

RESEARCH

Open Access



# Single-cell analysis of preimplantation embryonic development in guinea pigs

Tongxing Guan<sup>1,2,5†</sup>, Jing Guo<sup>2,5†</sup>, Runxia Lin<sup>2,5</sup>, Jinpeng Liu<sup>2,4,5</sup>, Rongping Luo<sup>2,5</sup>, Zhen Zhang<sup>2,5\*</sup>, Duanqing Pei<sup>1\*</sup> and Jing Liu<sup>2,3,5\*</sup>

## Abstract

**Background** Guinea pigs exhibit numerous physiological similarities to humans, yet the details of their preimplantation embryonic development remain largely unexplored.

**Results** To address this, we conducted single-cell sequencing on the transcriptomes of cells isolated from the zygote stage through preimplantation stages in guinea pigs. This study identified seven distinct cell types within guinea pig preimplantation embryos and pinpointed the timing of zygotic gene activation (ZGA). Trajectory analysis revealed a bifurcation into two lineage-specific branches, accompanied by alterations in specific pathways, including oxidative phosphorylation and vascular endothelial growth factor (VEGF). Additionally, co-expressed gene network analysis highlighted the most enriched functional modules for the epiblast (EPI), primitive endoderm (PrE), and inner cell mass (ICM). Finally, we compared the similarities and differences between human and guinea pig epiblasts (EPIs).

**Conclusion** This study systematically constructs a cell atlas of guinea pig preimplantation embryonic development, offering fresh insights into mammalian embryonic development and providing alternative experimental models for studying human embryonic development.

**Keywords** Guinea pigs, Single-cell sequencing, Early embryonic development

<sup>†</sup>Tongxing Guan and Jing Guo contributed equally to this work.

Jing Liu is the Lead contact.

\*Correspondence:

Zhen Zhang

zhang\_zhen@gibh.ac.cn

Duanqing Pei

peiduanqing@westlake.edu.cn

Jing Liu

liu\_jing@gibh.ac.cn

<sup>1</sup>School of Life Sciences, University of Science and Technology of China, Hefei 230026, China

<sup>2</sup>Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Guangdong-Hong Kong Joint Laboratory for Stem Cell and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China

<sup>3</sup>Joint School of Life Sciences, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou Medical University, Guangzhou 510530, China

<sup>4</sup>University of Chinese Academy of Science, Beijing 100049, People's Republic of China

<sup>5</sup>Center for Development and Regeneration, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China



## Introduction

Early embryonic development, especially mammalian embryonic development (including human), has been studied mainly using mouse and non-human primate embryos as models [1]. Mice are extensively used due to their quick reproductive cycle and the advancement of technologies such as gene editing and embryo manipulation. The preimplantation development in mice spans approximately five days, progressing from a single fertilized egg through multiple cell divisions, zygotic gene activation, compaction, polarization, and cavitation to form the blastocyst cavity. The initial lineage separation occurs during the transition from the morula to the early blastocyst stage [2]. Subsequently, the totipotent inner cell mass (ICM) differentiates into the primitive endoderm (PrE) and the pluripotent epiblast (EPI) [3–5]. Extensive research has focused on cell fate determination in mouse preimplantation embryos. Although human and mouse preimplantation embryos exhibit morphological similarities, mice develop early blastocysts at embryonic day 3.25 (E3.25) whereas humans do so at day post fertilization 5 (d.p.f 5). Late blastocysts form at E3.75 in mice and d.p.f 6–7 in humans. The first lineage separation begins in the morula of mice and the expanding blastocyst of humans, with the second lineage separation occurring in the late blastocyst [6, 7]. Human embryonic development research has been limited in part due to ethical and other constraints, and the maturity of technologies such as micromanipulation and gene editing has facilitated the use of mice and other alternative models for research. However, as mentioned above, mouse embryo development still exhibits differences from human embryo development in certain aspects. Therefore, it is essential to study embryos from various species to gain a comprehensive understanding of mammalian embryonic development.

Guinea pigs, also known as Dutch pigs, share several physiological similarities with humans. Firstly, their immune systems closely resemble those of humans, making guinea pigs excellent models for immunization studies [8–10]. Neuroanatomically mature, guinea pigs have well-characterized brain and neuroendocrine development, with earlier brain maturation than rats [11, 12]. They are among the few laboratory rodents that exhibit a complete estrous cycle, including follicular and luteal phases similar to humans, unlike most rodents such as mice, rats, and hamsters, which have distinct and incomplete estrous cycles lacking a functional luteal phase [13]. Crucially, guinea pigs have a placental pattern similar to humans [14–16] and possess unique developmental advantages over mice and rats [17–21]. Additionally, like humans, guinea pigs cannot synthesize their own vitamin C, and research indicates that vitamin C deficiency

in these animals can negatively impact fertility and pregnancy [22, 23].

Due to the numerous similarities between guinea pigs and humans, the use of guinea pigs was prevalent in the early 1970s, surpassing that of mice, rats, and hamsters. However, this usage has since declined [24]. The limitations in genome assembly and the immaturity of gene editing technologies have contributed to this decrease. Given these similarities and limitations, further investigation into guinea pig embryonic development is essential to bridge gaps in our understanding of mammalian development.

In this study, we hypothesize that guinea pig embryos undergo zygotic genome activation, followed by the first and second cell fate determinations, all occurring in a relatively orderly manner before implantation under physiological conditions. To address this, we propose to employ single-cell transcriptome sequencing technology to explore the lineage decisions and gene expression patterns during guinea pig preimplantation embryo development. Our objectives include characterizing cell types in these embryos, pinpointing the timing of zygotic genome activation (ZGA) in guinea pigs, and delineating two critical cell fate-determining branches with specific pathway alterations across developmental stages. Additionally, we aim to identify functional modules predominantly enriched during the second lineage segregation involving three distinct cell types. Lastly, we will investigate interspecies differences in the preimplantation epiblast (EPI) population between humans and guinea pigs.

## Methods

### Collection of guinea pig preimplantation embryos and isolation of single cell samples

Pregnant guinea pigs were euthanized using carbon dioxide, and the reproductive organs (ovaries, oviducts, and uterus) were excised and immersed in pre-warmed (37 °C) phosphate-buffered saline (PBS). To harvest zygotes, two-cell, four-cell, and eight-cell stage embryos, the fallopian tube was incised at the uterine horn junction, and 500 µl of pre-warmed embryo collection solution (DMEM/F12 supplemented with 2% fetal bovine serum, FBS) was injected using a 1 ml syringe. The embryos were then flushed out from the fallopian tube's fimbriae. For the collection of morulae and blastocysts, the uterus was cut near the cervix, and a syringe needle was inserted near the tubal-uterine horn junction to inject 1 to 2 ml of preheated embryo collection solution, facilitating the flushing of embryos from the uterus. The zona pellucida was subsequently removed by treating the embryos with Acidic Tyrode's Solution for 15–30 s. Embryos without zona pellucida were treated with Accutase alone or a mixture of Accutase and TrypLE (1:1) for 3–30 min until the cells were loosened. Single

cells were then dispersed in DPBS containing 0.2% BSA using a mouth pipette for sampling. To separately sample mural trophoctoderm (mTE) cells, some blastocyst-stage embryos were microdissected along the inner cell mass (ICM) using a laser membrane rupture device on a micromanipulation platform. The mTE and the ICM plus polar trophoctoderm (pTE) groups were then digested and sampled individually.

### Construction and sequencing of single-cell RNA-seq libraries

Cells were lysed, and RNA with polyA tails, primarily mRNA, was reverse transcribed using an oligo(dT) primer to synthesize the first strand of cDNA. A splice sequence was appended to the 3' end of the cDNA utilizing the template-switching activity of the reverse transcriptase. The second strand of cDNA was synthesized at this junction sequence using a TSO (template-switching oligo) primer, which displaced the RNA complementary to the first strand of cDNA. PCR amplification followed to achieve full-length cDNA products, effectively mitigating 3' bias and rRNA contamination during cDNA synthesis. The resultant single-cell cDNAs underwent quality control using a high-sensitivity chip, with fragment lengths ranging from 400 bp to 10,000 bp, and the peak of the library distribution was around 2,000 bp, indicating qualified cDNA libraries. These qualified single-cell cDNA libraries were fragmented using Tn5 transposase, which also added junction sequences to the cDNA ends. A final round of PCR amplification of the labeled cDNAs was conducted using sequencing primers, and the amplified cDNAs were purified with magnetic beads to yield on-board libraries of 200–600 bp. All steps of the cDNA library construction, including lysis, reverse transcription, amplification, and purification, were carried out in-house by our lab.

The sequencing platforms chosen for this study were Illumina's NovaSeq6000 and BGI's MGISEQ2000. The data volume generated from a single-cell library on the paired-end PE150/100 machines was approximately 1.5G–2G, sufficient for comprehensive analysis. The library construction method employed was Smart-seq2 [25].

### Bioinformatics workflow

Upon receiving raw FASTQ files, we conducted quality control with FastQC and MultiQC to generate comprehensive quality reports. The subsequent adapter trimming was performed using Trim Galore. Post-quality check, the data were aligned with HISAT2, and expression quantification was achieved using featureCounts. The resulting expression matrix was then processed through the "Seurat" package for standard single-cell workflow analysis. Annotated cells underwent

differential expression analysis, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis, developmental trajectory construction with Monocle2, and high-dimensional gene co-expression network analysis using hdWGCNA to examine lineage-specific gene expression characteristics. Additionally, we conducted interspecies comparisons. Detailed descriptions of the algorithms used and their corresponding parameters are provided in the following sections.

### Preprocessing of scRNA-seq data

The quality of the initial sequencing data, represented by the FASTQ file, is assessed, encompassing the GC content, base quality at each position, base distribution, presence of junctions, and junction content. Sequencing files that exhibit no evident quality issues, contamination from adapter sequences or other species, or sequences of identical length are retained. Adapter sequences are automatically detected and removed using "trim\_galore," which also eliminates low-quality bases scoring below q20 (--quality 20). The sequenced reads are aligned to the guinea pig reference genome (Cavpor3.0) using the Hisat2 (v2.2.1) software package [26]. Unmatched data is filtered out using "samtools" (v1.17) [27], resulting in bam files. The gene expression matrix is subsequently quantified by counting the number of molecules per gene, guided by the annotation file, through "featureCounts" (Liao et al., 2014).

### Dimensionality reduction clustering

The expression matrix derived from the upstream analysis was imported into R, and the Seurat package (v4.4.0) [28] was employed to create Seurat objects. Following a simple merge, three batches of data were filtered based on QC-specific criteria: genes expressed in at least 3 cells, and thresholds for nFeature\_RNA (>5000), nCount\_RNA (>1,000,000), and nCount\_RNA (<10,000,000). The standard Seurat workflow was then applied, including "NormalizeData," "FindVariableFeatures," and "ScaleData." Principal component analysis (PCA) was conducted using the "RunPCA" function. Clustering was achieved through "FindNeighbors" and "FindClusters," and the results were visualized using Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction. The "FindAllMarkers" function, with default parameters, was used for identifying characteristic genes of each cluster. The default parameters used in "FindAllMarkers" for this study are "only.pos=FALSE, min.pct=0.25, thresh.use=0.25". Cell type annotation was performed manually. Default parameters were used throughout, with dimensions selected from 1 to 15 (dims=1:15) and a resolution of 0.8.

### Identification of the period of occurrence of ZGA

To detect changes in gene expression during successive developmental stages, we conducted a differential analysis comparing each time point to its preceding one. Genes exhibiting significant differences were identified using the criteria of absolute average log<sub>2</sub> fold change (avg\_log<sub>2</sub>FC) greater than 1 and a *p*-value less than 0.05. Visualization of these differential genes was achieved through scatter plots and volcano plots, both of which were generated using the “ggplot2” package [29].

### DEG analysis

The analysis was conducted using the built-in “FindMarkers” function within the Seurat package, which employs the “wilcox” test for identifying differentially expressed genes. The specific parameters are “verbose=FALSE, test.use = ‘wilcox’, min.pct=0.1,” with all other parameters set to default.

### GO and KEGG enrichment analysis

The initial step involved utilizing the “AnnotationHub” package [30] to retrieve annotations specific to guinea pigs. Following this, GO and KEGG enrichment analyses were conducted on the outcomes of the differential expression analysis, employing the “clusterProfiler” package (v4.10.1) [31]. Visualization of the results was accomplished through the “GOplot” package [32].

### Pseudotime analysis

To investigate the two genealogical separations, single-cell trajectory analysis was conducted using the R Monocle2 package with the DDRTree method and default parameters. The top 3000 genes were selected as the input for this trajectory analysis. Subsequently, heatmaps were generated to display significantly changed genes, which were clustered based on their pseudo-temporal expression patterns. Pathway enrichment analysis was then performed using gene sets from KEGG. In a similar vein, the “AddModuleScore\_UCell” function [33] was employed to integrate the gene set scoring results into the corresponding matrix of the Monocle object, and the “plot\_cell\_trajectory” function was used to visualize the pathways.

### High dimensional weighted gene co-expression network analysis (hdWGCNA)

To pinpoint genes linked to the second lineage segregation in guinea pigs, a co-expression gene network analysis was executed using the “hdWGCNA” package [34]. The process involved constructing metacells individually for each sample and cell cluster using the hdWGCNA function, with 50 cells amalgamated per metacell. For each cell population, a subset of Seurat objects was established for the targeted cell population, followed by

the application of the standard hdWGCNA pipeline. This pipeline was implemented sequentially through the execution of the following functions with default parameters: TestSoftPowers, ConstructNetwork, ModuleEigengenes, ModuleConnectivity, and RunModuleUMAP.

### Species comparison

The dataset utilized in this study, sourced from GSE136447 [35], comprised scRNA-seq profiles of human preimplantation EPI cells aligned with the human genome. The dataset included FPKM values for 13,915 homologous genes, derived from a combination of 106 human ICM and EPI single cells and 315 guinea pig ICM and EPI scRNA-seq profiles. To refine the dataset, filtering was applied to retain 10,698 genes common to both human and guinea pig, based on the criterion that these genes had a log<sub>2</sub>(FPKM+1) value of at least 4 in at least one of the 106 human or 315 guinea pig cells. The raw FPKM values for these 10,698 genes from both species were subsequently normalized and scaled using the Seurat package in R. Principal Component Analysis (PCA) was conducted on these normalized and scaled FPKM values via the prcomp function in R. The PCA results for the first three major components were visualized using MatLab (MathWorks) [36]. Notably, 952 homologous genes in human and guinea pig significantly contributed to PC1, with absolute PC1 values greater than 0.8 standard deviations. Clustering and visualization of 12,475 genes were performed using the heatmap function on the R platform. Further analysis focused on genes with PC1 values greater than 0.8 and less than -1.75 standard deviations, encompassing 250 and 702 genes respectively, for KEGG enrichment using the clusterProfiler package.

### Animals

Females and males representing the outbred Hartley strain of guinea pigs (*Cavia porcellus*) were purchased from Beijing Vital River Laboratory.

### Results

#### Acquisition of high-quality guinea pig preimplantation embryos

To acquire high-quality single-cell data from guinea pig preimplantation embryos, we employed morphological observation to stage these embryos based on post-fertilization time and microscopic morphological features. Embryos were harvested from guinea pigs at various embryonic days ranging from E0.5 to E5.5, encompassing stages from Zygote to Blastocyst, including 2-cell, 4-cell, 8-cell, early morula, late morula, early blastocyst, and blastocysts stages.

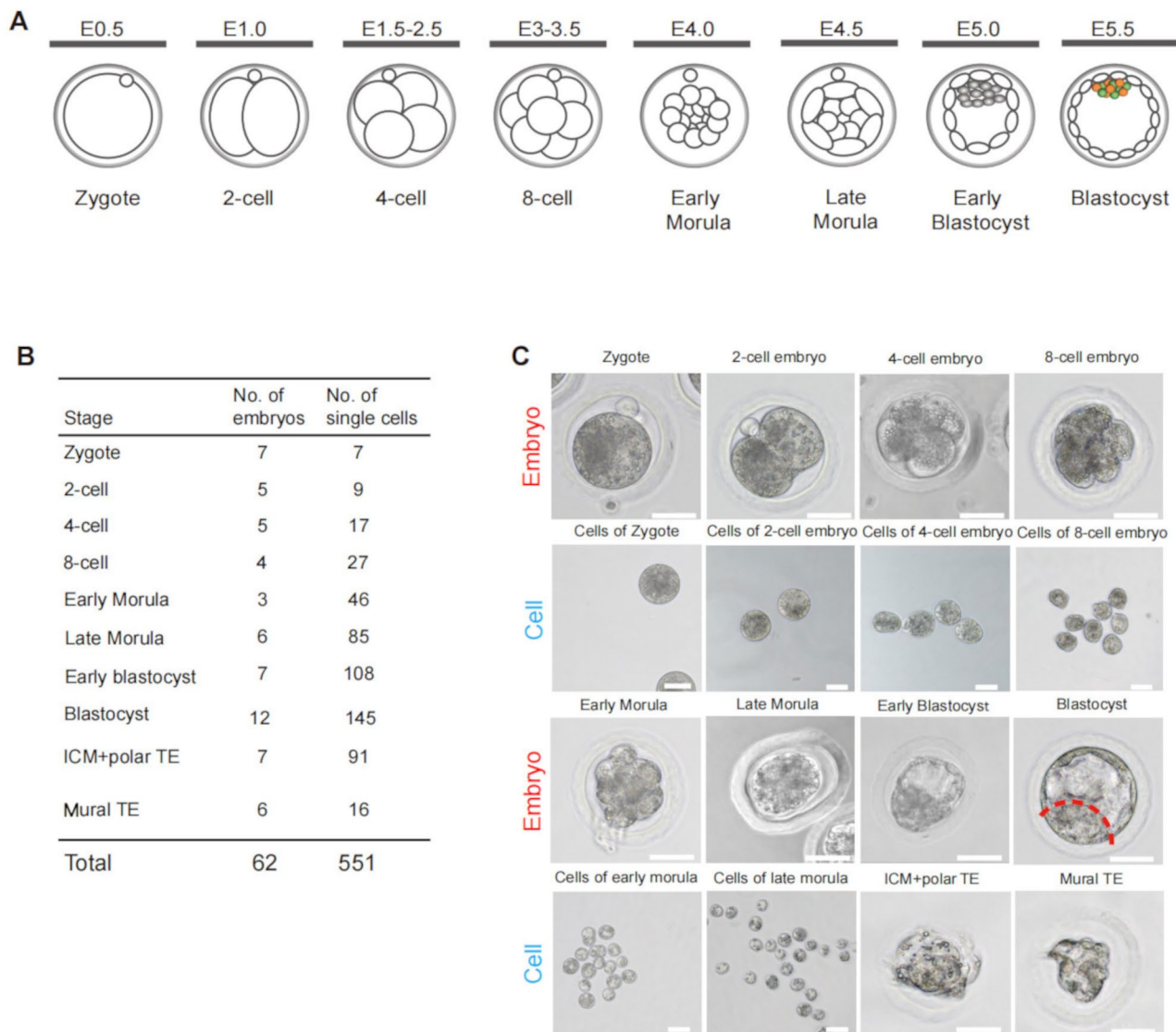
Some blastocysts were further dissected into Inner Cell Mass (ICM) plus polar Trophoblast Cells (pTE) and

mural Trophoblast Cells (mTE) groups using microdissection. These embryos were then enzymatically dissociated into single cells. We collected a total of 599 samples from 39 pregnant guinea pigs (Fig. 1A, B). Following quality control (QC), we obtained final data from 551 high-quality single-cell samples derived from guinea pig preimplantation embryos. The morphology of each stage of embryos and the dissociated single cells was documented (Fig. 1C, Fig.S1A, B).

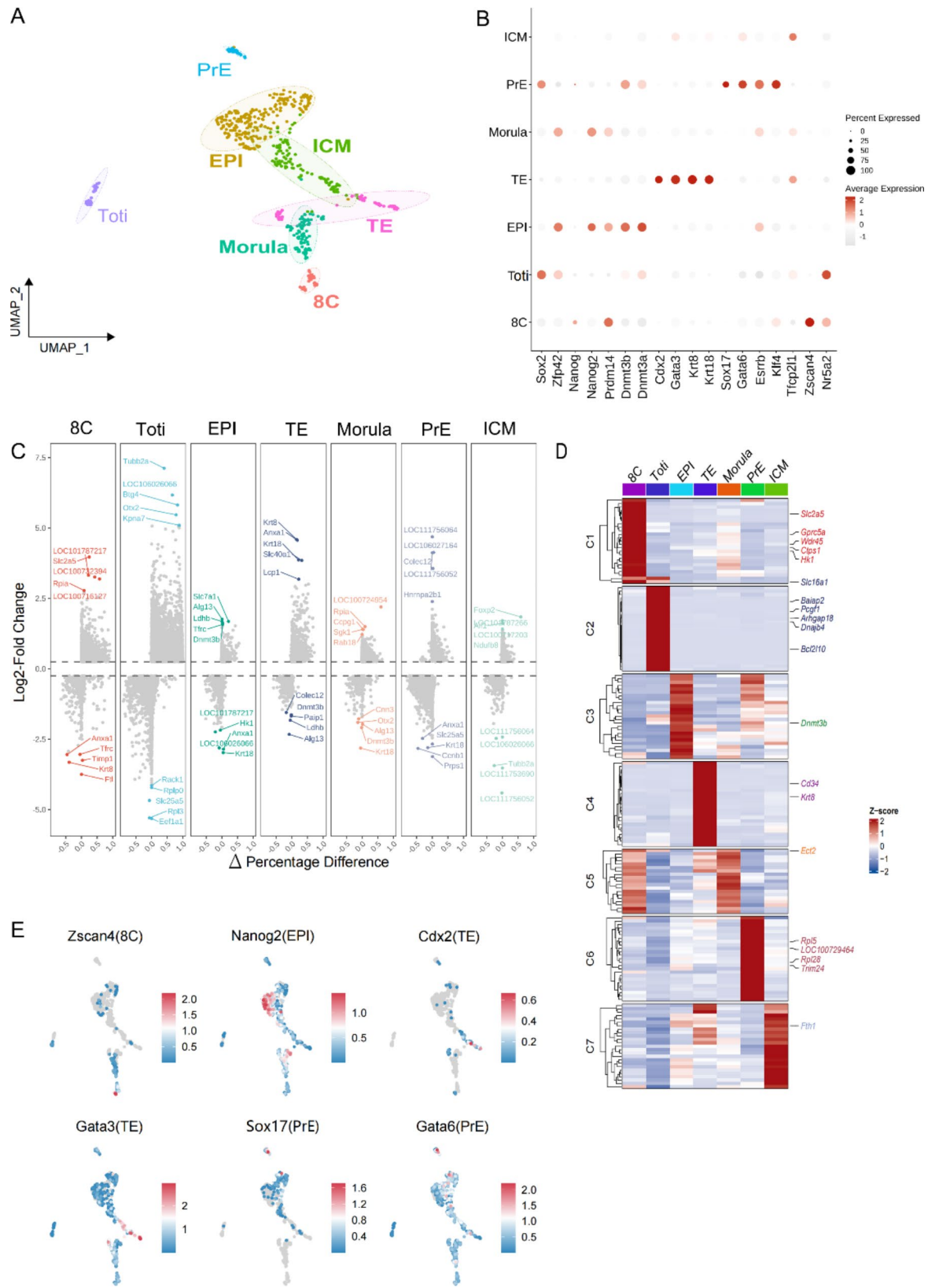
### Single-cell analysis of guinea pig preimplantation embryos

The post-QC single-cell data were subjected to UMAP dimensionality reduction and classified into a total of 9 clusters. Given the limited research on guinea pig preimplantation embryos, we relied on cell markers from

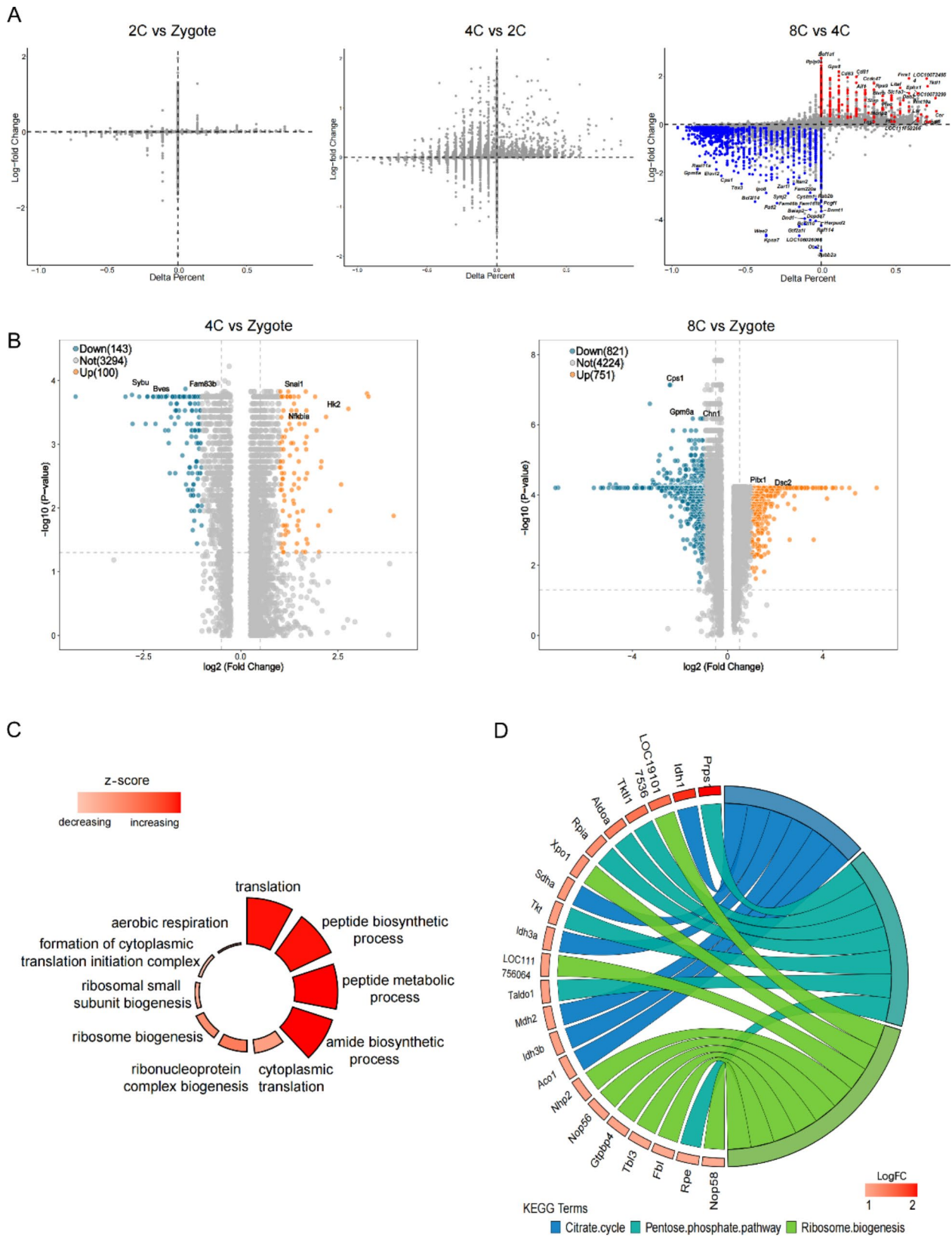
human and mouse preimplantation embryos to characterize the single-cell data of guinea pig embryos (Fig. 2B; Fig. S1D) [37–40]. The joint morphological pre-staging of sample information facilitated accurate cell annotation of clusters (Fig. S1C). We defined Totipotency (Toti) to include Zygote, 2-cell, and 4-cell stages, as well as 8-cell, morula, Trophoblastic Ectoderm (TE), and Primitive Endoderm (PrE) based on morphological observations and corresponding markers (Fig. 2A). However, the definitions of the Inner Cell Mass (ICM) and Epiblast (EPI) were more ambiguous. The regulation of the pre-implantation Epiblast (EPI) in various species has been shown to be regulated by *Nanog* [41–43]. Consequently, we aimed to use *Nanog* as a specific marker for EPI in cell cluster characterization. However, unexpectedly, *Nanog*



**Fig. 1** Preimplantation Embryo Sampling in Guinea Pigs (A) Embryo sampling conducted at different preimplantation stages. (B) Summary of the number of embryos and cells collected at each time point. (C) Representative bright-field microscopy images of in vivo acquired embryos. Scale bar: 50  $\mu$ m



**Fig. 2** Transcriptomic analysis of preimplantation embryos in guinea pigs **(A)** UMAP visualization of 551 single-cell transcriptomes from guinea pig preimplantation embryos. Points are colored by cell types. **(B)** Dot plot of selected marker genes for various cell types. **(C)** Visualization of the top 5 upregulated and downregulated genes in each cluster. **(D)** Heatmap showing the top 25 differentially expressed genes (DEGs) in each cluster. **(E)** UMAP plot of all cells with projected expression patterns of the following marker genes for each cluster: Zscan4 (8 C), Nanog2 (EPI), Cdx2, Gata3 (TE), and Sox17, Gata6 (PrE)



**Fig. 3** (See legend on next page.)

(See figure on previous page.)

**Fig. 3** The zygotic genome activation (ZGA) of guinea pigs **(A)** Percentage difference ( $\Delta$  percent of cells) and log-fold change based on the Wilcoxon rank-sum test results for differential gene expression comparing 2 C versus zygote (Left), 4 C versus 2 C (Middle), and 8 C versus 4 C (Right). Red: upregulated genes; Blue: downregulated genes. **(B)** Volcano plots of differential gene expression comparing 4 C versus zygote (Left) and 8 C versus zygote (Right). **(C)** Gene Ontology (GO) circle plot displaying gene-annotation enrichment analysis. Red indicates upregulated gene-associated GO terms, relative to the z-score of the analysis. **(D)** KEGG Chord plot illustrating the relationship between the list of selected KEGG terms (upregulated genes in 8 C versus 4 C) and their corresponding genes, along with the log fold change (logFC) of the genes. The left half of the KEGG Chord plot displays whether the gene expression is upregulated, and the right half represents different KEGG terms with various colors

expression was not detected in the available single-cell data (Fig. 2B). Given that guinea pig embryos initiate implantation at E6 [21], and theoretically, *Nanog* should be expressed in both ICM and EPI at E5.5, which was not observed. Further investigation revealed a gene named LOC100731016, which, upon inquiry in the NCBI database, was identified as the *Nanog2* gene of the guinea pig. *Nanog* belongs to a gene family that includes *NANOG1*, its tandem duplication gene *NANOG2*, and the pseudogenes *NANOGP2-P11*. *NANOG* and *NANOG2* have been demonstrated to be functionally equivalent and to activate specific stem cell regulatory pathways [44]. Thus, we adopted LOC100731016 (*Nanog2*) as an EPI-specific marker for cell cluster characterization (Fig. 2B).

Additionally, *Dnmt3b* was highly expressed in EPI, indicating a hypermethylation state, akin to findings in human embryo studies [45]. By subgrouping based on sampling information and marker expressions such as *Tfcp2l1* and LOC100731016, the ICM group was also defined. The final classification comprised 7 cell classes: Totipotency (Toti), 8 C, Morula, ICM, TE, EPI, and PrE (Fig. 2A). Multiple marker gene enrichment and specific up- and down-regulated genes were also clearly differentiated (Fig. 2C-E).

### Guinea pig ZGA occurs at the 4–8 cell stage

Mammalian embryonic development commences with the formation of a fertilized egg following sperm-egg fusion. The initial stage of embryonic development is predominantly governed by a series of maternal factors supplied by the egg cell's cytoplasm. During this phase, the zygotic genome remains transcriptionally silent, enabling the zygote to maintain a state of totipotency. As the maternally deposited mRNA degrades, the zygotic genome starts to be transcribed, leading to the activation of the genome, the mobilization of embryonic gene products, and the clearance of maternal factors. This process, known as the maternal-to-zygotic transition (MZT), involves two key events: the degradation of maternal mRNA and the production of nascent mRNA, which signifies the activation of the zygotic genome (ZGA) [46, 47]. The coordination of these events is crucial. MZT is a critical phase in developmental studies, as the transitions that occur during this period are essential for embryonic patterning, which is heavily influenced by the actions of transcriptional activators [48–50].

Advances in histological sequencing have enabled the development of a sensitive and accurate method for sequencing in continuous time to assess the activation of the zygotic genome. We analyzed differences in gene expression between adjacent time periods for samples ranging from the zygote to the morula stage. The onset of ZGA is marked by significant changes in gene expression patterns and transcript levels, transitioning the embryo from a state of minimal transcription to one where thousands of genes are actively transcribed. This shift triggers a cascade of gene expression changes that influence subsequent cell fates [51].

Analysis of the differentially expressed genes (DEGs) revealed that the transcriptomic differences in guinea pig embryos at the zygote, 2 C, and 4 C stages are minimal, suggesting a transcriptionally quiescent state. In stark contrast, the transition from 4 C to 8 C stages exhibited significant transcriptomic alterations (Fig. 3A), aligning with the anticipated dynamics associated with the onset of ZGA. We delved into the gene expression patterns during the 4 C to 8 C developmental stages (Fig. 3B; Fig. S2,B). Gene Ontology (GO) enrichment analysis indicated that the up-regulated genes were predominantly involved in protein synthesis and transport (Fig. 3C). Conversely, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis highlighted the pentose phosphate pathway (Fig. 3D), which is crucial for providing precursors for nucleic acid synthesis. The enrichment of this pathway during the 4 C to 8 C stages suggests vigorous DNA replication and synthesis. Additionally, Gene Set Enrichment Analysis (GSEA) and GO confirmed that numerous synthesis reactions are actively occurring at this stage. (Fig. S2).

**Pseudotime analysis of preimplantation embryonic development in guinea pigs** provides insights into lineage specification, a process extensively investigated in humans and mice. In mice, lineage specification typically unfolds in two stages. The initial stage occurs at the morula stage, delineating the inner cell mass (ICM) and trophectoderm (TE) cell fates, with TE later giving rise to extraembryonic tissues like the placenta. The subsequent stage transpires in the blastocyst phase, where the ICM differentiates into epiblast (EPI) and primitive endoderm (PrE) lineages, with EPI eventually forming the individual and PrE contributing to the yolk sac [3, 52]. In human studies, however, two contrasting views have emerged. One perspective posits that TE, EPI, and



PrE are specified concurrently during human preimplantation development, suggesting a one-step process [41]. The alternative view advocates for an intermediate ICM state during lineage specification, akin to the two-step model observed in mice [53–55].

To investigate the lineage specification of preimplantation embryonic development in guinea pigs, we analyzed embryonic data spanning from the 8 C to the blastocyst stages. The pseudotime analysis revealed a bifurcating trajectory with two primary branches emerging from an initial branch at the first branching point. Subsequently, one of these branches further diverged into two additional branches at the second branching point. Mapping cell types onto these branches indicated that the initial branch was predominantly composed of 8 C and 16 C morula cells. The first branching point delineated the trophoblast (TE) and inner cell mass (ICM) branches. Following the proposed developmental timeline, the ICM branch subsequently differentiated into epiblast (EPI) and primitive endoderm (PrE) branches at the second branching point (Fig. 4A). Additionally, we overlaid the sampling time and pseudotime values onto the branches, identifying the first lineage separation in guinea pig preimplantation embryonic development around day E4.5, with the second lineage separation occurring around day E5.0 (Fig. 4B). To elucidate the key drivers of lineage specification, we visualized the top five genes with the highest contribution to the trajectory construction. Among these, *Argfx*, a gene homologous to a frame gene frequently observed in mammalian embryonic development studies from 8 C to morula stages and known to influence developmental regulation [56], exhibited a comparable expression pattern in guinea pigs. The projection of specific marker genes further validated the accuracy of the trajectory construction (Fig. S3; Fig. S4). Notably, the ICM branch in guinea pigs was found to contain a mixture of EPI cells, differing from mice where distinct ICM cell clusters can be identified based on significant molecular features. Similar to humans, guinea pigs exhibit a high degree of similarity between ICM and EPI cell clusters. Overall, the pseudotime analysis of our data offers a comprehensive view of the three-lineage specification in guinea pig preimplantation embryos. Furthermore, we depicted the cell proportion dynamics at various time points, aligning with the developmental trajectory (Fig. 4C).

To explore the gene expression dynamics during lineage specification, we ranked the top 1000 genes by ascending q-values and generated a heatmap depicting differential gene expression across pseudotime, aiming to pinpoint genes with significant expression shifts over time. These genes were grouped into six clusters based on their expression patterns (Fig. 4D). The heatmap clusters illustrated distinct gene expression profiles that evolved

with pseudotime. To delve deeper into pathway expression trends, we utilized Ucell to score relevant pathways, focusing on key processes such as oxidative phosphorylation, purine metabolism, and the VEGF pathway (Fig. 4E–G). Notably, the oxidative phosphorylation pathway exhibited elevated scores in ICM and EPI cells, aligning with observations of enrichment in preimplantation EPI cells across various species [57, 58]. These findings are congruent with prior research. Furthermore, the VEGF pathway showed higher scores in TE and PrE cells, which is consistent with its role in angiogenesis, as TE and PrE later contribute to the vascular-rich placenta and yolk sac, respectively. The observed expression patterns were congruent with the pathway enrichment analysis, underscoring the relevance of these pathways in early embryonic development.

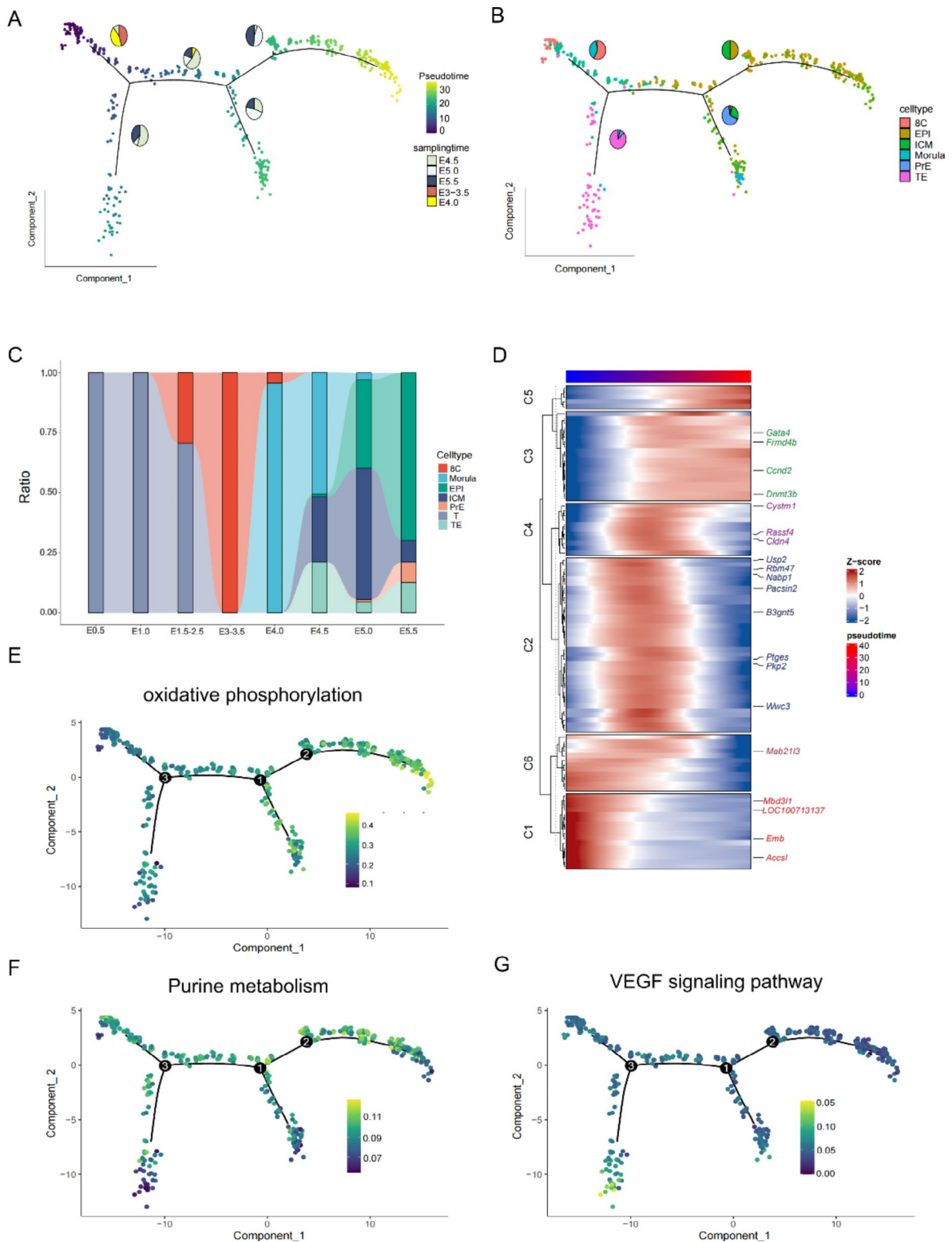
#### Functional characteristics of ICM, EPI and PrE

To delineate the molecular characteristics of each lineage, we employed hdWGCNA (high-dimensional Weighted Gene Co-expression Network Analysis) [34] to uncover gene expression patterns associated with lineage specification (Fig. 5A; Fig. S5B). We selected a soft threshold power of 7 to construct the co-expression network (Fig. S5A). Following high-dimensional gene co-expression analysis, we identified 19 functional modules (Fig. S5A–B; Fig. S5D), with particular attention to those associated with the second lineage segregation. Utilizing correlation-assisted enrichment module screening, we identified the most significant modules enriched for EPI, ICM, and PrE as 3, 6, and 13 modules, respectively (Fig. 5B; Fig. S5C). The hub genes of these modules were determined by selecting the top 25 hub genes for each. Visualization revealed that *Nanog2* was the central regulator in the 3 modules most significantly enriched for EPI, whereas *Sox17* was the predominant core regulator enriched in PrE (Fig. 5C), aligning with previous findings [59].

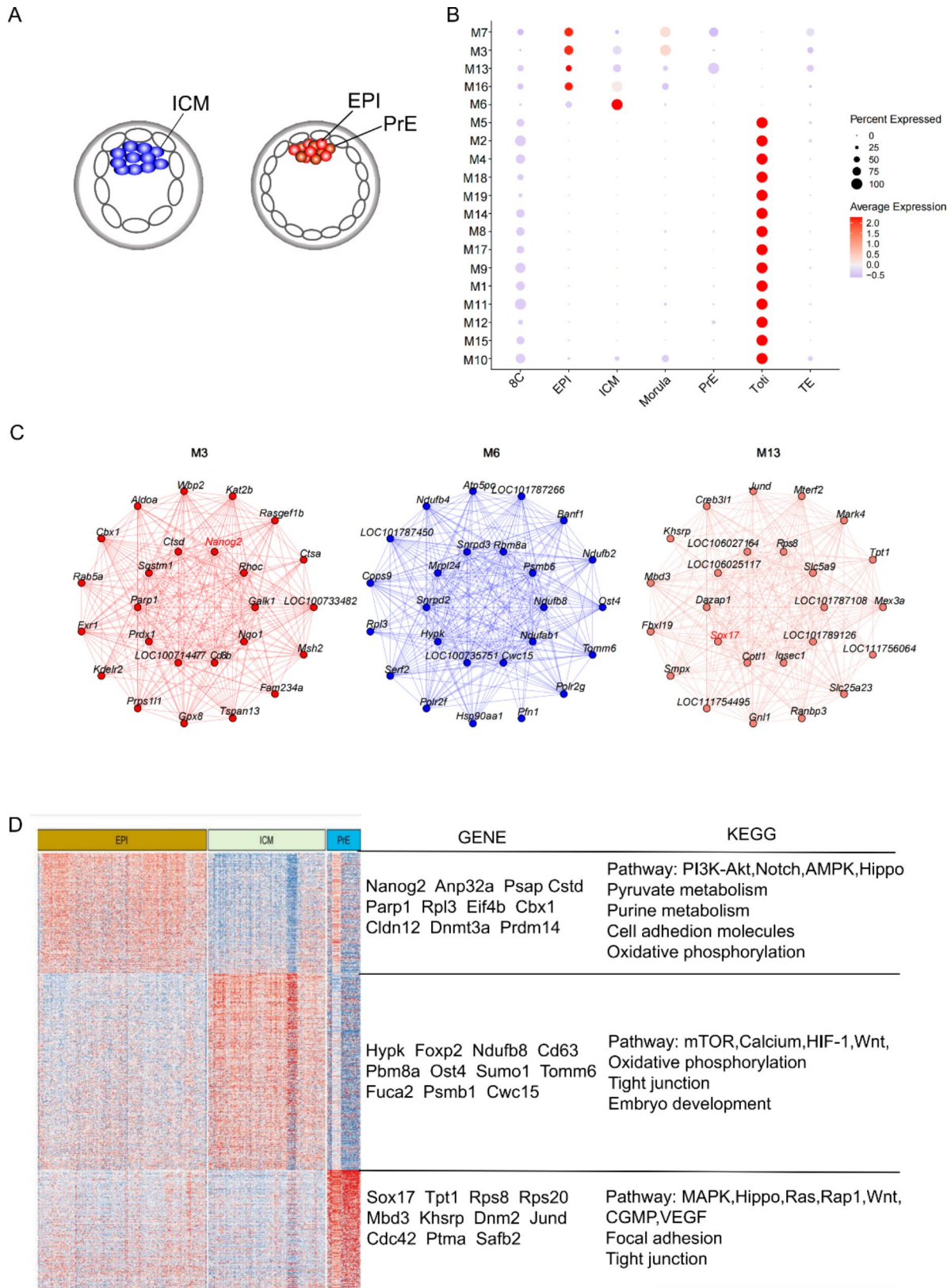
To understand the gene expression patterns associated with the second cell fate decision, we identified 1209 genes by intersecting the markers identified by hdWGCNA and the findmarker function. We then analyzed these genes using KEGG to plot and examine a heatmap. The analysis revealed that EPI cells primarily involve pathways such as Hippo, Notch, oxidative phosphorylation, embryonic development, and cell adhesion. In contrast, ICM cells were associated with embryonic development and oxidative phosphorylation, while PrE cells were enriched in pathways like Ras, Rap1, and VEGF. These cell types exhibited partially overlapping yet distinct pathway expression profiles (Fig. 5D).

#### Comparison of human and guinea pig EPI characteristics

It is now generally accepted that preimplantation EPI is characterized by a naïve state, with the potential



**Fig. 4** Two rounds of lineage separation in guinea pig preimplantation embryos **(A)** UMAP plot showing the developmental trajectory across different sampling times. **(B)** UMAP plot illustrating the developmental trajectory of 8-cell stage (8C), morula, trophectoderm (TE), inner cell mass (ICM), epiblast (EPI), and primitive endoderm (PrE) cells. **(C)** Bar plot representing the proportion of each cell type at different sampling times. **(D)** Heatmap displaying trends of differentially expressed genes as a function of inferred pseudotime. Genes in each row are grouped into six clusters based on their expression patterns. **(E-G)** These UMAP plots represent the distribution of pathway scores across different pseudotime points. The color gradient on the plot indicates the pathway scores, with dark blue representing low scores and light yellow representing high scores.



**Fig. 5** Identification of co-expressed gene modules associated with ICM, EPI, and PrE **(A)** Schematic representation of the second cell fate decision for preimplantation embryos. **(B)** Module activities in different clusters estimated using the hdWGCNA algorithm. **(C)** Co-expression plots for modules M3 (EPI), M6 (ICM), and M13 (PrE). **(D)** Differential gene heatmap and KEGG enrichment analysis results for EPI, ICM, and PrE

to produce a complete individual (but not to develop extra-embryonic tissues) and the ability to form chimeras [60, 61]. Also similar to humans, after implantation into the endometrium, both guinea pig and human EPIs show developmental events of cavitation and formation of a double-layered intervertebral disc [62]. EPI plays an important role in understanding the developmental mechanisms of pluripotent stem cells as well as the development of regenerative medicine. To compare the characteristics of EPI in human and guinea pig preimplantation embryos, we combined previously published human data with our guinea pig data. We focused on the dynamics of naïve genes and the EPI pluripotency transition process (EPST), which have been documented in humans [35]. Our analysis revealed that, compared to human EPST, guinea pigs lacked expression of *Il6r* and *Nr0b1*, showed reduced expression of *Utf1*, and exhibited increased expression of *Lifr* (Fig. 6C–D).

Furthermore, we performed PCA on the integrated data. The analysis indicated a.

clear separation of human and guinea pig cell data along the PC1 axis, highlighting species differences as the most significant factor (Fig. 6A). We identified 13,915 genes expressed in both species. Along the PC1 axis, we classified these genes into significantly positive and significantly negative categories. KEGG enrichment analysis revealed notable differences in pathways such as phosphatidylinositol metabolism, pyrimidine and purine metabolism, and oxidative phosphorylation (Fig. 6B). Along the PC2 and PC3 axes, the cellular projections of human and guinea pig overlapped, suggesting similar developmental trajectories during the EPI pluripotency transition process (Fig. 6A). These findings suggest both conservation and some differences in the development of human and guinea pig preimplantation EPI.

## Discussion

In our study, we systematically analyzed single-cell data from 62 guinea pig preimplantation embryos, encompassing 551 single cells, using single-cell sequencing technology to explore their developmental characteristics. Our analysis covered key developmental events from the zygote stage through to embryo implantation, including zygotic gene activation (ZGA), ICM and TE cell fate decisions, EPI and PrE cell fate decisions. We identified specific time windows for these critical events and observed corresponding gene expression changes. Additionally, we compared the similarities and differences between human and guinea pig preimplantation EPIs by integrating previously reported human embryonic data. Our findings offer fresh perspectives on guinea pig embryonic development and expand the range of species available for studying mammalian embryogenesis, thereby providing new options for comparative studies.

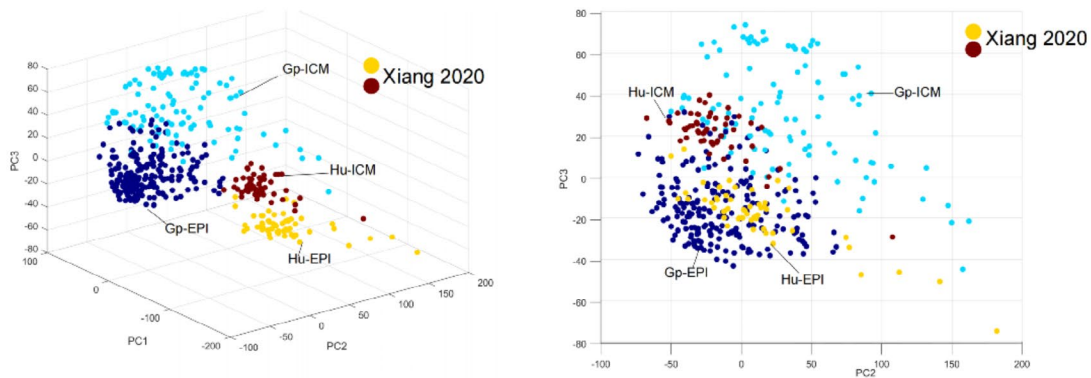
We employed single-cell RNA-seq for lineage tracing in guinea pig preimplantation embryos. Following QC, the data were subjected to dimensionality reduction and clustered into seven distinct cell types. The preimplantation EPI population in guinea pigs exhibits a unique core marker compared to other species, with high expression of the *Nanog2* gene, which is functionally similar to *Nanog*. Additionally, we observed that Totipotency (Toti) cells highly express *Sox2*, 8 C cells express *Zscan4*, and ICM cells express *Tfcp2l1*. In contrast, the TE population is characterized by high enrichment of *Krt8*, *Krt18*, *Gata3*, and *Cdx2*. The PrE signature marker *Sox17*, commonly found in human and mouse, is also highly enriched in the guinea pig PrE population. These findings resulted in high-resolution single-cell transcriptome profiles, providing detailed insights into the molecular signatures of different cell types during early embryonic development.

Embryonic development commences with a fertilized egg, initially governed primarily by maternal factors. Subsequently, there is a gradual decline in maternal deposition and the activation of the zygotic genome. The occurrence of ZGA is a conserved process across species; however, the timing varies significantly. For instance, human ZGA primarily occurs at the 8-cell stage, while in mice, it happens at the 2-cell stage, and in cattle, it is at the 8–16-cell stage [63]. Our findings indicate that guinea pig ZGA predominantly takes place during the 4-cell to 8-cell phase, aligning closely with the timing in humans. The enrichment of up-regulated genes during this ZGA window is associated with processes such as protein synthesis and transport, as well as ribose synthesis. The number of activated genes and the enriched functions collectively suggest that zygotic gene activation in guinea pig embryos occurs from the 4 C to 8 C stages.

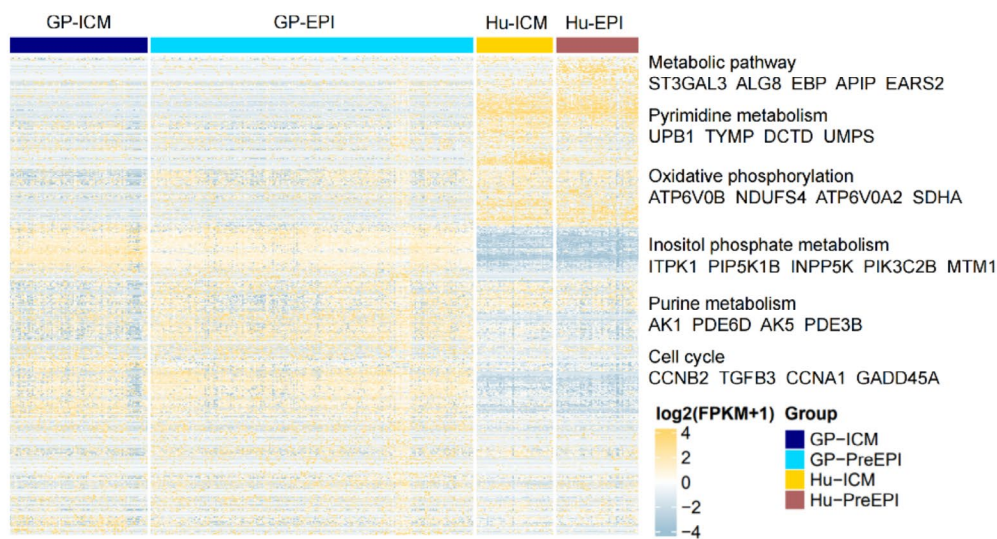
We conducted pseudotime analysis of transcriptome developmental trajectories in guinea pig embryos, which revealed a clear bifurcation into two distinct branches. Analysis of the embryo sampling points indicated that the ICM/TE cell fate decision occurs around E4.5, while the EPI/PrE cell fate decision takes place earlier, at E5.0. Despite the clear branching, we observed a high degree of similarity between early ICM and EPI, with significant overlap in trajectory branching. This challenge in distinguishing EPI/PrE lineages is also present in human preimplantation embryos. Our findings suggest that EPI may have emerged first, followed by PrE fate. In mice, the presence of *Fgf4* (secreted by EPI cells) is crucial for PrE fate decision [64]. A hypothesis regarding human EPI/PrE cell fate decision posits that ICM is closer to EPI at the transcriptome level, potentially leading to confusion between ICM and EPI [6]. Our data indicate that guinea pigs exhibit a similar pattern to humans in this aspect.

Furthermore, we examined the gene expression characteristics of the second lineage decision. Employing

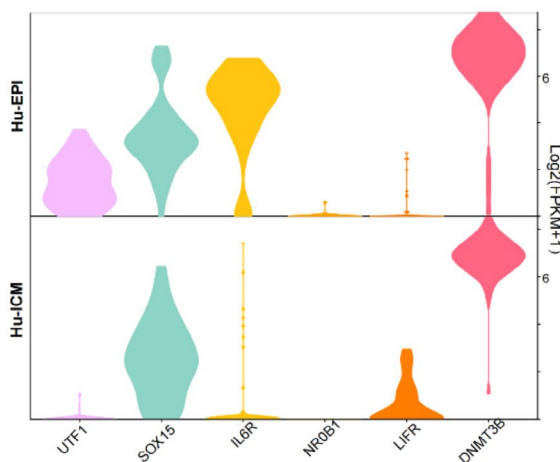
A



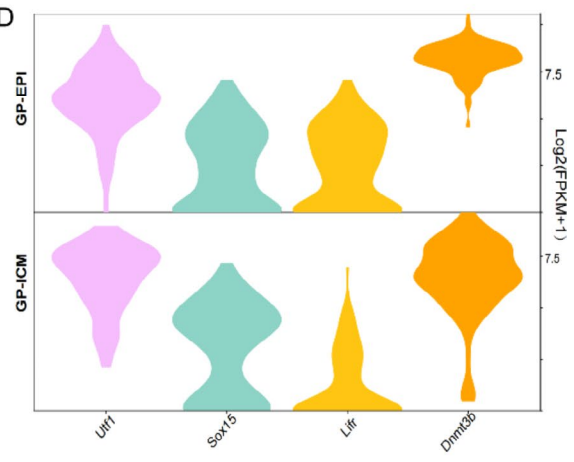
B



C



D



**Fig. 6** Comparison of preimplantation EPI in humans and guinea pigs **(A)** Three-dimensional Principal Component Analysis (PCA) of the EPI lineage, determined by the expressed genes among all groups of EPIs during development in humans (hu) and guinea pigs (GP). In total, 10,698 out of 13,915 annotated genes expressed in human and guinea pig cells (human, 106 cells; guinea pig, 315 cells) were used. **(B)** Heatmap of 952 genes that highly contributed to PC1 (greater than 0.8 standard deviation of PC1). **(C)** Violin plots showing expression levels of pluripotency genes during the EPI pluripotency transition in human embryos. **(D)** Violin plots showing expression levels of selected genes from (C) in guinea pig embryos

hdWGCNA, we identified functional modules most pertinent to ICM, EPI, and PrE. For instance, at the core of regulation for EPI and PrE are *Nanog2* and *Sox17*. *Nanog* and *Sox17* are key regulators in the cell fate determination of the EPI and PrE lineages in mice. During the specialization of the EPI lineage, *Nanog*, *Sox2*, and *Oct4* work synergistically, promoting the expression of *FGF4* while suppressing *Gata6* and *Fgf* signaling, thus driving the specification of the ICM lineage. In contrast, *Sox17* plays a crucial role in the specification of the PrE lineage. *FGF* signaling facilitates PrE formation by promoting *Gata6* expression and inhibiting *Nanog*, with *Sox17* acting as a downstream regulatory factor.

Given that *Nanog2* is functionally equivalent to *Nanog*, it is intriguing to explore whether the regulatory mechanism in guinea pig EPI/PrE follows a similar pathway. This question remains open and warrants further investigation.

Recognizing that the analyzed functional modules could be expressed across multiple cell populations, we intersected the most enriched module genes of ICM, EPI, and PrE with genes identified by the Findmarker function to pinpoint specific genes for these three cell types. We then enriched these for further analysis. The pathways enriched in the respective cells are similar and conserved compared to those in human and mouse, suggesting that the guinea pig offers certain advantages and is a rational choice as an experimental animal for studying human embryonic development.

Trophoblast (TE) cells play a crucial role in implantation. Traditionally, it was thought that all TE cells possess equal potential to interact with and invade the endometrium. However, recent research has revealed that polar TEs (pTE) and mural TEs (mTE), which have distinct localizations, exhibit different potentials [6, 41, 65]. In species such as mice, rats, pigs, and cows, embryos attach to the endometrium through mTE, while in humans and non-human primates, it is the pTE that facilitates attachment. In guinea pigs, similar to mice and rats, it is the mTE that is oriented towards attachment and invasion of the endometrium [66–69]. Despite these findings, the specific mechanisms by which mTE and pTE influence embryo attachment to the endometrium remain elusive, making this an intriguing area of inquiry.

The guinea pig genome available on NCBI have not been assembled to the level of individual chromosomes, which limits the resolution of genome assembly. Consequently, research on guinea pig chromatin accessibility, epigenetic modifications of transcription factor binding, and related aspects is constrained. The current study is predominantly confined to the transcriptome level, yet transcription factors and epigenetic modifications are crucial regulators during embryonic development. This represents a limitation of the study. Therefore,

constructing a more comprehensive and refined reference genome for the guinea pig is a critical step for advancing future research in this species.

## Conclusions

To summarize, the current research is the first to systematically establish a single-cell global transcriptional profile of guinea pig preimplantation embryos, which is characterized by a wide array of similarities with human embryos. Embryonic development is a pivotal topic in developmental biology and broader biological research. Ethical constraints limit the study of human embryos, and while mice have served as model organisms for embryonic development research, they exhibit differences from humans in certain aspects. The investigation of guinea pig preimplantation embryos offers a novel perspective and method for comprehending human preimplantation embryonic development. As a consequence, deciphering molecular mechanisms underlying inter-transcriptomic hallmarks correlated with early embryogenesis in guinea pigs might open up new possibilities for elaborating an excellent *ex vivo* embryological rodent-based model. The latter seems to be a highly suitable for exploration of both maternal-to-embryonic transition for controlling gene expression and developmental timeline arising from embryonic cell commitment/differentiation in humans. Moreover, this guinea pig model aimed at investigating experimental and applied embryology in humans might provide a powerful tool for transgenic and biomedical research with the use of such assisted reproductive technologies as cloning by somatic cell nuclear transfer and *in vitro* fertilization of oocytes either by their coincubation with capacitated spermatozoa or by their intracytoplasmic microinjection with single sperm cells.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10815-z>.

Supplementary Material 1

## Acknowledgements

We thank the members of the LiuJing lab for providing helpful discussions and guidances on this study.

## Author contributions

JL, DqP and ZZ designed and directed the project; JG and TxG designed and performed the study; The bioinformatic analysis was done by TxG; JpL and RxL participated in the construction of single-cell libraries; Animal feeding is handled by RpL; The embryo experimental part was done by JG and RxL, TxG wrote the manuscript.

## Funding

This work was supported by National Key R&D Program of China(2018YFE0204800), National Natural Science Foundation of China (32370791), Science and Technology Projects in Guangzhou (202201010510), Science and Technology Planning Project of Guangdong

Province (2023B1212060050, 2023B1212120009) and Basic Research Project of Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences (GIBHBRP23-02).

#### Data availability

Availability of data and materials Sequence data used by this study is available under Genome Sequence Archive (GSA; <https://bigd.big.ac.cn/gsa/>) with Bioproject ID PRJCA028188.

#### Declarations

##### Ethics approval and consent to participate

All procedures were conducted in accordance with the “Guiding Principles in the Care and Use of Animals(China)” and were approved by Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences and Laboratory Animal Welfare and Ethics Control Committee (IACUC). IACUC No. 2024008.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare no competing interests.

Received: 29 July 2024 / Accepted: 19 September 2024

Published online: 30 September 2024

#### References

- Gerri C, Menchero S, Mahadevaiah SK, Turner JMA, Niakan KK. Human embryogenesis: a comparative perspective. *Annu Rev Cell Dev Biol.* 2020;36:411–40.
- White MD, Zenker J, Bissiere S, Plachta N. Instructions for assembling the early mammalian embryo. *Dev Cell.* 2018;45(6):667–79.
- Chazaud C, Yamanaka Y, Pawson T, Rossant J. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell.* 2006;10(5):615–24.
- Rossant J, Tam PPL. Early human embryonic development: blastocyst formation to gastrulation. *Dev Cell.* 2022;57(2):152–65.
- Biondic S, Canizo J, Vandal K, Zhao C, Petropoulos S. Cross-species comparison of mouse and human preimplantation development with an emphasis on lineage specification. *Reproduction.* 2023;165(4):R103–16.
- Meistermann D, Bruneau A, Loubesac S, Reignier A, Firmin J, François-Campion V, Kilens S, Lelièvre Y, Lammers J, Feyeux M, et al. Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. *Cell Stem Cell.* 2021;28(9):1625–e16401626.
- Molè MA, Weberling A, Zernicka-Goetz M. Comparative analysis of human and mouse development: from zygote to pre-gastrulation. *Curr Top Dev Biol.* 2020;136:113–38.
- Liang L, He C, Lei M, Li S, Hao Y, Zhu H, Duan Q. Pathology of guinea pigs experimentally infected with a novel reovirus and coronavirus isolated from SARS patients. *DNA Cell Biol.* 2005;24(8):485–90.
- Orme IM. Mouse and guinea pig models for testing new tuberculosis vaccines. *Tuberculosis (Edinb).* 2005;85(1–2):13–7.
- Bowick GC, Fennewald SM, Elsom BL, Aronson JF, Luxon BA, Gorenstein DG, Herzog NK. Differential signaling networks induced by mild and lethal hemorrhagic fever virus infections. *J Virol.* 2006;80(20):10248–52.
- Dobbing J, Sands J. Growth and development of the brain and spinal cord of the guinea pig. *Brain Res.* 1970;17(1):115–23.
- Dobbing J, Sands J. Comparative aspects of the brain growth spurt. *Early Hum Dev.* 1979;3(1):79–83.
- Suzuki O, Koura M, Noguchi Y, Takano K, Yamamoto Y, Matsuda J. Optimization of superovulation induction by human menopausal gonadotropin in guinea pigs based on follicular waves and FSH-receptor homologies. *Mol Reprod Dev.* 2003;64(2):219–25.
- Mess A. The guinea pig placenta: model of placental growth dynamics. *Placenta.* 2007;28(8–9):812–5.
- Kaufmann P, Black S, Huppertz B. Endovascular trophoblast invasion: implications for the pathogenesis of intrauterine growth retardation and preeclampsia. *Biol Reprod.* 2003;69(1):1–7.
- Carter AM, Enders AC, Jones CJ, Mess A, Pfarrer C, Pijnenborg R, Soma H. Comparative placentation and animal models: patterns of trophoblast invasion – a workshop report. *Placenta.* 2006;27(Suppl A):S30–33.
- Suzuki O, Ogura A, Asano T, Noguchi Y, Yamamoto Y, Oike M. Development of preimplantation guinea-pig embryos in serum-free media. *Reprod Fertil Dev.* 1993;5(4):425–32.
- Dorsch MM, Glage S, Hedrich HJ. Collection and cryopreservation of preimplantation embryos of *Cavia porcellus*. *Lab Anim.* 2008;42(4):489–94.
- Hribal R, Guenther A, Rübensam K, Jewgenow K. Blastocyst recovery and multifactorial gene expression analysis in the wild guinea pig (*Cavia aperea*). *Theriogenology.* 2016;86(5):1299–307.
- Morrison JL, Botting KJ, Darby JRT, David AL, Dyson RM, Gatford KL, Gray C, Herrera EA, Hirst JJ, Kim B, et al. Guinea pig models for translation of the developmental origins of health and disease hypothesis into the clinic. *J Physiol.* 2018;596(23):5535–69.
- Kapoor A, Matthews SG. Short periods of prenatal stress affect growth, behaviour and hypothalamo-pituitary-adrenal axis activity in male guinea pig offspring. *J Physiol.* 2005;566(Pt 3):967–77.
- Drouin G, Godin JR, Pagé B. The genetics of vitamin C loss in vertebrates. *Curr Genomics.* 2011;12(5):371–8.
- Coker SJ, Dyson RM, Smith-Díaz CC, Vissers MCM, Berry MJ. Effects of low vitamin C intake on fertility parameters and pregnancy outcomes in guinea pigs. *Nutrients.* 2023;15(19).
- Maxeiner S, Gebhardt S, Schweizer F, Venghaus AE, Krasteva-Christ G. Of mice and men - and guinea pigs? *Ann Anat.* 2021;238:151765.
- Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using smart-seq2. *Nat Protoc.* 2014;9(1):171–81.
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37(8):907–15.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–9.
- Butler A, Hoffman P, Smibert P, Papalexis E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* 2018;36(5):411–20.
- Ito K, Murphy D. Application of ggplot2 to pharmacometric graphics. *CPT Pharmacometrics Syst Pharmacol.* 2013;2(10):e79.
- Silva TC, Colaprico A, Olsen C, D’Angelo F, Bontempi G, Ceccarelli M, Noushmehr H. TCGA Workflow: analyze cancer genomics and epigenomics data using bioconductor packages. *F1000Res.* 2016;5:1542.
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innov (Camb).* 2021;2(3):100141.
- Walter W, Sánchez-Cabo F, Ricote M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics.* 2015;31(17):2912–4.
- Zhang P, Pei S, Zhou G, Zhang M, Zhang L, Zhang Z. Purine metabolism in lung adenocarcinoma: a single-cell analysis revealing prognostic and immunotherapeutic insights. *J Cell Mol Med.* 2024;28(8):e18284.
- Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Methods.* 2023;3(6):100498.
- Xiang L, Yin Y, Zheng Y, Ma Y, Li Y, Zhao Z, Guo J, Ai Z, Niu Y, Duan K, et al. A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature.* 2020;577(7791):537–42.
- Jalalvand AR, Roushani M, Goicoechea HC, Rutledge DN, Gu HW. MATLAB in electrochemistry: a review. *Talanta.* 2019;194:205–25.
- Niakan KK, Eggan K. Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Dev Biol.* 2013;375(1):54–64.
- Wicklow E, Blij S, Frum T, Hirate Y, Lang RA, Sasaki H, Ralston A. HIPPO pathway members restrict SOX2 to the inner cell mass where it promotes ICM fates in the mouse blastocyst. *PLoS Genet.* 2014;10(10):e1004618.
- Frum T, Murphy TM, Ralston A. HIPPO signaling resolves embryonic cell fate conflicts during establishment of pluripotency in vivo. *Elife.* 2018;7.

40. Posfai E, Petropoulos S, de Barros FRO, Schell JP, Jurisica I, Sandberg R, Lanner F, Rossant J. Position- and Hippo signaling-dependent plasticity during lineage segregation in the early mouse embryo. *Elife*. 2017;6.
41. Petropoulos S, Edsgård D, Reinius B, Deng Q, Panula SP, Codeluppi S, Plaza Reyes A, Linnarsson S, Sandberg R, Lanner F. Single-cell RNA-Seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell*. 2016;165(4):1012–26.
42. Zhi M, Zhang J, Tang Q, Yu D, Gao S, Gao D, Liu P, Guo J, Hai T, Gao J, et al. Generation and characterization of stable pig pregastrulation epiblast stem cell lines. *Cell Res*. 2022;32(4):383–400.
43. Nakamura T, Okamoto I, Sasaki K, Yabuta Y, Iwatani C, Tsuchiya H, Seita Y, Nakamura S, Yamamoto T, Saitou M. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature*. 2016;537(7618):57–62.
44. Eberle I, Pless B, Braun M, Dingermann T, Marschalek R. Transcriptional properties of human NANOG1 and NANOG2 in acute leukemic cells. *Nucleic Acids Res*. 2010;38(16):5384–95.
45. Zhu P, Guo H, Ren Y, Hou Y, Dong J, Li R, Lian Y, Fan X, Hu B, Gao Y, et al. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat Genet*. 2018;50(1):12–9.
46. Newport J, Kirschner M. A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. *Cell*. 1982;30(3):687–96.
47. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development*. 2009;136(18):3033–42.
48. Jukam D, Shariati SAM, Skotheim JM. Zygotic genome activation in vertebrates. *Dev Cell*. 2017;42(4):316–32.
49. Schulz KN, Harrison MM. Mechanisms regulating zygotic genome activation. *Nat Rev Genet*. 2019;20(4):221–34.
50. Schultz RM. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Hum Reprod Update*. 2002;8(4):323–31.
51. Aoki F. Zygotic gene activation in mice: profile and regulation. *J Reprod Dev*. 2022;68(2):79–84.
52. Yao C, Zhang W, Shuai L. The first cell fate decision in pre-implantation mouse embryos. *Cell Regen*. 2019;8(2):51–7.
53. <The selection arena in early human blastocysts resolves the pluripotent. pdf>
54. Singh M, Widmann TJ, Bansal V, Cortes JL, Schumann GG, Wunderlich S, Martin U, Garcia-Canadas M, Garcia-Perez JL, Hurst LD et al. The selection arena in early human blastocysts resolves the pluripotent inner cell mass. *bioRxiv*. 2019;318329.
55. Stirparo GG, Boroviak T, Guo G, Nichols J, Smith A, Bertone P. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development*. 2018;145(3).
56. Maeso I, Dunwell TL, Wyatt CD, Marlétaz F, Vetó B, Bernal JA, Quah S, Irimia M, Holland PW. Evolutionary origin and functional divergence of totipotent cell homeobox genes in eutherian mammals. *BMC Biol*. 2016;14:45.
57. Niu Y, Sun N, Li C, Lei Y, Huang Z, Wu J, Si C, Dai X, Liu C, Wei J et al. Dissecting primate early post-implantation development using long-term in vitro embryo culture. *Science*. 2019;366(6467).
58. Liu T, Li J, Yu L, Sun HX, Li J, Dong G, Hu Y, Li Y, Shen Y, Wu J, et al. Cross-species single-cell transcriptomic analysis reveals pre-gastrulation developmental differences among pigs, monkeys, and humans. *Cell Discov*. 2021;7(1):8.
59. Liu L, Liu C, Quintero A, Wu L, Yuan Y, Wang M, Cheng M, Leng L, Xu L, Dong G, et al. Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity. *Nat Commun*. 2019;10(1):470.
60. Masaki H, Kato-Itoh M, Umino A, Sato H, Hamanaka S, Kobayashi T, Yamaguchi T, Nishimura K, Ohtaka M, Nakanishi M, et al. Interspecific in vitro assay for the chimera-forming ability of human pluripotent stem cells. *Development*. 2015;142(18):3222–30.
61. Zhou J, Hu J, Wang Y, Gao S. Induction and application of human naive pluripotency. *Cell Rep*. 2023;42(4):112379.
62. Pfeffer PL. Alternative mammalian strategies leading towards gastrulation: losing polar trophoblast (Rauber's layer) or gaining an epiblast cavity. *Philos Trans R Soc Lond B Biol Sci*. 2022;377(1865):20210254.
63. Xie D, Chen CC, Ptaszek LM, Xiao S, Cao X, Fang F, Ng HH, Lewin HA, Cowan C, Zhong S. Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. *Genome Res*. 2010;20(6):804–15.
64. Regin M, Essahib W, Demtschenko A, Dewandre D, David L, Gerri C, Niakan KK, Verheyen G, Tournaye H, Sterckx J, et al. Lineage segregation in human pre-implantation embryos is specified by YAP1 and TEAD1. *Hum Reprod*. 2023;38(8):1484–98.
65. Liu D, Chen Y, Ren Y, Yuan P, Wang N, Liu Q, Yang C, Yan Z, Yang M, Wang J, et al. Primary specification of blastocyst trophectoderm by scRNA-seq: new insights into embryo implantation. *Sci Adv*. 2022;8(32):eabj3725.
66. Lee KY, DeMayo FJ. Animal models of implantation. *Reproduction*. 2004;128(6):679–95.
67. Wang H, Dey SK. Roadmap to embryo implantation: clues from mouse models. *Nat Rev Genet*. 2006;7(3):185–99.
68. James JL, Carter AM, Chamley LW. Human placentation from nidation to 5 weeks of gestation. Part I: what do we know about formative placental development following implantation? *Placenta*. 2012;33(5):327–34.
69. Hertig AT, Rock J, Adams EC. A description of 34 human ova within the first 17 days of development. *Am J Anat*. 1956;98(3):435–93.

## Publisher's note

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