2 (*FGFR2*) mRNA is enriched in the bovine ICM. It is thus probable that *FGFR4* plays a more important role in bovine PED. Janus kinase (JAK) inhibition represses both EPI and hypoblast transcripts as well as naive pluripotency-related genes and the JAK substrate STAT3 [47]. The observation that *STAT3* mRNA is enriched in ICM cells of bovine blastocyst also supports the opinion that JAK/STAT3 activation is required for bovine ICM formation. Catenin beta like 1 (*CTNNBL1*) (also known as *B-CATENIN)*, as a downstream target of canonical WNT signal, which enhances pluripotency [48], is not expressed in bovine preimplantation embryos of late stages, whereas glycogen synthase kinase 3 beta (*GSK3B*) is highly expressed in bovine preimplantation embryos. This mRNA expression pattern corroborates previous studies on mice and bovines describing that activation of WNT signaling during early embryonic development inhibits development of embryos to the blastocyst stage [49, 50]. Bone morphogenetic protein (BMP) signaling regulates the correct development of both extraembryonic lineages, PE, and TE in mouse preimplantation embryos [51]. Although *BMP4* and bone morphogenetic protein receptor type 1A (*BMPR1A*), which encode a ligand and receptor for BMP signaling, respectively, are detectable in bovine early embryos (this study), they do not appear to be essential for bovine ICM development and gene expression [47]. Therefore, a detailed investigation is required to determine whether BMP signaling regulates extraembryonic lineages specification in bovine embryos as occurs in the mouse.

In summary, we performed a quantitative analysis of the expression of many genes in parallel at single-cell resolution throughout the time course of bovine PED. The combinatorial gene expression defines the cellular state and transitions in potency. Our single-cell analysis offers intriguing new insights into the early cell fate determination in bovines and offers theoretical support for deriving bovine ESCs.

Supplementary data

Supplementary data are available at *BIOLRE* online.

Supplemental Information contains Supplemental Experimental Procedures, two figures, and one table that are available online.

Supplementary Table S1. Summary of the genes analyzed in this study and their corresponding primer sequences information.

Supplementary Figure 1. Hierarchical clustering of single cells derived from early blastocysts based on the expression of 96 genes.

Supplementary Figure 2. Expression dynamics of each gene through preimplantation embryonic development. The boxed region represents the middle 50% of expression values, the black bar the median, and the whiskers the maximum and minimum values. Cells with outlying expression values are depicted as asterisks.

Supplementary Figure 3. Bars represent the absolute *CDX2*, *KRT8*, and *HNF4A* expression level in each cell of a representative 16- and 32-cell stage embryos.

Supplementary Figure 4. Schematic model of gene expression patterns in the segregation of the three blastocyst lineages. All the listed marker genes are enriched in the corresponding lineages. The mouse model is based on two published single-cell gene expression analysis researches, and the bovine model is based on our results. "Up arrow" marks upregulation of the gene expression. "Down arrow" marks downregulation of the gene expression. "Plus" marks maintaining high expression of the gene.

Acknowledgments

We would like to thank other members of the Dai lab for production of bovine in vitro fertilized embryos, and Ying Guo for technical assistance.

References

1. Hamatani T, Carter MG, Sharov AA, Ko MS. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell* 2004; **6**:117–131.
2. Cao SY, Han JY, Wu J, Li QY, Liu SC, Zhang W, Pei YL, Ruan XA, Liu ZH, Wang XM, Lim B, Li N. Specific gene-regulation networks during the pre-implantation development of the pig embryo as revealed by deep sequencing. *BMC Genomics* 2014; **15**:1471–2164.
3. Jarrell VL, Day BN, Prather RS. The transition from maternal to zygotic control of development occurs during the 4-cell stage in the domestic pig, Sus scrofa: quantitative and qualitative aspects of protein synthesis. *Biol Reprod* 1991; **44**:62–68.
4. Kues WA, Sudheer S, Herrmann D, Carnwath JW, Havlicek V, Besenfelder U, Lehrach H, Adjaye J, Niemann H. Genome-wide expression profiling reveals distinct clusters of transcriptional regulation during bovine preimplantation development in vivo. *Proc Natl Acad Sci USA* 2008; **105**:19768–19773.
5. Braude P, Bolton V, Moore S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* 1988; **332**:459–461.
6. Degrelle SA, Campion E, Cabau C, Piumi F, Reinaud P, Richard C, Renard JP, Hue I. Molecular evidence for a critical period in mural trophoblast development in bovine blastocysts. *Dev Biol* 2005; **288**:448–460.
7. Kirchhof N, Carnwath JW, Lemme E, Anastassiadis K, Scholer H, Niemann H. Expression pattern of Oct-4 in preimplantation embryos of different species. *Biol Reprod* 2000; **63**:1698–1705.
8. van Eijk MJ, van Rooijen MA, Modina S, Scesi L, Folkers G, van Tol HT, Bevers MM, Fisher SR, Lewin HA, Rakacolli D, Galli C, de Vaureix C et al. Molecular cloning, genetic mapping, and developmental expression of bovine POU5F1. *Biol Reprod* 1999; **60**:1093–1103.
9. Vejlsted M, Du Y, Vajta G, Maddox-Hyttel P. Post-hatching development of the porcine and bovine embryo–defining criteria for expected development in vivo and in vitro. *Theriogenology* 2006; **65**:153–165.
10. Kuijk EW, van Tol LT, Van de Velde H, Wubbolts R, Welling M, Geijsen N, Roelen BA. The roles of FGF and MAP kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos. *Development* 2012; **139**:871–882.
11. Meissner A, Jaenisch R. Generation of nuclear transfer-derived pluripotent ES cells from cloned Cdx2-deficient blastocysts. *Nature* 2006; **439**:212–215.
12. Wu G, Gentile L, Do JT, Cantz T, Sutter J, Psathaki K, Arauzo-Bravo MJ, Ortmeier C, Scholer HR. Efficient derivation of pluripotent stem cells from siRNA-mediated Cdx2-deficient mouse embryos. *Stem Cells Dev* 2011; **20**:485–493.
13. Berg DK, Smith CS, Pearton DJ, Wells DN, Broadhurst R, Donnison M, Pfeffer PL. Trophectoderm lineage determination in cattle. *Dev Cell* 2011; **20**:244–255.

14. Goissis MD, Cibelli JB. Functional characterization of CDX2 during bovine preimplantation development in vitro. *Mol Reprod Dev* 2014; **81**:962–970.
15. Hubner K, Fuhrmann G, Christenson LK, Kehler J, Reinbold R, De La Fuente R, Wood J, Strauss JF 3rd, Boiani M, Scholer HR. Derivation of oocytes from mouse embryonic stem cells. *Science* 2003; **300**: 1251–1256.
16. Geijsen N, Horoschak M, Kim K, Gribnau J, Eggan K, Daley GQ. Derivation of embryonic germ cells and male gametes from embryonic stem cells. *Nature* 2004; **427**:148–154.
17. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 1981; **292**:154–156.
18. Thomson JA, Kalishman J, Golos TG, Durning M, Harris CP, Becker RA, Hearn JP. Isolation of a primate embryonic stem cell line. *Proc Natl Acad Sci USA* 1995; **92**:7844–7848.
19. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. *Science* 1998; **282**:1145–1147.
20. Buehr M, Meek S, Blair K, Yang J, Ure J, Silva J, McLay R, Hall J, Ying QL, Smith A. Capture of authentic embryonic stem cells from rat blastocysts. *Cell* 2008; **135**:1287–1298.
21. Blomberg LA, Telugu BP. Twenty years of embryonic stem cell research in farm animals. *Reprod Domest Anim* 2012; **47**:80–85.
22. Wu X, Song M, Yang X, Liu X, Liu K, Jiao C, Wang J, Bai C, Su G, Liu X, Li G. Establishment of bovine embryonic stem cells after knockdown of CDX2. *Sci Rep* 2016; **6**:28343.
23. Keefer CL, Pant D, Blomberg L, Talbot NC. Challenges and prospects for the establishment of embryonic stem cell lines of domesticated ungulates. *Anim Reprod Sci* 2007; **98**:147–168.
24. Ozawa M, Sakatani M, Yao J, Shanker S, Yu F, Yamashita R, Wakabayashi S, Nakai K, Dobbs KB, Sudano MJ, Farmerie WG, Hansen PJ. Global gene expression of the inner cell mass and trophectoderm of the bovine blastocyst. *BMC Dev Biol* 2012; **12**:33.
25. Nagatomo H, Kagawa S, Kishi Y, Takuma T, Sada A, Yamanaka K, Abe Y, Wada Y, Takahashi M, Kono T, Kawahara M. Transcriptional wiring for establishing cell lineage specification at the blastocyst stage in cattle. *Biol Reprod* 2013; **88**:1–10.
26. Jiang Z, Sun J, Dong H, Luo O, Zheng X, Obergfell C, Tang Y, Bi J, O'Neill R, Ruan Y, Chen J, Tian XC. Transcriptional profiles of bovine in vivo pre-implantation development. *BMC Genomics* 2014; **15**:1471–2164.
27. Zhao XM, Cui LS, Hao HS, Wang HY, Zhao SJ, Du WH, Wang D, Liu Y, Zhu HB. Transcriptome analyses of inner cell mass and trophectoderm cells isolated by magnetic-activated cell sorting from bovine blastocysts using single cell RNA-seq. *Reprod Domest Anim* 2016; **51**: 726–735.
28. Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL, van Oudenaarden A, Jaenisch R. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 2012; **150**:1209–1222.
29. Moignard V, Macaulay IC, Swiers G, Buettner F, Schutte J, Calero-Nieto FJ, Kinston S, Joshi A, Hannah R, Theis FJ, Jacobsen SE, de Bruijn MF et al. Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. *Nat Cell Biol* 2013; **15**:363–372.
30. Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell* 2010; **18**:675–685.
31. Dietrich JE, Hiiragi T. Stochastic patterning in the mouse pre-implantation embryo. *Development* 2007; **134**:4219–4231.
32. Burton A, Muller J, Tu S, Padilla-Longoria P, Guccione E, Torres-Padilla ME. Single-cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the mammalian embryo. *Cell Rep* 2013; **5**: 687–701.
33. Van Soom A, Boerjan ML, Bols PE, Vanroose G, Lein A, Coryn M, de Kruif A. Timing of compaction and inner cell allocation in bovine embryos produced in vivo after superovulation. *Biol Reprod* 1997; **57**:1041–1049.
34. Schultz RM. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Hum Reprod Update* 2002; **8**:323–331.
35. Jedrusik A, Parfitt DE, Guo G, Skamagki M, Grabarek JB, Johnson MH, Robson P, Zernicka-Goetz M. Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. *Genes Dev* 2008; **22**:2692–2706.
36. Williams TJ, Elsden RP, Seidel GE, Jr. Pregnancy rates with bisected bovine embryos. *Theriogenology* 1984; **22**:521–531.
37. Rho GJ, Johnson WH, Betteridge KJ. Cellular composition and viability of demi- and quarter-embryos made from bisected bovine morulae and blastocysts produced in vitro. *Theriogenology* 1998; **50**:885–895.
38. Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, Rossant J. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* 2005; **123**:917–929.
39. Berg DK, van Leeuwen J, Beaumont S, Berg M, Pfeffer PL. Embryo loss in cattle between days 7 and 16 of pregnancy. *Theriogenology* 2010; **73**:250–260.
40. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature* 2011; **469**:343–349.
41. Yamaji M, Ueda J, Hayashi K, Ohta H, Yabuta Y, Kurimoto K, Nakato R, Yamada Y, Shirahige K, Saitou M. PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* 2013; **12**:368–382.
42. Chan YS, Goke J, Lu X, Venkatesan N, Feng B, Su IH, Ng HH. A PRC2-dependent repressive role of PRDM14 in human embryonic stem cells and induced pluripotent stem cell reprogramming. *Stem Cells* 2013; **31**:682–692.
43. Okashita N, Kumaki Y, Ebi K, Nishi M, Okamoto Y, Nakayama M, Hashimoto S, Nakamura T, Sugasawa K, Kojima N, Takada T, Okano M et al. PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development* 2014; **141**:269–280.
44. Klauzinska M, Castro NP, Rangel MC, Spike BT, Gray PC, Bertolette D, Cuttitta F, Salomon D. The multifaceted role of the embryonic gene Cripto-1 in cancer, stem cells and epithelial-mesenchymal transition. *Semin Cancer Biol* 2014; **29**:51–58.
45. Ding J, Yang L, Yan YT, Chen A, Desai N, Wynshaw-Boris A, Shen MM. Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. *Nature* 1998; **395**:702–707.
46. Fiorenzano A, Pascale E, D'Aniello C, Acampora D, Bassalert C, Russo F, Andolfi G, Biffoni M, Francescangeli F, Zeuner A, Angelini C, Chazaud C et al. Cripto is essential to capture mouse epiblast stem cell and human embryonic stem cell pluripotency. *Nat Commun* 2016; **7**:12589.
47. Meng F, Forrester-Gauntlett B, Turner P, Henderson H, Oback B. Signal inhibition reveals JAK/STAT3 pathway as critical for bovine inner cell mass development. *Biol Reprod* 2015; **93**:1–9.
48. Martello G, Sugimoto T, Diamanti E, Joshi A, Hannah R, Ohtsuka S, Gottgens B, Niwa H, Smith A. Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* 2012; **11**:491–504.
49. Denicol AC, Dobbs KB, McLean KM, Carambula SF, Loureiro B, Hansen PJ. Canonical WNT signaling regulates development of bovine embryos to the blastocyst stage. *Sci Rep* 2013; **3**:1266.
50. Xie H, Tranguch S, Jia X, Zhang H, Das SK, Dey SK, Kuo CJ, Wang H. Inactivation of nuclear Wnt-beta-catenin signaling limits blastocyst competency for implantation. *Development* 2008; **135**:717–727.
51. Graham SJ, Wicher KB, Jedrusik A, Guo G, Herath W, Robson P, Zernicka-Goetz M. BMP signalling regulates the pre-implantation development of extra-embryonic cell lineages in the mouse embryo. *Nat Commun* 2014; **5**:5667.