RESEARCH ARTICLE SUMMARY

DEVELOPMENTAL BIOLOGY

Dissecting primate early post-implantation development using long-term in vitro embryo culture

Yuyu Niu*†, Nianqin Sun†, Chang Li†, Ying Lei†, Zhihao Huang, Jun Wu, Chenyang Si, Xi Dai, Chuanyu Liu, Jingkuan Wei, Longqi Liu, Su Feng, Yu Kang, Wei Si, Hong Wang, E. Zhang, Lu Zhao, Ziwei Li, Xi Luo, Guizhong Cui, Guangdun Peng, Juan Carlos Izpisúa Belmonte*, Weizhi Ji*, Tao Tan*

INTRODUCTION: The period from peri-implantation to gastrulation is critical for mammalian embryogenesis. During this time, connections between embryonic and maternal tissues are set up, and the primary germ layers and body plan are established. There is a substantial gap in our knowledge of early human postimplantation development because of technological limitations and ethical considerations. To extend the study of human embryogenesis to the postimplantation period, an in vitro culture system has been established that extends human blastocyst development to the pregastrulation stage (up to12 days) after fertilization, and the molecular and cellular events are revealed.

RATIONALE: With the general prohibition of growing human embryos beyond 14 days, closely related surrogate species can be examined. In addition, improvements are needed for primate embryo culture to support extended growth periods. The establishment of an in vitro culture system that enables the development of primate embryos beyond the implantation period provides an accessible way to study molecular and cellular mechanisms that underlie postimplantation development, including gastrulation.

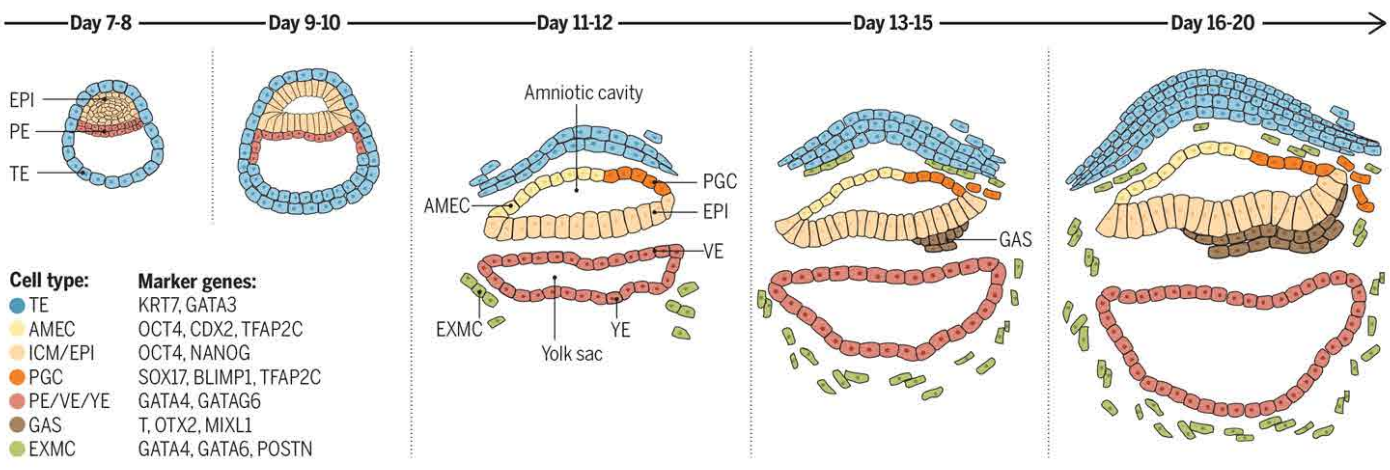
RESULTS: In this study, we have modified a human embryo in vitro culture protocol that enables the development of cynomolgus monkey (long-tailed macaque) embryos to develop up to 20 days after fertilization. The cultured cynomolgus embryos recapitulated key primate in vivo morphogenetic events, including amniotic and yolk sac cavitation, embryonic and extraembryonic lineage specification, specification of primordial germ-like cells (PGCLCs), and primitive streak cells. We demonstrated that the amniotic lumenogenesis is accompanied by the polarization of the epiblast (EPI); however, the polarization of PE was not observed during yolk sac cavitation. We used single-cell RNA-sequencing to delineate the developmental trajectories of EPI, trophoblast, and primitive endoderm (PE) lineages. We observed that accompanying the transition from the naïve to primer state, the metabolic mode of oxidative phosphorylation is no longer used in the EPI cells. Furthermore, the trophoblast differentiates in a stepwise manner, and expression of the trophoectoderm marker CDX2 decreases rapidly after day 11 in the trophoblast but maintains in the amniotic epithelium cells. In addition, we identified two types of PE lineage. Coordinated interactions were observed among EPI, trophoblast, PE, and extraembryonic mesenchyme cells during the postimplantation period. Furthermore, we showed that PGCLCs specified in vitro are similar to early-stage PGCs in vivo. Using the single-cell assay for transposase-accessible chromatin followed by sequencing (scATAC-seq), we also identified EPI, trophoblast, PE, and EXMC lineages. Last, scATAC-seq revealed that distal regions of chromatin in the EPI lineage exhibited higher accessibility than in other cell types.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aaw5754

CONCLUSION: In this study, we show that monkey embryos show robust development beyond 14 days after fertilization, surviving until day 20 without support from maternal tissue. We also provide insights into the transcriptional programs and chromatin dynamics that underlie monkey post-implantation development. Our system provides a platform to analyze molecular and cellular dynamics during primate early development. Last, our data may help guide the development of improved differentiation protocols for primate pluripotent stem cells. ■

The full list of author affiliations is available online.
*Corresponding author. Email: tant@lpbr.cn.com (T.T.); niuyy@lpbr.cn (Y.N.); wji@lpbr.cn (W.J.); belmonte@salk.edu (J.C.I.B.)
†These authors contributed equally to this work.
Cite this article as: Y. Niu *et al.*, *Science* **366**, eaaw5754 (2019).
DOI: 10.1126/science.aaw5754



Monkey embryos cultured in vitro recapitulate primate postimplantation embryogenesis in vivo. Scheme of monkey postimplantation embryogenesis cultured in vitro. ICM, inner cell mass; EPI, epiblast; PE, primitive endoderm; TE, trophectoderm; AMEC, amniotic epithelium cell; PGC, primordial germ cell; VE, visceral endoderm; YE, yolk-sac endoderm; EXMC, extra-embryonic mesenchyme cell; and GAS, gastrulating cell.

in PE cells (Fig. 6I). DLL3, a NOTCH pathway ligand, was highly expressed in EPI, and the receptor NOTCH2 was highly expressed in EXMC (Fig. 6I). Because PE (VE/YE), EXMC, and EPI may form a niche coordinating-lineage specification, we determined potential cellular communication interactions among PE (VE/YE), EXMC, EPI, and TE using public ligand-receptor databases (Materials and methods). EXMC and PE (VE/YE) frequently interacted with other cell types through ligand expression, and EPI cells mainly expressed receptors to receive the ligand from PE (VE/YE), EXMC, and TE (fig. S7D). For example, insulinlike growth factor 2 (IFG2) and granulin (GRN) were enriched in EXMC and PE (VE/YE), respectively. TGFBR3 was specifically expressed in EPI (fig. S7D). Overall, our study suggests the specific expression spectrum of ligands and receptors and corresponding interactions within PE (VE/YE), EXMC, EPI, and TE.

Transcriptional regulation of PGCLCs specification

To investigate potential transcriptional mechanisms underlying cynomolgus PGCLCs specification, PGCLC candidates (expressing SOX17, TFAP2C, and PRDM1/ BLIMP1 but not SOX2) (fig. S8A) (*14*) were compared with EPI cells because our data indicate that PGCLCs may originate from the amnion. Pseudotime analysis revealed two types of PGCLCs, one clustered with EPI cells (PGCLC-EPI) and one that did not (PGCLC) (fig. S8B). GO term analysis revealed that genes enriched in PGCLC-EPI cells were involved in ribosome biogenesis, whereas genes enriched in PGCLC cells were related to responses to fibroblast growth factor and cell-cell signaling by WNT (fig. S8C and table S5). These findings suggest heterogeneity in PGCLC population during development in vitro.

Genes differentially expressed between PGCLCs and the EPI included enrichment of the GO terms cell substrate adhesion and extracellular structure organization in PGCLCs (fig. S8D and table S6), implying that migration is a characteristic of PGCs, which is in agreement with expression of EMT-related genes in cynomolgus PGCs in vivo (*14*). We examined the correlation between PGCLCs and in vivo PGCs (early and late stage) (*14*). PGCLCs exhibited close correlation with in vivo early-stage PGCs (fig. S8E). Thus, PGCLCs specified in vitro exhibited a transcriptome similar to that of early in vivo PGCs.

Discussion

Gastrulation in primates involves extensive remodeling of the embryo's transcriptional landscape to facilitate the formation of the body plan (*1*, *7*). However, information about postimplantation development in primates remains limited because of ethical considerations, technical limitations, and high research costs. In vitro culture systems have been developed to study early postimplantation development in humans (*11*, *12*). Here, we modified the human embryo culture system for cynomolgus embryos, enabling them to develop in vitro for up to 20 days and recapitulate the specification of embryonic and extraembryonic lineages, the cavitation of amnion and yolk sac, PS formation, and PGC specification. The structure of the embryos, especially amnion and yolk sac cavity, varied among cultured embryos, which was not observed from in vivo derived embryos. This discrepancy can potentially be attributed to different microenvironments, such as mechanical and biochemical cues, between in vivo and in vitro. The structure of cultured embryos eventually collapsed around 20 days, which highlights the need to further improve culture parameters to enable more faithful and extended monkey embryo development in vitro.

Single-cell -omics analyses of the cultured cynomolgus embryos revealed molecular details of how the EPI, PE, TE, and EXMC lineage are specified. We found that cynomolgus PE cells differentiated into putative visceral and yolk sac endoderm and also EXMCs, during which PE cells undergo the EMT, regulated by the TGFβ signaling pathway. In mice, PE differentiate into parietal and visceral endoderm. Then the latter further differentiates into distal and anterior visceral endoderm (DVE/AVE) (*10*). Whether monkey PE cells follow a similar developmental path warrants further studies. We saw evidence of the metabolic state transition from early- to late-stage EPI, accompanying the switch from naïve to primed pluripotency. This is consistent with the notion that interactions between the metabolome and histone modifications drive the metabolic switch from naïve to primed pluripotency in ESCs (*44*). Our results support the notion that PGCs originate from the amnion in vivo (*14*). The transcriptional profiles of a small group of the PGCLCs were similar to that of the EPI, suggesting that these PGCLCs may also have been derived from the EPI.

One of the advantages of in vitro monkey embryo culture is that it provides an accessible platform to gain in-depth insights into the dynamic molecular and cellular changes during primate early development on the basis of the improvement of culture conditions that could support more homogenous and extended embryo development. Combining the embryo culture system with live-cell imaging, lineage tracing, signaling pathway perturbation through activators and inhibitors, and single-cell -omics analyses will help determine key factors and pathways underlying the specification and development of different lineages and thereby guiding the development of improved differentiation protocols from primate PSCs, including humans.

We have cultured primate embryos beyond day 14 and provide a comprehensive analysis of the transcriptional features of postimplantation embryos using scRNA-seq and scATAC-seq. This culture system provides a platform with which to dissect the mechanisms underlying primate embryonic development.

Materials and methods

Embryo in vitro culture

The monkey blastocysts were collected at 6 to 7 days post fertilization. Zona pellucida of blastocyst was removed by exposure to hyaluronidase from bovine testes about 30s. Zona-freed blastocysts were seeded on 8-well µ-plates (ibidi 80826, ibidi) with pre-equilibrated in vitro culture medium 1 (IVC1). After embryos attached to the well, half of the IVC1 was exchanged with in vitro culture medium 2 (IVC2). Thereafter, half of the medium were replaced daily with fresh IVC2, and monkey embryos were harvested at the indicated stage. Embryo culture was performed at 37°C in 5% CO_2. All chemicals were from Sigma Chemicals unless otherwise stated.

IVC1: DMEM/F12 (11320-033, Thermo Fisher Scientific) supplemented with 20% (vol/vol) heat-inactivated fetal bovine serum (FBS) (35-076-CV, Corning), 2mM L-glutamine (25030, Thermo Fisher Scientific), Penicillin (25 units/ml) streptomycin (25 µg/ml) (15070-063, Thermo Fisher Scientific), 1× ITS-X (51500-056, Thermo Fisher Scientific), 8 nM β-estradiol, 200 ng/ml progesterone and 25 µM N-acetyl-L-cysteine.

IVC2: 20% (vol/vol) heat-inactivated FBS of IVC1 was replaced with 30% (vol/vol) Knock-Out Serum Replacement (10828010, Thermo Fisher Scientific).

Single cell collection and embryo laser capture microdissection

After washing with phosphate buffered saline (PBS) (MA0008, meilunbio), the embryo was

dehydrated in sucrose (57-50-1, meilunbio) solution, embedded in OCT, frozen and stored at -80°C. Then the samples were prepared as cryosections with 8 µm thickness on pre-treated glass slides (CA1005, CITOTEST), and air-dried for 1 hour. After permeabilized in PBST (PBS with 0.3% Triton X-100) (0694, Amresco) for 30 min at room temperature, samples were blocked with 3% (w/v) BSA and 10% (v/vol) FBS (04-001-1ACS, Biological Industries) in PBS for overnight at 4°C. Slides were then incubated with primary antibodies overnight at 4°C. After wash, fluorescence-conjugated secondary antibodies, and 4′,6-diamidino-2-phenylindole (DAPI) were incubated with the slides in dark at 25°C for 2 hours. Images were taken using Leica TCS SP8 confocal microscope (Leica). For antibodies used, see table S7.

scRNA-seq

Plate-based single-cell cDNA synthesis, amplification and library construction were performed as previously described (*43*, *46*). Briefly, single cell was placed into the lysis buffer by mouth pipette. The reverse transcription reactions and pre-amplification were performed using SuperScript II (18064-071, Invitrogen), and KAPA HiFi HotStart ReadyMix (KK2602, KAPA biosystems,), respectively. For library construction, the cDNA was fragmented by Tn5 transposase (BGE005, BGI) mix at 55°C for 15 min. After that, 0.25% SDS was added to each well to stop the fragmentation, followed by polymerase chain reaction (PCR) amplification using KAPA HiFi HotStart polymerase mix (KK2602, KAPA biosystems) and primer sets with 25 cycles of denaturation. Then the libraries were purified with AMPure XP magnetic beads (A63881, Beckman Coulter), and were circularized by incubating with T4 DNA ligase (EL0011, Fermentas) at 37°C for 1 hour. After removal of linear DNA, the libraries were then purified by PEG32 (MP07123, CG) and sequenced on BGISEQ-500 (BGI).

Fig. 6. Transcription landscape of monkey peri- and postimplantation embryos. (**A**) t-distributed stochastic neighbor embedding (t-SNE) plot of cells at representative stages (days 9, 11, 13, 15, 17, 19, and 20). Cells were identified as epiblast (EPI), primitive endoderm (PE), trophectoderm (TE), and extraembryonic mesenchyme (EXMC) cells. (**B**) Expression of lineage-specific marker genes exhibited on t-SNE plots. A gradient of gray, yellow, and red indicates low to high expression. (**C**) t-SNE plot of TE cells at seven time points. (**D**) (Left) Pseudotime construction of single TE cells colored according to embryonic stage. (Right) Expression patterns of CDX2, TCEAL4, GCM1, GATA3, SOX2, and GATA4 exhibited on pseudotime construction. Dot-size gradient indicates low to high expression. (**E**) (Left) t-SNE plot of PE cells at seven time points, revealing two cell types, with (right) expressions of cell-type specific genes. (**F**) Gene Ontology (GO) term analysis: Cluster 1, enriched for lipid metabolism and transport; cluster 2, enriched for transcription and protein synthesis. (**G**) t-SNE plot of EPI cells at seven time points. Cells are designated as EPI-A, EPI-B, EPI-C, and gastrulating cell (Gast). (Inset) Single cells colored according to embryonic stage. (**H**) GO term analysis of classified EPI cells. (**I**) Violin plot of FGF/WNT/NOTCH signaling components expression in the various lineages.

IF analysis

Monkey in vivo and in vitro postimplantation embryos were harvested and fixed in 4% paraformaldehyde in PBS for 30 min at room temperature. Then embryos were washed with PBS, and then day 9 and day 10 embryos were stained directly. Other stages of embryos were dehydrated in sucrose (57-50-1, meilunbio) solution, embedded in OCT, frozen and stored at -80°C. Then the samples were prepared as cryosections with 8 µm thickness on pre-treated glass slides (CA1005, CITOTEST), and air-dried for 1 hour. After permeabilized in PBST (PBS with 0.3% Triton X-100) (0694, Amresco) for 30 min at room temperature, samples were blocked with 3% (w/v) BSA and 10% (v/vol) FBS (04-001-1ACS, Biological Industries) in PBS for overnight at 4°C. Slides were then incubated with primary antibodies overnight at 4°C. After wash, fluorescence-conjugated secondary antibodies, and 4′,6-diamidino-2-phenylindole (DAPI) were incubated with the slides in dark at 25°C for 2 hours. Images were taken using Leica TCS SP8 confocal microscope (Leica). For antibodies used, see table S7.

scRNA-seq

Plate-based single-cell cDNA synthesis, amplification and library construction were performed as previously described (*43*, *46*). Briefly, single cell was placed into the lysis buffer by mouth pipette. The reverse transcription reactions and pre-amplification were performed using SuperScript II (18064-071, Invitrogen), and KAPA HiFi HotStart ReadyMix (KK2602, KAPA biosystems,). respectively. For library construction, the cDNA was fragmented by Tn5 transposase (BGE005, BGI) mix at 55°C for 15 min. After that, 0.25% SDS was added to each well to stop the fragmentation, followed by polymerase chain reaction (PCR) amplification using KAPA HiFi HotStart polymerase mix (KK2602, KAPA biosystems) and primer sets with 25 cycles of denaturation. Then the libraries were purified with AMPure XP magnetic beads (A63881, Beckman Coulter), and were circularized by incubating with T4 DNA ligase (EL0011, Fermentas) at 37°C for 1 hour. After removal of linear DNA, the libraries were then purified by PEG32 (MP07123, CG) and sequenced on BGISEQ-500 (BGI).

scATAC-Seq

Plate-based single-cell transposition reactions were performed with transposase mixture at 37°C for 15 min with agitation at 300 rpm, following with a modified FAST-ATAC method (*47*). The release buffer was added and the reaction was maintained at 50°C for 15 min. Then plasmid DNA (30 ng) was added as the [illegible]DNA to the mixture. Afterwards, the [illegible]as purified with Ampure XP beads. Then the DNA was pre-amplified with NEBNext High-Fidelity 2× PCR Master Mix (M0541, New England Biolabs) and transposase adapters. The pre-amplified DNA was purified with Ampure XP beads and used for libraries construction as previously described (*48*). Briefly, DNA was amplified for 14 cycles using the NEBNext High-Fidelity 2× PCR Master Mix and barcode primers. Then the libraries were size-selected with Ampure XP beads for fragments between 150 and 700 base pairs (bp) and sequenced on BGISEQ-500.

scRNA-seq and scATAC-seq data pre-processing

For RNA-seq data, adapters and low quality reads (*N* rate > 0.2) were removed by Cutadapt (v1.15) (*49*). Then raw reads were mapped to Macaca fascicularis (Macaca_fascicularis_5.0) genome by STAR (v2.5.3) program (*50*). We calculated the transcripts per million mapped reads (TPM) as expression level using RSEM v1.3.0 with default parameters (*51*). The single cells with mapped reads >1 million and more than 2000 genes with TPM value >1 were used to further analysis.

For ATAC-seq data, the inherent sequence of Tn5 for each cell was removed using cutadapt (v1.16) (*49*) and aligned to the cynomolgus reference genome (Macaca_fascicularis_5.0) using bowtie2 (v2.2.5) with the parameter –X 2000 (*52*). SAMtools (v1.9) (*53*) was used to filter reads for alignment quality of >Q30 and Picard tools (v2.6.0) (Broad Institute; http://broadinstitute.github.io/picard) was used to remove duplicate reads. Cells with usable fragments under 10000 and promoter regions (500 bp around transcriptional start site) with ratio of fragments under 10% were filtered out. The fragments were aggregated and reference peak were called using Sambamba (v0.6.6) (*54*) and MACS2 (v2.1.2) (*55*) respectively. Then the number of fragments per reference peak per cell was counted using chromVAR (v1.4.0) (*56*) to construct the matrix of fragment counts in peaks X.

Cell clustering and differentially expressed gene analysis

Cell clustering and differential gene expression analysis were performed by Seurat (v2.3.4), a package in R (*57*). Principal components analysis (PCA) was performed to select principal components for clusters finding based on a jack straw method. Then graph-based clustering approaches were applied to identify cell clusters and t-distributed stochastic neighbor embedding (t-SNE) was used for visualization of distance between cells in the reduced 2D space.

t-SNE analysis of scATAC-seq data

chromVAR (*56*) were used to perform t-SNE analysis and Cicero (v1.0.11) (*58*) were used to define clusters in t-SNE embedding.

Transcription factor motifs analysis and differentially accessible sites identification

TF motif deviations were obtained from "deviationScores" function in chromVAR (*56*). To identify differentially accessible sites, we adopt the method in previously report (*59*). We randomly sampled 20 cells from each of the clusters identified above to generate the reference panel. Then we implement "binomialff" test using Monocle 2 package (*60*, *61*) to get differentially accessible sites at a 1% FDR threshold (Benjamini-Hochberg method). We calculated specificity scores for differentially accessible sites based on Jensen-Shannon divergence using method in previous report (*59*). We set 0.02 for specificity score threshold to determine cluster specific accessible sites. To investigate chromatin accessibility of WNT signaling family members, we calculated "gene activity scores" to evaluate the chromatin accessibility degree of genes using Cicero (*58*).

Functional enrichment analysis

GO analysis of differentially expressed genes (DEGs) was performed using ClueGO (v2.5.2) (*62*, *63*) as previously described (*64*).

Single cell pseudotime analysis

For scRNA-seq data, the pseudo-temporal analysis was performed using the R package Monocle2 (*60*, *61*, *65*).

Similarity analysis between in vivo and in vitro embryos

scRNA-seq dataset of our cultured and previous published in vivo embryos were aligned by