Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology) 2023 24(5):418-429
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Research Article
https://doi.org/10.1631/jzus.B2200440

Multiplexed single-cell transcriptome analysis reveals molecular characteristics of monkey pluripotent stem cell lines

Shuang LI[1,3], Zhenzhen CHEN[1,3], Chuanxin CHEN[1,4,5]✉, Yuyu NIU[1,2,3]✉

[1]*State Key Laboratory of Primate Biomedical Research, Institute of Primate Translational Medicine, Kunming University of Science and Technology, Kunming 650500, China*
[2]*Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming 650500, China*
[3]*Yunnan Key Laboratory of Primate Biomedical Research, Kunming 650500, China*
[4]*Bioland Laboratory, Guangzhou 510005, China*
[5]*Guangzhou Laboratory, Guangzhou 510005, China*

Abstract: Efforts have been made to establish various human pluripotent stem cell lines. However, such methods have not yet been duplicated in non-human primate cells. Here, we introduce a multiplexed single-cell sequencing technique to profile the molecular features of monkey pluripotent stem cells in published culture conditions. The results demonstrate suboptimized maintenance of pluripotency and show that the selected signaling pathways for resetting human stem cells can also be interpreted for establishing monkey cell lines. Overall, this work legitimates the translation of novel human cell line culture conditions to monkey cells and provides guidance for exploring chemical cocktails for monkey stem cell line derivation.

Key words: Monkey pluripotent stem cell; Multiplexed single-cell sequencing; Naive pluripotency; Extended pluripotent stem cells

1 Introduction

Human pluripotent stem cells (hPSCs) are currently well recognized for holding gigantic promise for regenerative medicine and studies on development (Moris et al., 2020; Girgin et al., 2021; Yu et al., 2021). However, research on hPSCs is challenged by ethical concerns at both workbench and bedside (Bredenoord et al., 2017). Recently, groundbreaking work on the non-human primate stem cell systems has made it an alternative for hPSC research and applications in the future (Ma et al., 2019; Niu et al., 2019; Chen et al., 2021).

The current method of maintaining non-human primate stem cells is largely based on the use of hPSC medium. Unfortunately, although various culture conditions have been developed to induce primed/naive/extended/formative pluripotency for hPSCs (Theunissen et al., 2014; Yang et al., 2017; Kinoshita et al., 2021), efficient and defined conditions optimized for maintaining non-human primate PSCs still have not been attained. Whether or to what degree the current protocols for hPSCs can be translated to non-human primate systems is a fundamental question for promotion of this field. Therefore, it is urgent to profile the identity and molecular signatures of non-human primate cells in human-optimized culture medium.

Single-cell transcriptome sequencing technologies have recently gained great interest in stem cell research for determining the molecular signatures of various cell types and defining cell identities (Theunissen et al., 2016; Posfai et al., 2021; Tyser et al., 2021). However, such investigations are often hampered by high costs, batch variance, and labor-consuming sampling processes when performed on multiple samples. Therefore, multiplexed sampling of single-cell transcriptome sequencing using oligo-tagged antibodies has been employed to overcome these limitations

✉ Chuanxin CHEN, redpanda00@126.com
Yuyu NIU, niuyy@lpbr.cn
Chuanxin CHEN, https://orcid.org/0000-0003-4693-6851
Yuyu NIU, https://orcid.org/0000-0002-8797-4152

Received Sept. 4, 2022; Revision accepted Jan. 15, 2023; Crosschecked Apr. 14, 2023

(Stoeckius et al., 2018). In this work, we validate for the first time the capability of a multiplexed single-cell transcriptome system to label and analyze monkey PSCs. We address the capacity of different human culture conditions to induce or maintain the pluripotency of monkey PSCs based on bio-informatic analysis, and provide guidance for improving non-human primate stem cell culture systems.

2 Results

2.1 Multiplexing various monkey pluripotent stem cell (PSC) lines based on 10× genomics

The "Cell Hashing" method developed by Stoeckuis et al. (2017) uses single-stranded DNA (ssDNA)-integrated antibodies (termed "Hashtags") that bind to ubiquitous cell-surface membrane proteins to mark and pool experimental samples. The barcoded oligos with antibodies are captured along with messenger RNAs (mRNAs). This allows the obtained data to be demultiplexed according to the barcoded antibody signals. In this work, we used the Cell Hashing method for multiplexed analysis of various monkey PSC lines.

Currently, commercial Hashtags are designed for human and mouse systems, so we first sought to verify their feasibility for cynomolgus monkey cell lines. The commercial Hashtags (Biolegend) were comprised of a pair of antibodies aimed at surface markers cluster of differentiation 298 (CD298) and β2-microglobulin (2M2). This ensured a high labeling rate of human cells. Considering the species similarities between humans and monkeys, we first sorted out the protein-sequence alignment of CD298 and 2M2 between the two species via the public database the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov). We found that CD298 was strictly conserved between human and monkey cells, but for 2M2 there was a ten-amino-acid difference. To investigate the specificity of Hashtags for labeling monkey cells, we used individual antibodies against human surface markers CD298 and 2M2 for flow cytometry analysis (Fig. S1a). Of note, although both human and cynomolgus monkey cells showed a strong human CD298-positive ($hCD298^+$) signal, cynomolgus monkey cells were not detected by human 2M2 (h2M2) antibody. Moreover, the flow results confirmed that the corresponding signals to CD298 lacked noise from mouse embryonic fibroblast (MEF)-derived feeders because MEF showed no binding capacity to either antibody (Fig. S1a). Although the single-cell sequencing system and flow cytometry system had different technical processes, the antibody-specificity data have been proved to be mutually translatable (Stoeckius et al., 2017). In conclusion, the commercial hybridized Hashtags were sufficient to specifically label monkey stem cells.

Next, we established eight monkey PSCs and captured them for the experiment (Fig. S2). The details of the eight cell lines are displayed in Table 1. pooled monkey cells were labeled with distinct Hashtags and analyzed by 10× genomics (Fig. 1a). The 10× genomics analysis results indicated that each of eight samples could be clustered separately based on Hashtags, demonstrating good demultiplexing of the samples (Fig. 1b).

For quality control, we looked at the unique molecular identifier (UMI) family read count of the

Table 1 Denotation of different monkey pluripotent stem cell lines

No.	Hashtag	Sample denotation	Cell source*	Culture condition*	Reference
1	Hashtag1	P	2-4iPS	hPSM	Ai et al., 2020
2	Hashtag2	PEP	2-4iPS	LCDM	Yang et al., 2017
3	Hashtag4	PN	2-4iPS	5iLA	Theunissen et al., 2014
4	Hashtag5	E	ABE-CES1	hPSM	Ai et al., 2020
5	Hashtag6	NE	ABE-CES1	5iLA	Theunissen et al., 2014
6	Hashtag7	NTE	NT1	hPSM	Ai et al., 2020
7	Hashtag8	EP	Dan4-1	LCDM	Yang et al., 2017
8	Hashtag10	EPT	Dan10-2	TS	Dong et al., 2020

* For detailed information, please refer to supplement materials and methods. P denotes monkey induced pluripotent stem cells (iPSCs) cultured in conventional human stem cell medium; PEP denotes monkey iPSCs cultured in extended pluripotent stem medium; PN denotes monkey iPSCs cultured in human naive medium; E denotes monkey embryonic stem cells (ESCs) derived from blastocyst in conventional human stem cell medium; NE denotes monkey ESCs derived from blastocyst in human naive medium; NTE denotes nucleus transferred monkey PSCs cultured in conventional human stem cell medium; EP denotes monkey ESCs derived from blastocyst in extended pluripotent stem medium; EPT denotes monkey cells derived from blastocyst in trophoblast stem cell medium.

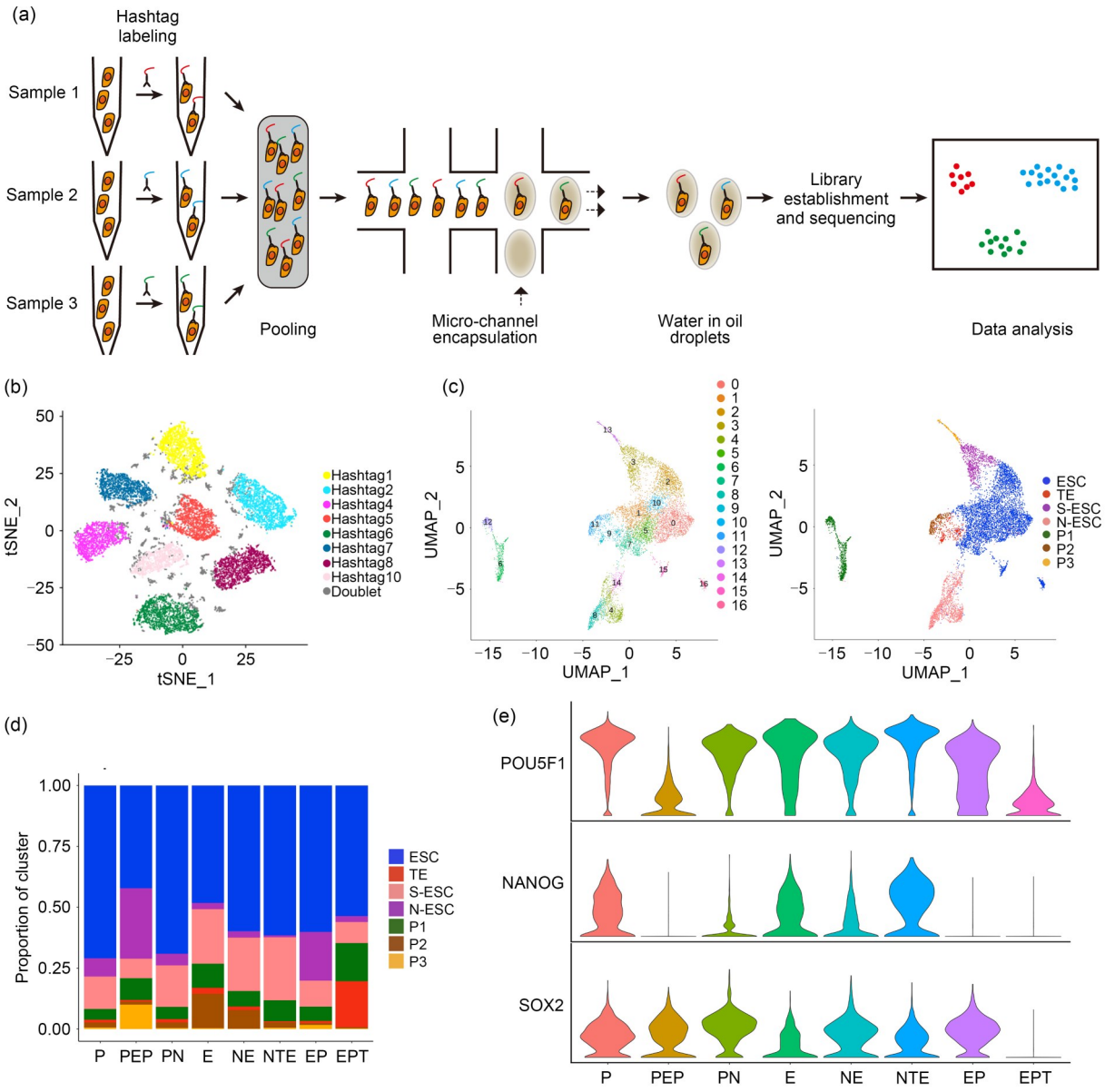

Fig. 1 Sample multiplexing of various monkey pluripotent stem cell (PSC) lines based on 10× genomics. (a) Multiplexing of monkey PSCs based on 10× genomics; (b) Demultiplexing of eight monkey PSC samples via Hashtags in t-distributed stochastic neighbor embedding (tSNE); (c) Uniform manifold approximation and projection (UMAP) analysis of unsupervised (left) and supervised (right) clustering of samples according to different cell identities; (d) Proportion of cell-identity subpopulations in each sample; (e) Expression of pluripotent markers POU5F1, NANOG, and SOX2 in each cell line. SOX2: SRY-box transcription factor 2; POU5F1: POU domain class 5 transcription factor 1, also known as octamer-binding transcription factor 4 (OCT4). The sample denotations were as listed in Table 1.

mixed pool and observed that the majority of UMI reads for each Hashtag barcode occurred in singlet droplet (stringent tested), indicating efficient single-cell encapsulation (Fig. S1b). Moreover, comparable counts were observed between double and single droplets in terms of feature genes, while significantly fewer feature genes were detected in negative droplets, suggesting that our method enabled accurate assignment of Hashtags to original samples without altering the total gene numbers among samples in the library.

2.2 Complex pluripotency of monkey PSC lines in hPSC media

Unsupervised clustering of the established 10× library gave 17 clusters (Fig. 1c left), and further convergence on embryonic development reduced the

clusters to seven groups (Figs. S1d and 1c right). We defined the group with the highest level of pluripotency-associated genes (POU domain class 5 transcription factor 1 (*POU5F1*), *NANOG*, and SRY-box transcription factor 2 (*SOX2*)) and epigenetic modification-associated genes (DNA methyltransferase 3 alpha (*DNMT3A*), *DNMT3B*, LINE-1 type transposase domain-containing 1 (*L1TD1*)) as authentic embryonic stem cells (ESCs), while we named the group with high trophoblast-associated genes (GATA-binding protein 3 (*GATA3*), Keratin 7 (*KRT7*), and Keratin 19 (*KRT19*)) as TE, due to its trophoblast-like genotype. Moderate levels of pluripotent genes and simultaneously upregulated developmental genes in some groups suggested that these subpopulations were multipotent stem cells (Fig. S1d). In particular, the group that had randomly activated expression of markers like Nestin (*NES*, ectoderm), Erb-B2 receptor tyrosine kinase 2 (*ERBB2*), and podocalyxin-like (*PODXL*, mesoderm) seemed to exhibit a progressed ESC state that was uncommitted to a specific lineage. Therefore, we termed this group "scrambled ESCs" (S-ESCs). Co-upregulation of pluripotent genes (octamer-binding transcription factor 4 (*OCT4*) and *NANOG*) and neuron development-related genes (paired box 6 (*PAX6*) and *NES*) in another group indicated neural-biased stemness, and therefore we termed it "N-ESC." Other groups were designated as P1–3 because they showed gene-expression patterns biased for specific cell fates and were not our focus in this work. Briefly, they showed deceased pluripotent genes and upregulated lineage-specific genes, which included markers for fibroblast (P1; actin α2 (*ACTA2*), collagen type I α1 chain (*COL1A1*)), endoderm (P2; *CXCR4*, Cerberus 1 (*CER1*), left-right determination factor 1 (*LEFTY1*)), and neuron (P3; tubulin β3 (*TUBB3*)) (Fig. S1d). P1–3 groups had more developed identities as compared to the ESC group.

On the whole, all the samples consisted of ESC, S-ESC, and N-ESC cell identities, while P1–3 identities only made up a minor fraction (Figs. 1c and 1d). This suggested that our samples were in complex pluripotent states even though they were cultured in state-defined conditions. It was surprising to note that even under trophoblast-directing conditions, only less than 25% of the EPT samples (monkey EPSCs derived in trophoblast stem cell (TSC) medium) exhibited a trophoblast identity, while over 50% of the population still had an ESC identity, implying inefficient differentiation. One reason could be that the medium was not optimized for monkey cells and cell-identity change toward trophoblast identity was incomplete.

P, which was maintained in conventional human PSC medium, showed a minimal percentage of developed identity (P1, P2, and P3) and mainly consisted of ESC, S-ESC, and N-ESC (Fig. 1d). The small proportion of P1–3 demonstrated the capacity of hPSC medium (hPSM) to suppress differentiation in monkey stem cells, yet the hybrid pluripotency indicated that knockout serum replacement (KSR) and fibroblast growth factor-2 (FGF-2) were not sufficient to maintain a pure primed state. Interestingly, although PEP (monkey induced PSC (iPSCs) cultured in extended pluripotent stem medium) and PN (monkey iPSCs cultured in human naive medium) were both derived from P with the aim of reprogramming cell identity from primed pluripotency towards extended or naive pluripotency, PN presented similar cell-identity composition to that of P, while PEP had a reduced ESC component (less than 50%). This indicated that 5iLA conditions (for naive pluripotency) exclusively impacted the ESC component of P, while LCDM conditions (for extended pluripotency) affected lineage-specific genes and caused differentiation of ESC in P.

The non-optimization of these culture conditions for non-human primate cells was also evident in other samples (Figs. 1d and 1e). Greater heterogeneity was obtained in lines derived from embryos using the same culture conditions for iPSC-derived lines. Compared with P and PN, smaller proportions of the ESC group were found in E (monkey ESCs derived from blastocyst in conventional human stem cell medium) or NTE (nucleus transferred monkey PSCs cultured in conventional human stem cell medium) primed lines and in the blastocyst-derived naive line (NE, monkey ESCs derived from blastocyst in human naive medium). Interestingly, although E and NTE were both derived under the same culture conditions, NTE had more ESC and less P2, suggesting that the origin of materials had some influence on the quality of derived cell lines. Moreover, P1–3 cells were apparently less abundant in NE compared to the proportion in E, suggesting that the culture system designed to induce the naive network in hPSCs also improved pluripotency in non-human primate PSCs. Whether the increased proportion of P2 in NE compared to NTE was a result

of the glycogen synthase kinase 3 (GSK3) signaling manipulation in 5iLA medium (Jiang et al., 2021) required further validation. In contrast, when compared with PEP, EP (monkey ESCs derived from blastocyst in extended pluripotent stem medium) showed a higher proportion of ESC, indicating that LCDM conditions were better optimized for deriving ESCs from blastocysts than from iPSCs. This was also evidenced by the lower proportion of P1–3 in EP compared with that in PEP. Notably, NANOG was the most strongly influenced pluripotent marker among the eight cell lines (Fig. 1e), consistent with the fact that its requirement could be bypassed by modulating other pathways for maintaining or inducing pluripotency (Takahashi et al., 2007; Stuart et al., 2014).

Additionally, we focused on the ESC commitment of all eight cell lines. Spearman correlation analysis of ESC components between all samples illustrated that all of them were highly correlated (>0.90) (Fig. S1c). The ESC component from EPT and PEP exhibited a lower correlation to primed ESC identity. Although the difference was minor and the details were not clarified, the data suggested that ESCs in TSC medium and EPS conditions displayed a certain degree of difference from the primed state.

In sum, culture conditions that were optimized for maintaining or inducing human pluripotency were shown to be inefficient for non-human primate PSCs, resulting in a complex hybrid of pluripotent cells. Of the culture conditions tested, P had the highest proportion of cells with high pluripotent gene levels.

2.3 Molecular features of monkey PSC under different culture conditions

In order to elucidate the impact of environmental factors on cell identity, we first looked into the detailed molecular features of samples that were derived from the same origin (P, PEP, and PN). In the Uniform Manifold Approximation and Projection (UMAP) plot, P and PN were both enriched for the ESC compartment, indicating good maintenance of pluripotency (Figs. 1c and 2a left & middle). We noticed colocalization of P and PN dots, which implied that cells in these two conditions shared some similarities. However, PEP, which was derived and maintained in EPS culture conditions, showed a distinct pattern (Figs. 1c and 2a right).

Since the culture conditions in these samples were designed to induce distinct pluripotent states, we specifically analyzed the ESC components of P, PEP, and PN to depict their molecular features. Fig. 2b shows the differentially expressed gene (DEG) patterns of ESC components in P, PEP, and PN. Intriguingly, although PEP had the greatest proportion of N-ESC identity in comparison to the other two samples, PN_ESCs were found to have the highest level of *SOX2*. Unexpectedly, TSC22 domain family member 1 (TSC22D1) and microtubule-associated protein 1B (MAP1B) were more actively expressed in PEP_ESCs and PN_ESCs than in P_ESCs. MAP1B plays a role in cytoskeleton structure changes, which is associated with neuronal migration and development. TSC22D1 is a transforming growth factor-β (TGFβ) downstream target. While these two genes are closely related to differentiation and development, their expression was found to be activated in response to naive media. This phenomenon indicated that different from the mouse system, in which TGFβ signaling was required for primed pluripotency, a fine tuning of TGFβ signaling was required for inducing naive or extended pluripotency in primates (Tesar et al., 2007; Pastor et al., 2016; Ai et al., 2020; Bayerl et al., 2021).

Gene Ontology (GO) analysis of the DEGs from Fig. 2b manifested higher enrichment of genes in association with metabolic processes and response to hypoxia in P_ESCs versus PEP_ESCs and PN_ESCs (Fig. 2c). In addition, PEP_ESCs showed enrichment of mitosis-related processes compared to P_ESCs, while negative regulation of RNAs and positive regulation of cell death were found to be more active in P_ESCs versus PN_ESCs, suggesting that P_ESCs had the lowest self-renewal capability. Interestingly, terms related to mitosis and chromosome organization were also enriched in PEP_ESCs versus PN_ESCs, suggesting the nucleus activities were more robust in PEP_ESCs. In contrast, GO terms related to differentiation and development were enriched in PN_ESCs versus PEP_ESCs, indicating that LCDM conditions were more suppressive for the progression of cell identity. Whether this was a sign of naive network establishment remained to be discovered. Variances were not enriched in these terms between PN_ESCs and P_ESCs, suggesting that the two groups shared similar pluripotent states.

In all, our data demonstrated that LCDM and 5iLA conditions virtually reset cells from the same origin to distinct pluripotent states. ESCs in primed condition had a higher metabolic level and were more

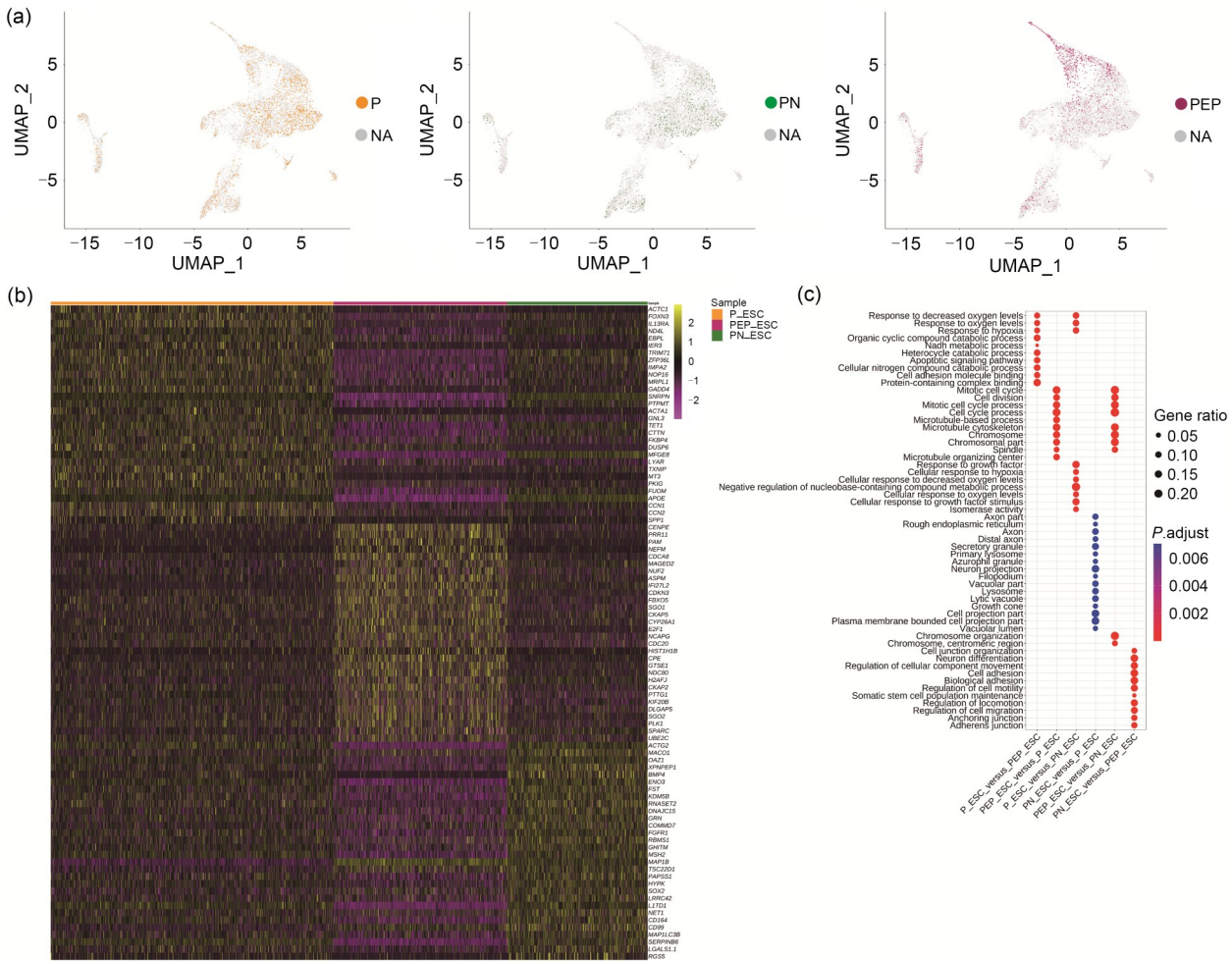

Fig. 2 Molecular features of monkey pluripotent stem cells in different culture environments. (a) Highlighted region of P/PN/PEP in the UMAP dataset; (b) Differentially expressed genes in the ESC component of P/PN/PEP; (c) Dot plot of the Gene Ontology (GO) analysis of the ESC component of P/PN/PEP. UMAP: Uniform Manifold Approximation and Projection; ESC: embryonic stem cell; NA: not available. The sample denotations were as listed in Table 1.

biased to development, and LCDM and 5iLA conditions were sufficient to promote self-renewal and reprogram monkey stem cells to less-developed cell identities.

2.4 Molecular features of different monkey PSC lines derived under the same environment

Next, we investigated how different contexts from non-human primates would respond to the same human culture conditions. Considering that hPSM yielded the highest ratio of the ESC group (Fig. 1d), we analyzed the single-cell transcriptomes of different lines maintained under these environments. As expected, cells directly derived from the embryos (E) showed different patterns in the UMAP compared to iPSC line (P) (Figs. 1c and 3a). Transcriptome analysis of the ESC components of P and E displayed distinct signatures (Fig. 3b). When we included another type of embryo-derived cell line (NTE) in hPSM, NTE_ESCs and E_ESCs were close to each other in the UMAP plot (Fig. 3a) and had similar gene signatures (Fig. 3b). This phenomenon indicated that original contexts had a profound impact on the transcriptomic features of the cell lines.

GO analysis revealed enrichment of genes associated with energy and nucleic-acid metabolism in P_ESCs (Fig. 3c). Additionally, E exhibited enrichment of activated genes in the terms in relation with cell-substrate junction when compared to P_ESCs. As expected, the same result was found when we analyzed NTE_ESCs against P_ESCs, suggesting that it was the result of the different materials rather than the culture

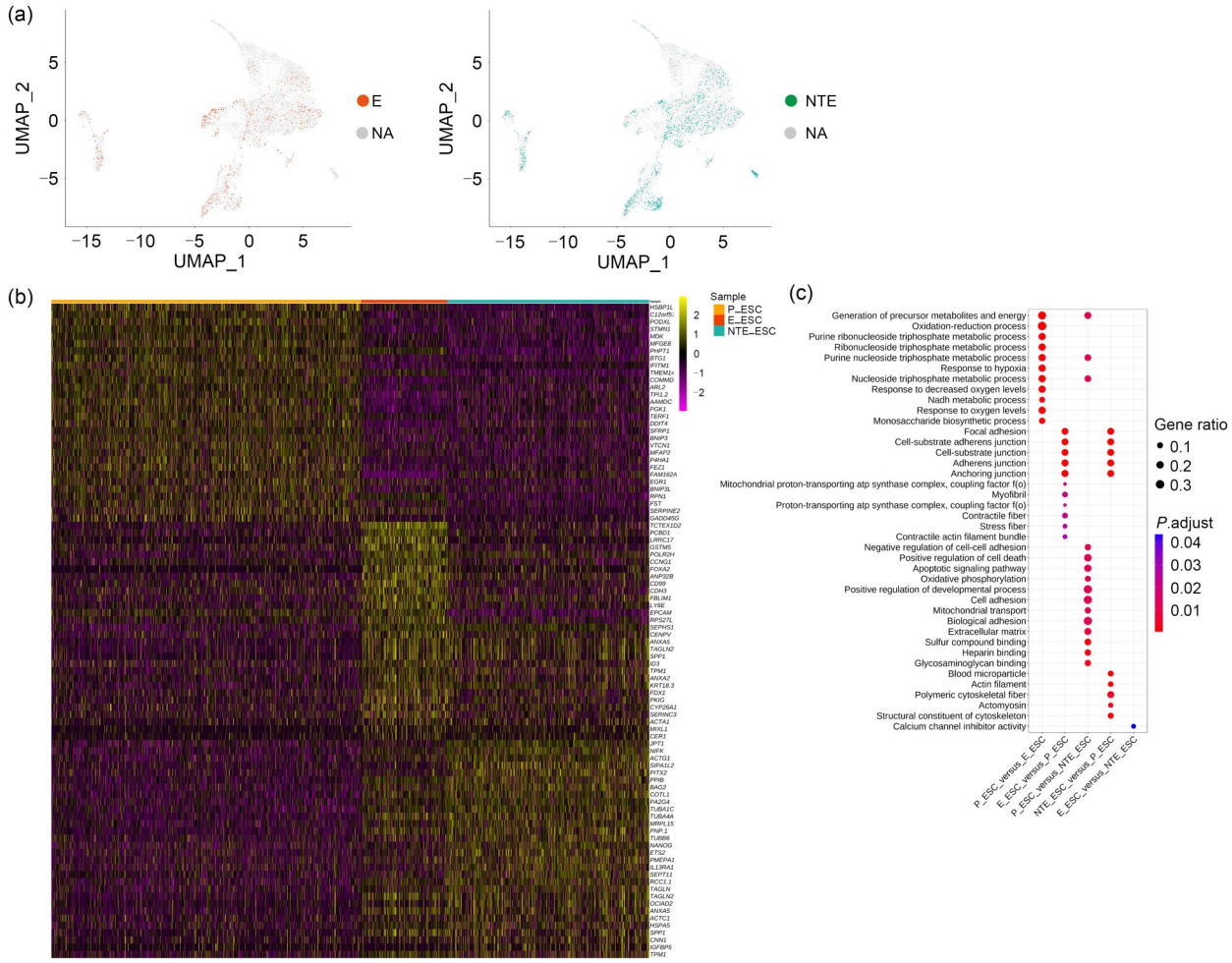

Fig. 3 Molecular features of different monkey pluripotent stem cells derived under the same culture conditions. (a) Highlighted region of E/NTE in the UMAP dataset; (b) Differentially expressed genes in the ESC component of P/E/NTE; (c) Dot plot of the gene ontology analysis of the ESC component of P/E/NTE. UMAP: Uniform Manifold Approximation and Projection; ESC: embryonic stem cell; NA: not available. The sample denotations were as listed in Table 1.

conditions. This indicated that ESCs derived from embryos robustly adapted to the physical environment in vitro. The fact that the selected GO terms barely showed any difference between E_ESCs and NTE_ESCs validated our former conclusion that the identified DEGs between different cell lines might reflect context variance rather than culture environments (Fig. 3c).

In sum, our data demonstrated that the origin of the cells could significantly affect the molecular features of the derived lines. Conversion or maintenance of cell identity under certain environments was dependent on cell context. This points to the need for cautious filtering to remove the noise from material variance when evaluating culture conditions.

2.5 Mapping monkey PSC cell lines to their in vivo counterparts during embryogenesis

In order to investigate the in vivo counterparts of the non-human primate PSCs, we collected single-cell RNA-sequencing (RNA-seq) data from monkey embryos representing different developmental stages (Liu et al., 2018) (Figs. 4b and 4c). Visualization of the data showed segregation of distinct clusters representing early blastocyst (EB), late blastocyst (LB), and hatching blastocyst (HB) stages in both two and three dimensions. Reducing the level of dimension did not obscure the border of these clusters (Fig. 4c). In particular, medium blastocyst (MB) colocalized with EB, suggesting a close resemblance of the two types of

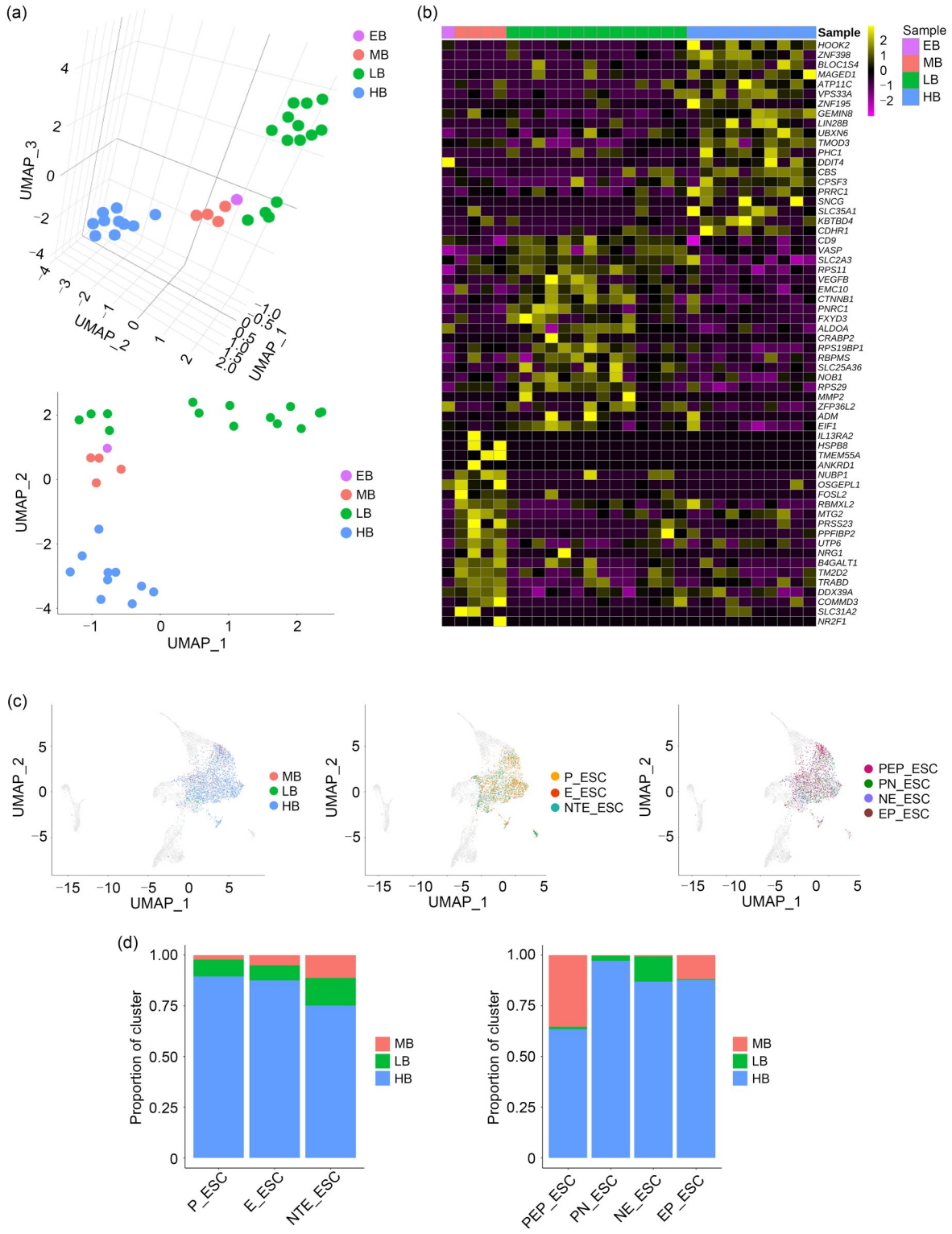

Fig. 4 Mapping monkey pluripotent stem cell lines to in vivo counterparts during embryogenesis. (a) 3D (top) and 2D (bottom) UMAP analyses of in vivo embryo data; (b) Heatmap of differentially expressed genes in EB, MB, LB, and HB; (c) Monkey pluripotent stem cells were mapped to in vivo features (left) and compared to the ESC identities of each sample (middle and right); (d) Proportion of cells with embryo features in ESCs of each sample. UMAP: Uniform Manifold Approximation and Projection; ESC: embryonic stem cell; EB: early blastocyst; MB: medium blastocyst; LB: late blastocyst; HB: hatching blastocyst. The sample denotations were as listed in Table 1.

blastocysts. This was in accordance with their developmental status. Interestingly, LB had an elongated distribution in projection to UMAP_1, representative of the occurrence of progressive development in this stage. Also, a small part of LB was close to EB and MB, indicating that a minor fraction of LB still possessed features similar to MB. The segregation of HB from all the other clusters suggested significant identity variance between HB and other embryo stages.

Mapping the transcriptome features of in vivo embryos to our data showed that a majority of cells with MB, LB, and HB features fell into the ESC identity region (Fig. 4a). Further mapping the ESC clusters to the embryo data illustrated that most of our obtained ESCs had a molecular profile similar to that of HB (Fig. 4d). Comparison between samples cultured under the same conditions showed similar proportions of MB, LB, and HB for P_ESC, E_ESC, and NTE_ESC. Surprisingly, PEP and EP had significantly higher proportions of MB than the other groups, suggesting that EPS conditions reset cells to earlier developmental stages. However, the different ratios of MB in PEP_ESC and EP_ESC also supported our earlier conclusion that conversion or maintenance of cell identity under the same conditions could be altered using different starting materials.

In conclusion, our data demonstrated that non-human primate cells in culture conditions currently used to maintain human stem cell pluripotency closely resembled HBs, indicating their primed-state pluripotency. Of note, EPS conditions tended to induce earlier developmental stages of cell identity than all the other conditions tested.

3 Discussion

Here, we developed a multiplexed single-cell RNA-seq platform and successfully performed analysis of eight monkey cell lines that were cultured in conditions optimized for hPSCs. By utilizing DNA barcodes coupled with antibodies (Hashtag) to label cells, we were able to distinguish samples from a single organism without disturbing cell status. Hashtags containing a hybrid of two antibodies were applied to label and demultiplex cynomolgus monkey cells. A flow chart (Fig. S1a) demonstrated specificity to monkey cells, ensuring the results that reflected cynomolgus monkey cells in defined culture conditions.

Various culture conditions have been used in stem cell research to maintain or induce different states of pluripotency for hPSCs. Since non-human primate models have recently become a promising platform for clinical translation, optimization of culture conditions for monkey cells is desirable rather than simply using hPSC-based media. To this end, we profiled the molecular features of eight monkey PSC lines cultured under hPSC conditions. Annotation of the sample clusters demonstrated heterogeneity of outcomes (Fig. 1d). While hPSC-optimized medium was shown to direct cell identity to some extent, its efficiency is suboptimal for monkey PSCs. hPSM is the most efficient medium, as it generates the least upregulation of developmental genes. Although 5iLA culture conditions have been proven to reset human primed PSCs to a naive-like state (Theunissen et al., 2014, 2016), in our study these conditions did not seem to cause any significant variation in the ESC ratio compared to conventional hPSM (Fig. 1d). However, a deeper look into the ESC transcriptome profiles of P and PN reveals distinct molecular signatures (Fig. 2b). GO analysis confirms the capacity of 5iLA conditions to suppress differentiation of associated genes (Fig. 2c); however, mapping the molecular profile to in vivo data reveals a resemblance of these ESCs to HB, which is a sign of primed pluripotency (Fig. 4d). At the least, we found evidence that 5iLA conditions select the exact signaling pathways to suppress differentiation and that they promote expansion of the proportion of LB and MB in monkey PSCs. Therefore, future work may require manipulation of inhibitor concentration in 5iLA media to improve monkey naive-cell culture. Another possible reason that the naive state was not established is the presence of FGF signaling (Fig. 2b), since it has been reported that adaptation to FGF-2-free conditions enhances upregulation of naive-related genes during reprogramming (Chen et al., 2015).

The PEP line used LCDM conditions as an EPS-directing medium. Given the fact that ESCs in LCDM or 5iLA environments show little variance in GO terms associated with differentiation and development (Fig. 2c), the two conditions are proven to reset the primed ESCs to more primitive and undeveloped pluripotent states. Simultaneously, ESCs in LCDM conditions were endowed with better self-renewal ability, even though FGF-2 was not included in the medium. Interestingly, overall analysis indicated that PEP had the highest rate of N-ESCs (Fig. 1d) and the lowest

level of NANOG (Fig. 1e). However, mapping the ESC from PEP to the in vivo counterpart revealed the highest frequency of early embryos (MB) identity, which suggests improved pluripotency (Fig. 4d). Therefore, we conclude that LCDM conditions are capable of bypassing NANOG and FGF signaling pathways to enhance self-renewal of primate ESCs while rewiring the molecular signatures towards an earlier pluripotent state.

Analysis of the molecular profiles of cells cultured in the same environments illustrates that the material origins have a profound impact on single-cell techniques. Indeed, the differences in DEGs between reprogrammed cells (P_ESCs) and embryo-derived cells (E_ESCs and NTE_ESCs) reflect variations in material rather than cell identity. Notably, the data in Fig. 3b show that the ESCs in E in general had higher levels of forkhead box protein A2 (*FOXA2*), mix paired-like homeobox 1 (*MIXL1*), and *CER1*. This suggests a primed state biased to endoderm lineages, as these genes are in close association with mesendoderm (Kempf et al., 2016; Su et al., 2020). This phenomenon might account for the higher P2 ratio in E, as the E_ESCs are prone to auto-differentiating into endoderm cells due to robust expression of *FOXA2*.

GO analysis confirmed the presence of enhanced energy and nucleic acid metabolism in the iPSCs (Fig. 3c). It is worth noting that both lines derived from embryos (E and NTE) showed enrichment of GO terms in association with cell–substrate junction, indicating that these ESCs retained some degree of polarity and had robust interactions with extracellular matrix. This polarity was lost in iPSCs as a result of reprogramming. Interestingly, NTE_ESCs showed elevated activity in GO terms for mesodermal differentiation, perhaps due to the imprinted pattern inherited from the fibroblasts from which the NTE nucleus was extracted. However, this elevated activity of mesoderm-biased gene expression was not observed in the iPSC line P, which was reprogrammed from fibroblasts, demonstrating efficient removal of the imprints from the parents by the reprogramming factors (Takashima et al., 2015; Pastor et al., 2016). Our results therefore demonstrate the direct effects of nuclear reprogramming in erasing original molecular features; neither nuclear transfer nor switching culture conditions was effective in completely disentangling parental genome patterns. Nevertheless, in vivo mapping of the transcriptome profile indicates similar developmental status for P and E, while NTE shows a slightly increased fraction of population with similarity to earlier developmental status (Fig. 4d). The contrary conclusions drawn from GO analysis and in vivo mapping might be due to misunderstanding of the functionalities of mesoderm-related genes. Stuart et al. (2019) reported that during nuclear reprogramming in the murine system, stem cells could be cultivated into a mesoderm-like state and then returned to naive pluripotency. This routine showed the activation of multiple mesoderm-related genes that were likely to be categorized into GO terms associated with mesodermal differentiation. We do not know whether the same scenario was taking place in the NTE line, and further work would be required to unravel the role of these marker genes in lineage specification.

To conclude, this work demonstrates that non-human primate PSCs exhibit a complicated cell-identity mixture in hPSC conditions due to species variation. On the whole, primed pluripotency was obtained under most culture conditions, yet the selected signaling pathways in hPSC conditions were found to only partially direct monkey stem cells to the designated identities. This work provides an in silico reference for establishing optimized conditions for non-human primate PSC culture.

Materials and methods

Detailed methods are provided in the electronic supplementary materials of this paper.

Acknowledgments

This work was supported by the National Key R&D Program of China (Nos. 2021YFA0805700 and 2021YFA1102000), the National Natural Science Foundation of China (No. U2102204), and the Natural Science Foundation of Yunnan Province, China (Nos. 202001BC070001 and 202102AA100053).

Author contributions

Yuyu NIU and Chuanxin CHEN conceived the structure of manuscript. Yuyu NIU and Shuang LI designed the experiments. Shuang LI performed the experiments and analyzed data. Zhenzhen CHEN prepared the cell lines. Chuanxin CHEN and Shuang LI wrote the context and generated the figures. Yuyu NIU, Chuanxin CHEN, and Shuang LI contributed to revision of the manuscript. Bioinformatics analysis was supported by SequMed BioTechnology Inc. (Guangzhou, China). All authors have read and approved the final manuscript, and therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

Compliance with ethics guidelines

Shuang LI, Zhenzhen CHEN, Chuanxin CHEN, and Yuyu NIU declare that they have no conflict of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Ai ZY, Niu BH, Duan K, et al., 2020. Modulation of Wnt and Activin/Nodal supports efficient derivation, cloning and suspension expansion of human pluripotent stem cells. *Biomaterials*, 249:120015. https://doi.org/10.1016/j.biomaterials.2020.120015

Bayerl J, Ayyash M, Shani T, et al., 2021. Principles of signaling pathway modulation for enhancing human naive pluripotency induction. *Cell Stem Cell*, 28(9):1549-1565.E12. https://doi.org/10.1016/j.stem.2021.04.001

Bredenoord AL, Clevers H, Knoblich JA, 2017. Human tissues in a dish: the research and ethical implications of organoid technology. *Science*, 355(6322):eaaf9414. https://doi.org/10.1126/science.aaf9414

Chen CX, Ji WZ, Niu YY, 2021. Primate organoids and gene-editing technologies toward next-generation biomedical research. *Trends Biotechnol*, 39(12):1332-1342. https://doi.org/10.1016/j.tibtech.2021.03.010

Chen HW, Aksoy I, Gonnot F, et al., 2015. Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat Commun*, 6:7095. https://doi.org/10.1038/ncomms8095

Dong C, Beltcheva M, Gontarz P, et al., 2020. Derivation of trophoblast stem cells from naïve human pluripotent stem cells. *eLife*, 9:e52504. https://doi.org/10.7554/eLife.52504

Girgin MU, Broguiere N, Hoehnel S, et al., 2021. Bioengineered embryoids mimic post-implantation development in vitro. *Nat Commun*, 12:5140. https://doi.org/10.1038/s41467-021-25237-8

Jiang YQ, Chen CX, Randolph LN, et al., 2021. Generation of pancreatic progenitors from human pluripotent stem cells by small molecules. *Stem Cell Rep*, 16(9):2395-2409. https://doi.org/10.1016/j.stemcr.2021.07.021

Kempf H, Olmer R, Haase A, et al., 2016. Bulk cell density and Wnt/TGFbeta signalling regulate mesendodermal patterning of human pluripotent stem cells. *Nat Commun*, 7:13602. https://doi.org/10.1038/ncomms13602

Kinoshita M, Barber M, Mansfield W, et al., 2021. Capture of mouse and human stem cells with features of formative pluripotency. *Cell Stem Cell*, 28(3):453-471.e8. https://doi.org/10.1016/j.stem.2020.11.005

Liu DH, Wang XY, He DJ, et al., 2018. Single-cell RNA-sequencing reveals the existence of naive and primed pluripotency in pre-implantation rhesus monkey embryos. *Genome Res*, 28(10):1481-1493. https://doi.org/10.1101/gr.233437.117

Ma HX, Zhai JL, Wan HF, et al., 2019. In vitro culture of cynomolgus monkey embryos beyond early gastrulation. *Science*, 366(6467):eaax7890. https://doi.org/10.1126/science.aax7890

Moris N, Anlas K, van den Brink SC, et al., 2020. An in vitro model of early anteroposterior organization during human development. *Nature*, 582(7812):410-415. https://doi.org/10.1038/s41586-020-2383-9

Niu YY, Sun NQ, Li C, et al., 2019. Dissecting primate early post-implantation development using long-term in vitro embryo culture. *Science*, 366(6467):eaaw5754. https://doi.org/10.1126/science.aaw5754

Pastor WA, Chen D, Liu WL, et al., 2016. Naive human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory. *Cell Stem Cell*, 18(3):323-329. https://doi.org/10.1016/j.stem.2016.01.019

Posfai E, Schell JP, Janiszewski A, et al., 2021. Evaluating totipotency using criteria of increasing stringency. *Nat Cell Biol*, 23(1):49-60. https://doi.org/10.1038/s41556-020-00609-2

Stoeckius M, Hafemeister C, Stephenson W, et al., 2017. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*, 14(9):865-868. https://doi.org/10.1038/nmeth.4380

Stoeckius M, Zheng SW, Houck-Loomis B, et al., 2018. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol*, 19:224. https://doi.org/10.1186/s13059-018-1603-1

Stuart HT, van Oosten AL, Radzisheuskaya A, et al., 2014. NANOG amplifies STAT3 activation and they synergistically induce the naive pluripotent program. *Curr Biol*, 24(3):340-346. https://doi.org/10.1016/j.cub.2013.12.040

Stuart HT, Stirparo GG, Lohoff T, et al., 2019. Distinct molecular trajectories converge to induce naive pluripotency. *Cell Stem Cell*, 25(3):388-406.e8. https://doi.org/10.1016/j.stem.2019.07.009

Su J, Morgani SM, David CJ, et al., 2020. TGF-β orchestrates fibrogenic and developmental EMTs via the RAS effector RREB1. *Nature*, 577(7791):566-571. https://doi.org/10.1038/s41586-019-1897-5

Takahashi K, Tanabe K, Ohnuki M, et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5):861-872. https://doi.org/10.1016/J.CELL.2007.11.019

Takashima Y, Guo G, Loos R, et al., 2015. Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell*, 162(2):452-453. https://doi.org/10.1016/j.cell.2015.06.052

Tesar PJ, Chenoweth JG, Brook FA, et al., 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*, 448(7150):196-199. https://doi.org/10.1038/nature05972

Theunissen TW, Powell BE, Wang HY, et al., 2014. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell*, 15(4):471-487. https://doi.org/10.1016/j.stem.2014.07.002

Theunissen TW, Friedli M, He YP, et al., 2016. Molecular

criteria for defining the naive human pluripotent state. *Cell Stem Cell*, 19(4):502-515. https://doi.org/10.1016/j.stem.2016.06.011

Tyser RCV, Mahammadov E, Nakanoh S, et al., 2021. Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature*, 600(7888):285-289. https://doi.org/10.1038/s41586-021-04158-y

Yang Y, Liu B, Xu J, et al., 2017. Derivation of pluripotent stem cells with in vivo embryonic and extraembryonic potency. *Cell*, 169(2):243-257.e25. https://doi.org/10.1016/j.cell.2017.02.005

Yu LQ, Wei YL, Duan JL, et al., 2021. Blastocyst-like structures generated from human pluripotent stem cells. *Nature*, 591(7851):620-626. https://doi.org/10.1038/s41586-021-03356-y

Supplementary information

Materials and methods; Figs. S1 and S2