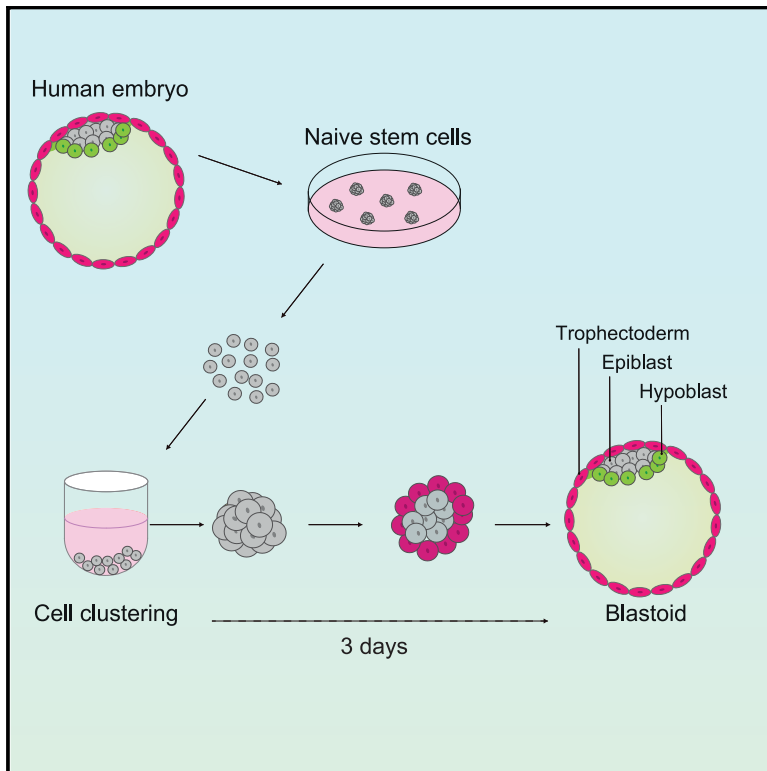


# Cell Stem Cell

## Naive stem cell blastocyst model captures human embryo lineage segregation

### Graphical abstract



### Authors

Ayaka Yanagida, Daniel Spindlow,  
Jennifer Nichols, Anish Dattani,  
Austin Smith, Ge Guo

### Correspondence

austin.smith@exeter.ac.uk (A.S.),  
g.guo@exeter.ac.uk (G.G.)

### In brief

Reliable stem cell models of the early embryo would open new approaches to studying human development and infertility. Yanagida and colleagues used naive stem cells to generate human blastocyst-like structures comprising the three founding tissue layers for the extraembryonic membranes and the embryo.

### Highlights

- Generation of human blastoids with structure and composition similar to the embryo
- Segregation of trophectoderm, hypoblast, and naive epiblast lineages
- Transcriptome fidelity with the embryo
- Simple, efficient, and scalable procedure



Brief Report

# Naive stem cell blastocyst model captures human embryo lineage segregation

Ayaka Yanagida,<sup>1,4</sup> Daniel Spindlow,<sup>1</sup> Jennifer Nichols,<sup>2,3</sup> Anish Dattani,<sup>1</sup> Austin Smith,<sup>1,\*</sup> and Ge Guo<sup>1,4,5,\*</sup>

<sup>1</sup>Living Systems Institute, University of Exeter, Stocker Road, Exeter EX4 4QD, UK

<sup>2</sup>Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Puddicombe Way, Cambridge CB2 0AW, UK

<sup>3</sup>Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 1GA, UK

<sup>4</sup>Present address: Division of Stem Cell Therapy, Centre for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan

<sup>5</sup>Lead contact

\*Correspondence: [austin.smith@exeter.ac.uk](mailto:austin.smith@exeter.ac.uk) (A.S.), [g.guo@exeter.ac.uk](mailto:g.guo@exeter.ac.uk) (G.G.)

<https://doi.org/10.1016/j.stem.2021.04.031>

## SUMMARY

Human naive pluripotent cells can differentiate into extraembryonic trophoderm and hypoblast. Here we describe a human embryo model (blastoid) generated by self-organization. Brief induction of trophoderm leads to formation of blastocyst-like structures within 3 days. Blastoids are composed of three tissue layers displaying exclusive lineage markers, mimicking the natural blastocyst. Single-cell transcriptome analyses confirm segregation of trophoderm, hypoblast, and epiblast with high fidelity to the human embryo. This versatile and scalable system provides a robust experimental model for human embryo research.

## INTRODUCTION

Natural development of the human embryo is challenging to study *in vivo* and few embryos are available for research *in vitro*. Scientists have therefore relied heavily on observations and experiments in other mammals, in particular mice. However, although embryology unfolds according to a similar overall program in all mammals, there are many distinctions between species. Even the first morphogenetic process—formation of the blastocyst—is regulated differently in human and mouse (Guo et al., 2021; Roode et al., 2012; Rossant, 2018). The blastocyst is a landmark of eutherian development that is essential for uterine implantation. Blastocyst formation initiates with delamination of epithelial trophoderm cells on the surface of the unspecified morula. The trophoderm forms a fluid-filled cavity and the internal inner cell mass (ICM) cells differentiate into two further lineages, epiblast and hypoblast (also known as primitive endoderm). The mature blastocyst formed by embryonic day 6 (E6) in human is a simple cavitated structure comprising three topologically and molecularly segregated lineages, each of which is critical for further development.

In recent provocative research, blastocyst-like structures termed blastoids have been generated from mouse stem cells (Rivron et al., 2018; Sozen et al., 2019). Approaches have subsequently been devised to manufacture representations of the human blastocyst (Liu et al., 2021; Yu et al., 2021). To be a useful model, however, blastoids must accurately recapitulate the cellular organization and lineage composition of the natural human blastocyst (Rossant and Tam, 2021).

Here we investigated the potential of human naive stem cells (Brendenkamp et al., 2019b; Guo et al., 2016; Takashima

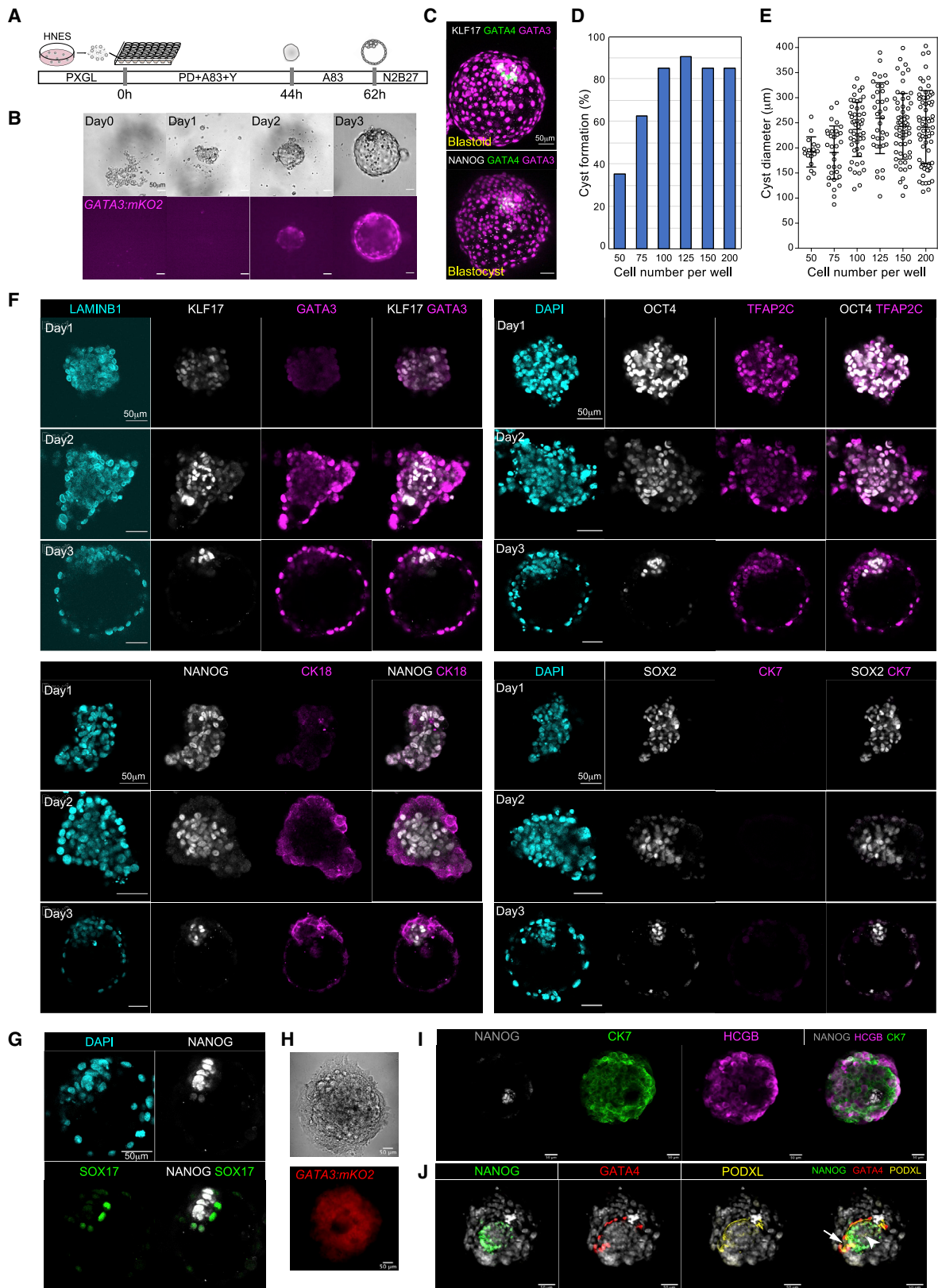
et al., 2014) to generate blastoids that show high fidelity to the human embryo and fulfil key criteria for an experimental model system.

## RESULTS

Recently we found that human naive pluripotent stem cells can generate mixed cultures comprising the three founding lineages of the blastocyst (Guo et al., 2021). Trophoderm differentiation is initiated by inhibition of ERK and NODAL signaling using small molecules PD0325901 and A83-01 (PD+A83). We also observed that treatment with PD+A83 in suspension resulted in appearance of epithelial cysts expressing the trophoderm reporter *GATA3:mKO2*. We noticed that cysts often contained inner cells that were initially mKO2 negative but became positive upon continued exposure to PD+A83. These findings suggested the possibility of generating blastocyst-like structures from naive stem cells by modulating PD+A83 treatment.

We set up trials in non-adherent 96-well U-bottom plates using *GATA3:mKO2* to monitor the formation and distribution of trophoderm cells. Cells were dissociated and seeded in PD+A83 plus Rho-associated kinase inhibitor Y-27632 to aid survival. The day after seeding, loose aggregates formed. *GATA3:mKO2* expression was apparent in surface cells on day 2 and cavitation initiated. At the end of day 2, we exchanged the medium to culture in A83 only. On day 3 many of the aggregates formed expanded cysts of *GATA3:mKO2*-positive cells with a *GATA3*-negative inner population (Figures 1A and 1B). We detected non-overlapping expression of naive epiblast (*KLF17*) and hypoblast (*GATA4*) markers in the inner population, as in the mature human blastocyst (Figure 1C).





(legend on next page)

The initial cell seeding number influenced the efficiency of cyst formation. Between 100 and 150 cells, we obtained single cavi- tated structures in more than 80% of wells (Figure 1D). Seeding fewer cells gave mainly compact aggregates of GATA3-positive cells (Figures S1A–S1C) and at higher seeding density multiple cysts formed per well. The size of cysts varied, with an average diameter of about 250  $\mu\text{m}$ , similar to the late human blastocyst (Figures 1E and S1A–S1C). Blastocyst-like structures main- tained integrity on day 4 after removal of A83.

By immunofluorescence staining we characterized expression of trophectoderm and epiblast markers in day 3 cysts (Figure 1F). GATA3 protein was readily detected from day 2 and localized in the outer layer together with the epithelial protein keratin 18 (CK18). On day 3, GATA3 was prominent in nuclei of all outer cells and absent from inside cells. In contrast, epiblast markers KLF17, NANOG, OCT4, and SOX2 that are ubiquitous in naive pluripotent cells (Bredenkamp et al., 2019b) were confined to inner cells by day 3. Expression of trophectoderm and epiblast markers was generally exclusive, as observed in the human mature blastocyst. Notably the core pluripotency factor OCT4 was restricted to inner cells that were all GATA3 negative. OCT4 can be detected in trophectoderm in the early human blas- tocyst (E5), but it is confined to the ICM by the fully expanded E6 blastocyst (Deglincerti et al., 2016; Niakan and Eggan, 2013; Shahbazi et al., 2016). TFAP2C (AP2 $\gamma$ ) is known as a trophoblast marker in mouse but in human is also expressed in naive epiblast and naive stem cells (Pastor et al., 2018). We detected moderate expression of TFAP2C in inner cells with upregulation in the outer epithelium. Presence of KLF17 in inner cells indicated that they remain in the pre-implantation stage of naive pluripotency (Bla- keley et al., 2015; Boroviak et al., 2015; Rostovskaya et al., 2019). Finally, barely detectable keratin 7 (CK7) is consistent with pre-implantation trophectoderm in contrast to post-implan- tation cytotrophoblast (Deglincerti et al., 2016).

By E6 the human ICM has segregated into naive epiblast and hypoblast (Niakan and Eggan, 2013; Petropoulos et al., 2016; Roode et al., 2012; Stirparo et al., 2018). We assessed pres- ence of hypoblast in day 3 cysts. We detected GATA4-positive cells in 36 of 44 cysts (Figures S1D and S1E). The number of GATA4-positive cells varied (Figure S1F), as observed for hypo- blast cells in human embryos (Niakan and Eggan, 2013; Roode et al., 2012). We also detected expression of hypoblast markers SOX17 and OTX2 (Boroviak et al., 2018) in cells that typically appeared to underly the SOX2-positive epiblast (Figures 1G and S1H).

We examined cyst formation from other embryo-derived (HNES2, HNES3), chemically reset (cR\_Shef6) (Guo et al., 2017), and directly reprogrammed (niPSC\_HDF75) (Bredenkamp et al., 2019b) naive stem cells. We obtained expanded blasto- cyst-like structures within 3 days in all cases although efficiency ranged from around 30% for cR-Shef6 to greater than 80% for niPSC\_HDF75 (Figures S1G). By day 4 all three lineages were consistently present (Figures S1I and S1J).

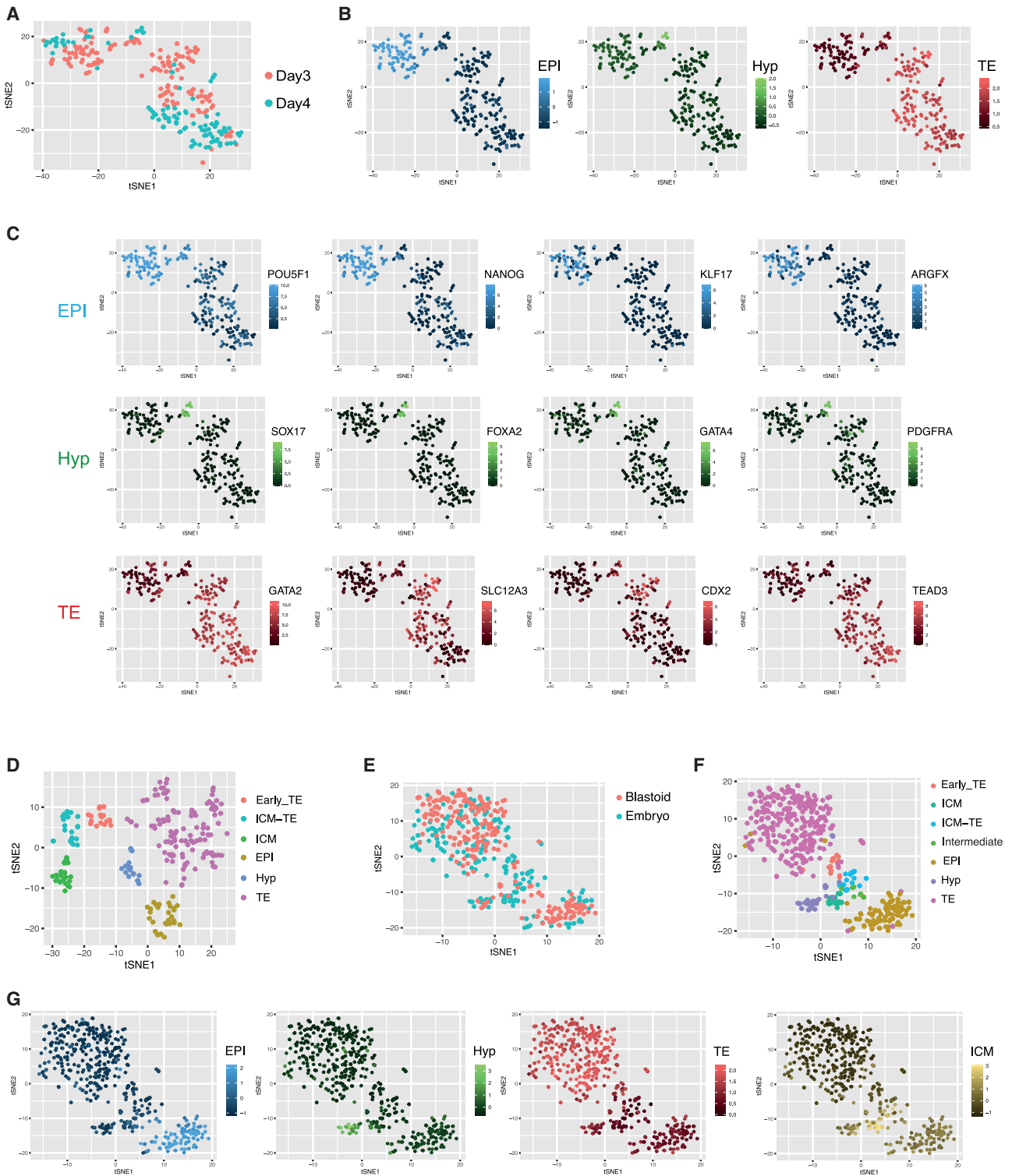
Overall, the localization of epiblast, hypoblast, and trophecto- derm markers indicates that the cysts attain a developmental stage comparable to the late human blastocyst (E6/7) in which the three founding embryo lineages are fully segregated. We therefore termed these cysts human blastoids.

After day 4 blastoid cavities collapsed, similar to human blas- tocysts by the end of day E7 in suspension (Shahbazi et al., 2016). When transferred to Geltrex-coated dishes, blastoids attached and proliferated to produce large outgrowths of GA- TA3:mKO2-positive cells (Figure 1H). Outgrowing cells dis- played cytokeratin 7 (CK7) and included regions expressing the syncytiotrophoblast marker human  $\beta$  chorionic gonadotro- phin (hCGB) (Figure 1I), similar to human blastocyst outgrowths (Deglincerti et al., 2016). In a portion of the cultures, a compact area of small cells was apparent that expressed the epiblast marker NANOG with adjacent cells expressing GATA4 (Figures 1I and 1J). In these regions we detected podocalyxin (PODXL) stained lumen structures (Bedzhov and Zernicka-Goetz, 2014; Shahbazi et al., 2016), suggestive of emergent yolk sac cavities and amniotic cavities (Figure 1J; Video S1). The cavities were bounded by NANOG-positive cells and embedded within exten- sive trophoblast outgrowths. Conventional hPSCs can self-orga- nize to form lumens and amnion-like cavities (Taniguchi et al., 2015; Zheng et al., 2019). Integration with differentiating tropho- blast is therefore necessary to assert progression from the blas- toid ICM. Nonetheless, more comprehensive analyses are required to make definitive statements regarding the emergence of these structures and their relationship to development of the bilaminar embryonic disc.

We performed single-cell transcriptome analysis, using SMART-seq2 (Picelli et al., 2013) to achieve suitable sequencing depth and facilitate reliable comparison with human embryo scRNA-seq data. To avoid over-representation of trophecto- derm cells, we manually excised most of the abembryonic region (mural trophectoderm) prior to dissociation. After quality control we analyzed transcriptomes from 159 day 3 and 108 day 4 blas- toid cells.

### Figure 1. Formation of model human blastocysts (blastoids)

- (A) Schematic of procedure.  
 (B) Phase contrast and fluorescence images showing formation of a GATA3:mKO2 positive cyst.  
 (C) z-projections of a day 3 blastoid and a human late blastocyst (E7) stained for epiblast (KLF17, NANOG), hypoblast (GATA4), and trophectoderm (GATA3) markers.  
 (D) Quantification of caviated mKO2-positive cyst formation related to cell number seeded.  
 (E) Quantification of diameter of cysts as in (D). Error bars are SD.  
 (F) Immunofluorescence staining for epiblast and trophectoderm markers during blastoid formation.  
 (G) Immunofluorescence staining for hypoblast marker SOX17 in day 3 blastoid.  
 (H) Phase contrast and fluorescence images of GATA3:mKO blastoid outgrowth after 4 days.  
 (I) z-projections of immunofluorescence staining of day 4 outgrowth stained for differentiated trophoblast (CK7) and syncytiotrophoblast (hCGB) markers and NANOG.  
 (J) z-projections of immunofluorescence staining of day 4 outgrowth stained for PODXL, NANOG, GATA4, and DAPI. Arrows point to cavities (see Video S1). Scale bars in all images are 50  $\mu\text{m}$ .



**Figure 2. Single-cell transcriptome analysis**

(A) t-distributed stochastic neighbor embedding (tSNE) of day 3 and day 4 blastoid cells.

(B) tSNE plot in (A) showing averaged expression scores of classifier genes (Table S1) enriched in human E6/7 embryo lineages, epiblast (EPI), hypoblast (Hyp), and trophoblast (TE).

(C) Expression (FPKM) of selected lineage markers on tSNE plot in (A).

(legend continued on next page)

t-distributed stochastic neighbor embedding (tSNE) revealed largely overlapping distributions on day 3 and day 4 (Figure 2A). Using exclusive lineage markers from the human E6/7 blastocyst (Guo et al., 2021; Petropoulos et al., 2016; Stirparo et al., 2018), almost all cells were assigned trophectoderm, epiblast, or hypoblast identities (Figures 2B and S2A). A small cluster of six cells (5 from day 3, 1 from day 4) was indeterminate between epiblast and hypoblast (Figure S2A). Averaged expression of classifier genes (Table S1) substantiated the lineage demarcations (Figure 2B). Illustrative marker gene profiles are shown in Figure 2C. Within the trophectoderm cluster, GATA2 and TEAD3 were uniformly high while early markers CDX2 (Niakan and Eggan, 2013) and SLC12A3 showed heterogeneous expression and were more prominent on day 3 than day 4. Naive status of the epiblast cluster was evidenced by KLF17 and ARGFX. PDGFRA, which discriminates hypoblast from definitive endoderm, authenticated the hypoblast cluster.

We also sequenced cells from E5, E6, and E7 human embryos, bisected to reduce the mural trophectoderm. Consistent with previous analyses (Blakeley et al., 2015; Petropoulos et al., 2016; Stirparo et al., 2018), tSNE separated early (E5) from late (E6/7) blastocyst stages (Figures 2D and S2B). Expression of lineage classifiers identified ICM, trophectoderm, epiblast, and hypoblast clusters (Figure S2B).

We performed integrated tSNE analysis on blastoid and embryo data (Figures 2E and S2C). With few exceptions, identities independently determined for the blastoid and embryo datasets aligned in the integrated plot (Figure 2F). That is, almost all blastoid cells clustered together with embryo cells assigned to the same lineage. The few misaligned cells mostly originated from the embryos. Blastoid cells clustered more densely with E6/7 embryo cells (Figure 2E), consistent with late blastocyst similarity. However, the small group of intermediate cells in the blastoid samples were most related to E5 embryo ICM. CDX2, SLC12A3, and GPRC5A were expressed in most E5 trophectoderm cells but were heterogenous in E6 and E7 samples (Figure S2E). These observations suggest that blastoids may transit through a stage with features of the E5 blastocyst.

Overall, transcriptome analyses show a high degree of fidelity in lineage segregation between our blastoids and human blastocysts. Trophectoderm, authenticated by multiple markers, is well represented and there is no indication of abundant mismatched cells.

We inspected scRNA-seq data from recent reports of human blastoids (Liu et al., 2021; Yu et al., 2021) (Figure S3). Trophectoderm is by far the major cell type in the expanded human blastocyst. However, tSNE analysis of data from Yu et al. (2021) showed only a minor population of trophectoderm-related cells with a preponderance of EPI-related cells (Figure S3A). Two distinct clusters expressed OCT4, SOX2, and NANOG but differed in expression of many genes including the embryo naive epiblast marker ETV4 (Figure S3B). Together they comprised more than 80% of cells, indicating a massive failure of pluripotent cells to differentiate or to acquire a consistent embryonic identity.

Furthermore, the trophectoderm-like cluster entirely lacked expression of the general trophoblast lineage marker TP63. Blastoids created by reprogramming (Liu et al., 2021) contained a majority of cells with ambiguous identities (Figures S3C and S3D).

## DISCUSSION

Availability of a reliable human blastocyst model will enable molecular and genetic dissection of early human embryogenesis, allow systematic exploration of culture conditions for assisted reproduction, and open a path for research into peri- and early post-implantation development. Our findings demonstrate that the innate lineage plasticity of human early epiblast and naive stem cells (Guo et al., 2021; Io et al., 2021) can be efficiently channelled into formation of entities with strong resemblance to the natural human blastocyst. Crucially, blastoid cells exhibited topologically restricted and mutually exclusive expression of markers of the three lineages, while scRNA-seq analysis confirmed the assignment of almost all blastoid cells to blastocyst stage trophectoderm, hypoblast, or epiblast.

Ability to form an authentic blastocyst-like entity solely from naive stem cells is a marked difference between human and mouse. Mouse embryonic stem cells are developmentally restricted from forming trophectoderm and blastoids are therefore created by combining cells of different lineages (Posfai et al., 2021; Rivron et al., 2018). In contrast, human naive stem cells can readily be induced into trophectoderm. In cell clusters, the trophectoderm cells form an outer epithelial monolayer that encloses persisting undifferentiated naive cells and initiates polarized ion transport. The resulting fluid uptake leads to formation of an expanded cavity. Induction of trophectoderm with PD+A83 should be limited in duration to preserve a naive inner cell population and allow hypoblast differentiation. Other epithelia can form cystic structures and cavitation alone is not sufficient to assert trophectoderm identity. The GATA3:mKO2 reporter, which is highly expressed in all trophectoderm cells, provides reassurance.

Blastoid formation results in an entity with high similarity to the human blastocyst developed after *in vitro* fertilization. Interestingly, however, the naive epiblast starting stage for blastoids is distinct from the embryological origin, compacted morula. Therefore, the initial differentiation and morphokinetic pathways are dissimilar. Nonetheless, a three-dimensional structure of appropriate size and shape is formed in which the three lineages are correctly specified and segregated. This outcome implies a flexible coupling between differentiation and morphogenesis. Understanding how structures with equivalent topology and composition are generated from different embarkation points will shed light into the resilience of tissue formation.

Overall, our findings demonstrate the capability for generation of human blastoids at scale in simple and defined conditions. The blastoids fulfil several key criteria for a useful model of the human blastocyst: (1) correct topological segregation of lineages

(D) tSNE analysis of human embryos at E5, E6, E7 with lineage assignment by marker profile. ICM\_TE denotes ambiguous ICM and TE identity.

(E) Integrated tSNE plot for blastoid and embryo cells.

(F) tSNE plot in (E) colored to show cell lineage assignments from Figures S1A and S1D.

(G) tSNE in (E) showing averaged expression scores of lineage classifiers.

evidenced by appropriately localized and mutually exclusive expression of multiple marker proteins; (2) clustering of scRNA-seq data into three unambiguous lineages; (3) high transcriptome fidelity with the human embryo, with few or no unasigned cells; (4) single epiblast population with naive features; (5) timescale of morphogenesis similar to human blastocyst formation (3–4 days); (6) coordinated progression of morphogenesis and lineage segregation; (7) robust and scalable procedure with a high yield of cavitated tri-lineage structures (>80%); and (8) consistency across multiple stem cell lines.

### Limitations of the study

Differentiation dynamics vary slightly between human naive stem cell lines. For high efficiency, the cell seeding number and duration of PD+A83 treatment may need fine-tuning for individual lines. Microfluidics and bioengineering approaches (Zheng et al., 2019) may yield a more standardized procedure. We also observe that hypoblast cell number varies and is not proportionate to the number of epiblast cells in blastoids. However, the presentation of hypoblast cells in the human embryo is variable. The blastoid system provides an opportunity to delineate stimuli and mechanisms that regulate specification and sorting of hypoblast in human embryos. Gene expression is similar but not identical between blastoids and blastocysts and this may also be a fruitful area for investigation. Finally, the present study focuses on characterization of a blastocyst-stage entity with only preliminary analysis of further developmental potential. A future prospect is to use the blastoid model to develop improved culture conditions that reliably mimic peri- and early post-implantation embryogenesis. This will require specific, objective, and quantifiable measures of morphogenesis and differentiation, ideally combined with functional assays. Better systems for extended culture can then be applied to human embryos for direct comparison of developmental trajectories and potential.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Human embryos
  - Cell culture
- **METHOD DETAILS**
  - Human naive pluripotent stem cell culture
  - Generation of blastoids
  - Attachment culture of blastoids
  - Human embryos
  - Immunostaining
  - Imaging
  - Transcriptome sequencing
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - RNA-Seq data processing
  - Transcriptome analysis

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.stem.2021.04.031>.

### ACKNOWLEDGMENTS

We are grateful to Takuya Azami and Elena Corujo-Simon for assistance with human embryo thawing, James Clarke and Francesca Carlisle for laboratory assistance, Giuliano Stirparo for advice on transcriptome analysis, Maike Parmor and Vicki Murray for generating sequencing libraries, and Peter Humphreys, Darran Clements, and Corin Liddle for imaging support. Sequencing was performed by the CRUK Cambridge Institute Genomics Core Facility. We thank support staff in Cambridge and Exeter for enabling research to continue during the COVID-19 pandemic. This research was funded by the Medical Research Council (MRC) of the United Kingdom (MR/P00072X/1). The Cambridge Stem Cell Institute receives core funding from Wellcome (203151/Z/16/Z) and the MRC (MC\_PC\_12009). A.S. is an MRC Professor (G1100526/1).

### AUTHOR CONTRIBUTIONS

Conceptualization, G.G.; Methodology, G.G., A.Y.; Investigation, A.Y., J.N., A.D., G.G.; Formal analysis, D.S.; Writing, A.Y., G.G., A.S.; Supervision, G.G., A.S.

### DECLARATION OF INTERESTS

G.G. and A.S. are inventors on a patent relating to human naive pluripotent stem cells filed by the University of Cambridge.

### INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. One or more of the authors of this paper self-identifies as living with a disability.

Received: April 9, 2021

Revised: April 27, 2021

Accepted: April 28, 2021

Published: May 5, 2021

### REFERENCES

- Bedzhov, I., and Zernicka-Goetz, M. (2014). Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* 156, 1032–1044.
- Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaitha, S.E., Hu, T.X., Elder, K., Snell, P., Christie, L., Robson, P., and Niakan, K.K. (2015). Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* 142, 3151–3165.
- Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., and Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev. Cell* 35, 366–382.
- Boroviak, T., Stirparo, G.G., Dietmann, S., Hernando-Herraez, I., Mohammed, H., Reik, W., Smith, A., Sasaki, E., Nichols, J., and Bertone, P. (2018). Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* 145, dev167833.
- Bredenkamp, N., Stirparo, G.G., Nichols, J., Smith, A., and Guo, G. (2019a). The Cell-Surface Marker Sushi Containing Domain 2 Facilitates Establishment of Human Naive Pluripotent Stem Cells. *Stem Cell Reports* 12, 1212–1222.
- Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., von Meyenn, F., Dietmann, S., Baker, D., Drummond, R., Ren, Y., Li, D., et al. (2019b). Wnt Inhibition Facilitates RNA-Mediated Reprogramming of Human Somatic Cells to Naive Pluripotency. *Stem Cell Reports* 13, 1083–1098.

- Deglincerti, A., Croft, G.F., Pietila, L.N., Zernicka-Goetz, M., Siggia, E.D., and Brivanlou, A.H. (2016). Self-organization of the in vitro attached human embryo. *Nature* **533**, 251–254.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
- Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., and Nichols, J. (2016). Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**, 437–446.
- Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., et al. (2017). Epigenetic resetting of human pluripotency. *Development* **144**, 2748–2763.
- Guo, G., Stirparo, G.G., Strawbridge, S.E., Spindlow, D., Yang, J., Clarke, J., Dattani, A., Yanagida, A., Li, M.A., Myers, S., et al. (2021). Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell*. Published online April 3, 2021. <https://doi.org/10.1016/j.stem.2021.02.025>.
- Io, S., Kabata, M., Iemura, Y., Semi, K., Morone, N., Minagawa, A., Wang, B., Okamoto, I., Nakamura, T., Kojima, Y., et al. (2021). Capturing human trophoblast development with naive pluripotent stem cells in vitro. *Cell Stem Cell*. Published online April 5, 2021. S1934-5909(21)00119-3.
- Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930.
- Liu, X., Tan, J.P., Schröder, J., Aberkane, A., Ouyang, J.F., Mohenska, M., Lim, S.M., Sun, Y.B.Y., Chen, J., Sun, G., et al. (2021). Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* **591**, 627–632.
- Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Dev. Biol.* **375**, 54–64.
- Pastor, W.A., Liu, W., Chen, D., Ho, J., Kim, R., Hunt, T.J., Lukianchikov, A., Liu, X., Polo, J.M., Jacobsen, S.E., and Clark, A.T. (2018). TFAP2C regulates transcription in human naive pluripotency by opening enhancers. *Nat. Cell Biol.* **20**, 553–564.
- Petropoulos, S., Edsgård, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165**, 1012–1026.
- Picelli, S., Björklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1098.
- Posfai, E., Schell, J.P., Janiszewski, A., Rovic, I., Murray, A., Bradshaw, B., Yamakawa, T., Pardon, T., El Bakkali, M., Talon, I., et al. (2021). Evaluating totipotency using criteria of increasing stringency. *Nat. Cell Biol.* **23**, 49–60.
- Rivron, N.C., Frias-Aldeguer, J., Vrij, E.J., Boisset, J.C., Korving, J., Vivié, J., Truckenmüller, R.K., van Oudenaarden, A., van Blitterswijk, C.A., and Geijsen, N. (2018). Blastocyst-like structures generated solely from stem cells. *Nature* **557**, 106–111.
- Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A., and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. *Dev. Biol.* **361**, 358–363.
- Rossant, J. (2018). Genetic Control of Early Cell Lineages in the Mammalian Embryo. *Annu. Rev. Genet.* **52**, 185–201.
- Rossant, J., and Tam, P.P.L. (2021). Opportunities and Challenges with Stem Cell-Based Embryo Models. *Stem Cell Reports*. Published online March 4, 2021. S2213-6711(21)00082-5.
- Rostovskaya, M., Stirparo, G.G., and Smith, A. (2019). Capacitation of human naive pluripotent stem cells for multi-lineage differentiation. *Development* **146**, dev172916.
- Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.N.M., Campbell, A., Devito, L., Ilic, D., et al. (2016). Self-organization of the human embryo in the absence of maternal tissues. *Nat. Cell Biol.* **18**, 700–708.
- Sozen, B., Cox, A.L., De Jonghe, J., Bao, M., Hollfelder, F., Glover, D.M., and Zernicka-Goetz, M. (2019). Self-Organization of Mouse Stem Cells into an Extended Potential Blastoid. *Dev. Cell* **51**, 698–712.e8.
- Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., and Bertone, P. (2018). Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* **145**, dev158501.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21.
- Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficzi, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269.
- Taniguchi, K., Shao, Y., Townshend, R.F., Tsai, Y.H., DeLong, C.J., Lopez, S.A., Gayen, S., Freddo, A.M., Chue, D.J., Thomas, D.J., et al. (2015). Lumen Formation Is an Intrinsic Property of Isolated Human Pluripotent Stem Cells. *Stem Cell Reports* **5**, 954–962.
- Yu, L., Wei, Y., Duan, J., Schmitz, D.A., Sakurai, M., Wang, L., Wang, K., Zhao, S., Hon, G.C., and Wu, J. (2021). Blastocyst-like structures generated from human pluripotent stem cells. *Nature* **591**, 620–626.
- Zheng, Y., Xue, X., Shao, Y., Wang, S., Esfahani, S.N., Li, Z., Muncie, J.M., Lakins, J.N., Weaver, V.M., Gumucio, D.L., and Fu, J. (2019). Controlled modelling of human epiblast and amnion development using stem cells. *Nature* **573**, 421–425.



STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Rabbit polyclonal anti-KLF17	Atlas Antibodies	Cat# HPA024629, RRID:AB_1668927
Goat polyclonal anti-NANOG	R&D System	Cat# AF1997, RRID:AB_355097
Rabbit polyclonal anti-NANOG	Abcam	Cat# ab21624, RRID:AB_446437
Rabbit polyclonal anti-Oct-4	Cell Signaling Technology	Cat# 2750, RRID:AB_823583
Mouse monoclonal anti-Sox2	Santa Cruz	Cat# sc-365823, RRID:AB_10842165
Rat monoclonal anti-Gata-4	Thermo Fisher Scientific	Cat# 14-9980-82, RRID:AB_763541
Goat polyclonal anti-SOX17	R&D System	Cat# AF1924, RRID:AB_355060
Goat polyclonal anti-Otx2	R&D System	Cat# AF1979, RRID:AB_2157172
Mouse monoclonal anti-Gata3	Thermo Fisher Scientific	Cat# MA1-028, RRID:AB_2536713
Mouse monoclonal anti-AP2 gamma	Santa Cruz	Cat# sc-12762, RRID:AB_667770
Rabbit monoclonal anti-CK7	Abcam	Cat# ab181598, RRID:AB_2783822
Mouse monoclonal anti-CK18	Abcam	Cat# ab668, RRID:AB_305647
Rabbit monoclonal anti-CK18	Abcam	Cat# ab133263, RRID:AB_11155542
Goat polyclonal anti-Lamin B	Santa Cruz	Cat# sc-6217, RRID:AB_648158
Donkey anti-Goat Alexa Fluor 405	Abcam	Cat# ab175664, RRID:AB_2313502
Donkey anti-Goat Alexa Fluor 488	Thermo Fisher Scientific	Cat# A32814, RRID:AB_2762838
Donkey anti-Rabbit Alexa Fluor 488	Thermo Fisher Scientific	Cat# A32790, RRID:AB_2762833
Donkey anti-Rat Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-21208, RRID:AB_2535794
Donkey anti-Goat Alexa Fluor 555	Thermo Fisher Scientific	Cat# A-21432, RRID:AB_2535853
Donkey anti-Rabbit Alexa Fluor 555	Thermo Fisher Scientific	Cat# A-31572, RRID:AB_162543
Donkey anti-Mouse Alexa Fluor 555	Thermo Fisher Scientific	Cat# A-31570, RRID:AB_2536180
Donkey anti-Goat Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-21447, RRID:AB_2535864
Donkey anti- Rabbit Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-31573, RRID:AB_2536183
Donkey anti- Mouse Alexa Fluor 647	Thermo Fisher Scientific	Cat# A-31571, RRID:AB_162542
Rabbit monoclonal anti-CK7	Abcam	Cat# ab181598 RRID:AB_2783822
Mouse monoclonal anti-hCGB	Abcam	Cat# ab9582 RRID:AB_296507
Mouse monoclonal anti-Podocalyxin	R&D System	Cat# MAB1658, RRID:AB_2165984
<b>Chemicals, peptides, and recombinant proteins</b>		
MEK inhibitor PD0325901	ABCR	Cat#AB 253775
Tankyrase inhibitor XAV939	Cell Guidance Systems	Cat#SMS38-200
aPKC inhibitor Gö6983	Bio-Techne	Cat#2285
Rho associated kinase inhibitor Y-27632	Merck Chemicals	Cat#688000-100MG
Human leukemia inhibitory factor (LIF)	Made in-house	N/A
Activin/nodal receptor inhibitor A83-01	Generon	Cat#A12358-50
<b>Complete Culture Media and Cell Dissociation Reagent</b>		
N2B27	Made in-house	N/A
Accutase	Millipore	Cat#SCR005
TrypLE™ Express Enzyme	Thermo Fisher Scientific	Cat#12605028
<b>Cell Attachment Proteins and Peptides</b>		
Geltrex	Thermo Fisher Scientific	Cat#A1413302
<b>Deposited data</b>		
scRNaseq	This study	GSE171820

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
scRNaseq	(Petropoulos et al., 2016)	E-MTAB-3929
scRNaseq	(Yu et al., 2021)	GSE150578
scRNaseq	(Liu et al., 2021)	GSE156596
<b>Experimental models: cell lines</b>		
HNES1-GATA3:mKO2	(Guo et al., 2021)	N/A
niPSC HDF75	(Bredenkamp et al., 2019b)	N/A
cR-Shef6	(Guo et al., 2017)	N/A
<b>Software and algorithms</b>		
STAR v2.7.7a	(Dobin et al., 2013)	<a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>
featureCounts (Subread v2.0.1)	(Liao et al., 2014)	<a href="http://subread.sourceforge.net/">http://subread.sourceforge.net/</a>
Rtsne v0.15		<a href="https://github.com/jkrijthe/Rtsne">https://github.com/jkrijthe/Rtsne</a>
Harmony v1.0	(Korsunsky et al., 2019)	<a href="https://github.com/immunogenomics/harmony">https://github.com/immunogenomics/harmony</a>
Seurat v4.0	(Stuart et al., 2019)	<a href="https://satijalab.org/seurat/">https://satijalab.org/seurat/</a>
Trim-Galore! v0.6.6		<a href="https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/">https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/</a>
R v4.0.3		<a href="https://www.r-project.org/">https://www.r-project.org/</a>

**RESOURCE AVAILABILITY**

**Lead Contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ge Guo, [g.guo@exeter.ac.uk](mailto:g.guo@exeter.ac.uk)

**Materials Availability**

All stable reagents generated in this study are available from the Lead Contact without restriction except for human embryo-derived cell lines, for which permission must be requested from the UK Stem Cell Steering Committee and a Materials Transfer Agreement completed.

**Data and Code Availability**

The RNaseq data generated in this paper are deposited in Gene Expression Omnibus with accession code GSE171820

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

**Human embryos**

The use of supernumerary human embryos in this research is approved by the Multi-Centre Research Ethics Committee, approval O4/MRE03/44, and licensed by the Human Embryology & Fertilization Authority of the United Kingdom, research license R0178.

**Cell culture**

Cell lines are listed in the Key Resources Table. Cells were cultured in humidified incubators at 37°C in 7% CO<sub>2</sub> and 5% O<sub>2</sub>. Cells were cultured without antibiotics and tested negative for mycoplasma by periodic PCR screening.

**METHOD DETAILS**

**Human naive pluripotent stem cell culture**

Human naive pluripotent stem cells were propagated in PXGL medium as described (Bredenkamp et al., 2019a). PXGL comprises 1 μM PD0325901 (P), 2 μM XAV939 (X), 2 μM Gö6983 (G) and 10 ng/mL human LIF (L) in N2B27 basal medium. Cells were cultured on irradiated MEF feeders. Rho associate kinase inhibitor (Y-27632) and Geltrex (0.5 μL per cm<sup>2</sup> surface area, Thermo Fisher Scientific, A1413302,) were added to media during replating. Cells were passaged by dissociation with Accutase (Biolegend, 423201) every 3-4 days.

### Generation of blastoids

PXGL cultures of naive stem cells in exponential growth were dissociated with TrypLE for 5 min and pelleted in washing medium (DMEM/F12 supplemented with 0.1% BSA). The pellet was resuspended in PD+A83+Y medium (N2B27 supplemented with 1.5  $\mu$ M PD0325901, 1  $\mu$ M A83-01 and 10  $\mu$ M Y-27632). Cell number was counted using a hemacytometer and cell density was adjusted to 1000 cell per ml. 50–200 cells were dispensed into each well of an ultra-low attachment multiple-well plate (Corning Coster) with a multichannel pipette. The plates were centrifuged at 300 g for 4 min at room temperature to cluster cells at the bottom of the wells. After 42–48 h, cell aggregates were manually transferred into a non-adherent, 'U'-bottomed 96-well (Greiner) containing pre-warmed N2B27 supplemented with 0.5  $\mu$ M A83-01) using a mouth-controlled pipette. At the end of day 3, cysts were manually transferred into N2B27 medium without either inhibitor.

### Attachment culture of blastoids

Day 4 blastoids were moved manually from a multiwell plate to a  $\mu$ -Slide 8-well (ibidi) coated with geltrex and cultured in N2B27 medium. Half of the medium was replaced every day. Blastoids usually attached to the plate and began to outgrow within two days. Cultures were fixed for staining after 4 days of attachment culture

### Human embryos

Supernumerary frozen blastocysts (E5 or E6) were thawed and cultured in N2B27 medium under mineral oil. Embryos that were well-expanded after thawing were judged to be E6 and processed immediately, while the majority of embryos were cultured for 24 h for development to expanded blastocysts (E6 or E7). The zona pellucida was removed using acid Tyrode's solution.

### Immunostaining

Blastoids, outgrowths and embryos were fixed with 4% PFA in PBS for 15 min at room temperature. Samples were rinsed in PBS containing 3 mg/mL polyvinylpyrrolidone (PBS/PVP) and permeabilized with PBS/PVP containing 0.25% Triton X-100 for 30 min. Blocking was performed in embryo blocking buffer comprising PBS supplemented with 0.1% BSA, 0.01% Tween20 and 2% donkey serum for 2–3 h at 4°C. Samples were incubated in the blocking buffer with or without 500 ng/mL DAPI for 1 h at room temperature in the dark. DAPI-stained samples were rinsed three times for 15 min in blocking buffer. The staining process for blastoids and embryos was performed in microwell miniTrays.

### Imaging

Blastoids and embryos were transferred in small drops of blocking buffer onto poly-D-lysine-coated glass-bottom dishes under mineral oil or  $\mu$ -slide 18 well-Flat and imaged using a Leica TCS SP5 confocal microscope. Images were processed using Fiji. Widefield images were taken using Leica DMI4000 or DMI8. Numbers of GATA4+ or KLF17+ cells were counted manually from the confocal images. The diameter of each blastoid was measured using Fiji.

### Transcriptome sequencing

Embryos and blastoids were placed singly in drops of M2 medium and the mural trophectoderm (mTE) was manually excised using finely drawn glass needles. mTE fragments were transferred to drops of Accutase and the ICM plus polar trophectoderm samples were transferred to TrypLE using glass capillaries or micropipettes. After incubation for 7–10 min at 37°C for Accutase or room temperature for TrypLE. Samples were transferred to drops of filtered M2 and dissociated by trituration using a finely drawn glass capillary with diameter just larger than a single cell. Single cells were transferred into individual wells of a 96-well plate containing Smart-seq2 single-cell lysis buffer and immediately frozen. Smart-seq2 libraries were prepared as described (Picelli et al., 2013) and sequenced using the Illumina Novaseq platform in paired end format.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### RNA-Seq data processing

Adaptor sequences and low-quality base calls were trimmed using Trim Galore! v0.6.6 with default parameters. Trimmed reads were aligned to the Human reference genome (GRCh38) using STAR v2.7.7a (Dobin et al., 2013) in paired-end mode with default parameters. Gene level counts were generated using FeatureCounts from Subread v2.0.1 (Liao et al., 2014).

### Transcriptome analysis

Cells with detection of fewer than 6000 expressed genes were discarded and the remaining samples were log<sub>2</sub>FPKM normalized using custom scripts. Only protein coding genes were considered in the analyses. tSNE (t-distributed stochastic neighbor embedding) plots were computed using variable genes across the dataset under analysis with Rtsne v0.15 (M. Krijthe, 2015) and visualized using ggplot2 in R v4.0.3. The AddModuleScore function from Seurat v4.0 (Stuart et al., 2019) was used to score and color cells according to the expression of signature gene lists for different cell lineages derived from (Petropoulos et al., 2016; Stirparo et al., 2018). To

obtain the signature gene list, FPKM values were used to compute Pearson's correlation coefficients with a binary feature vector for each cell population consisting of values of 1 for the considered cell type and 0 for other cell types. Genes with a correlation coefficient of greater than 0.4 were used as lineage classifiers ([Table S1](#)).

Harmony v1.0 ([Korsunsky et al., 2019](#)) was used for batch correction between the embryo and blastoid datasets using the first 50 principle components produced via principal component analysis.

**Cell Stem Cell, Volume 28**

**Supplemental Information**

**Naive stem cell blastocyst model**

**captures human embryo lineage segregation**

**Ayaka Yanagida, Daniel Spindlow, Jennifer Nichols, Anish Dattani, Austin Smith, and Ge Guo**

## SUPPLEMENTAL INFORMATION INDEX

Figure S1. Formation of human blastocyst models (blastoids), relates to Figure 1

Figure S2. Single cell transcriptome analyses, relates to Figure 2

Figure S3. Single cell transcriptome analysis of previously reported blastoids, relates to Figure 2

## SUPPLEMENTAL LEGENDS

### Figure S1. Formation of model human blastocysts (blastoids), relates to Figure 1

- A. Images of day 3 HNES1-*GATA3:mKO2* cysts formed from initial cell numbers of 100 or 125. Images were taken from a continuous area of a 96 well plate to represent the range of morphologies.
- B. Images of disorganized, aggregate, single cyst or multiple cyst structures on day 3.
- C. Quantification of structures on day 3.
- D. Images of day 3 cysts with no ICM (left, GATA3+), no hypoblast (middle, GATA3+, KLF17+), and with three lineages (right, KLF17+, GATA4+, GATA3+).
- E. Pie chart of lineage marker expression in day 3 blastoids.
- F. Quantification of the number of KLF17 or GATA4 positive inner cells in blastoids.
- G. Images of day 3 niPSC\_HDF75 blastoids in contiguous wells. Initial cell number seeded was 150.
- H. Immunofluorescence staining for OTX2 in day 3 HNES1mKO2 blastoid.
- I. Z-projections of day 3 blastoids formed from reset cR-Shef6 naïve cells
- J. Z-projections of day 4 blastoids formed from niPSC\_HDF75 naïve cells. All scale bars are 50  $\mu\text{m}$ .

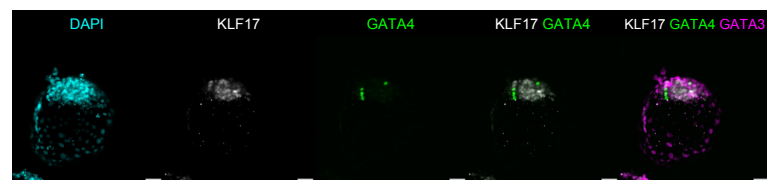
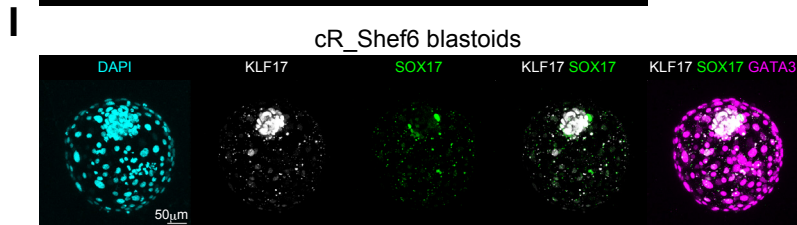
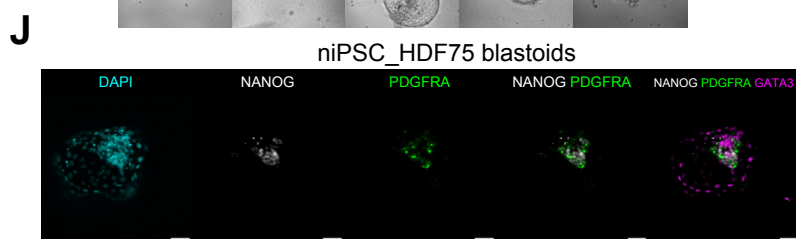
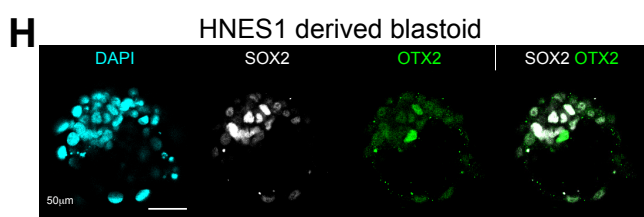
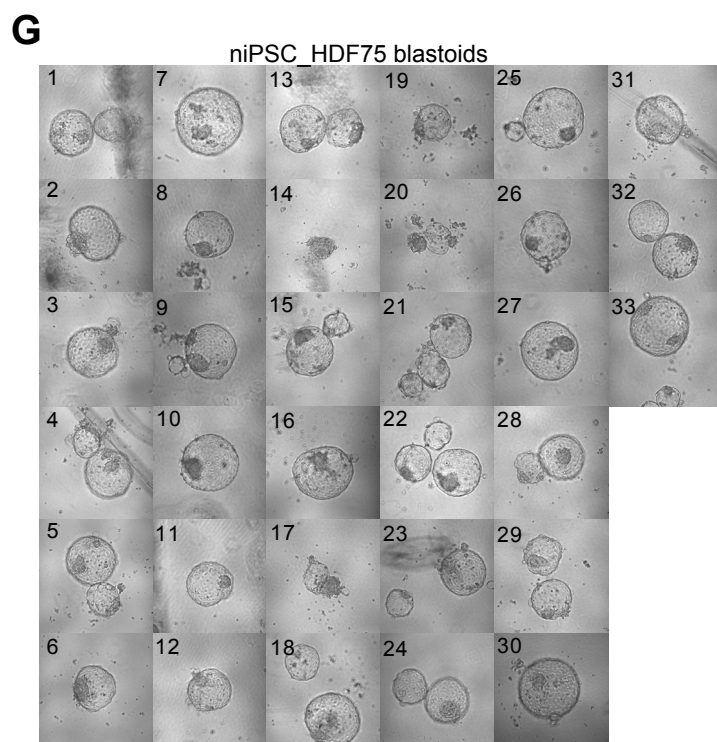
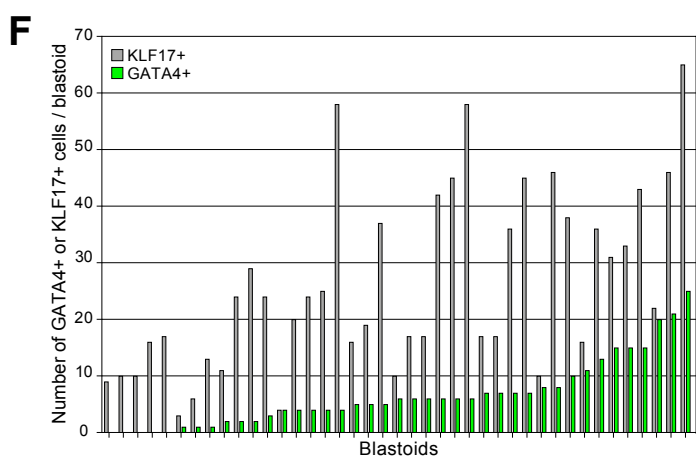
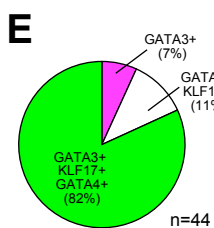
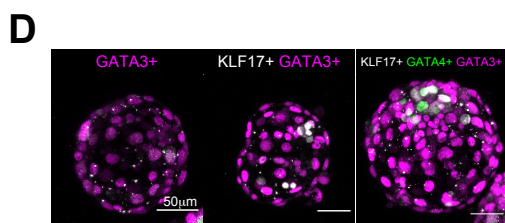
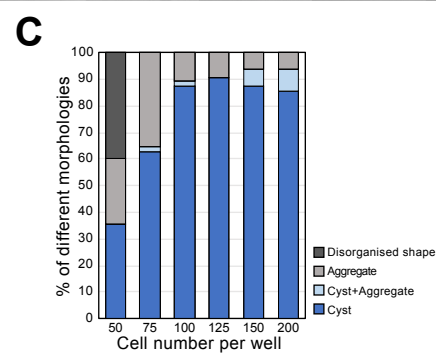
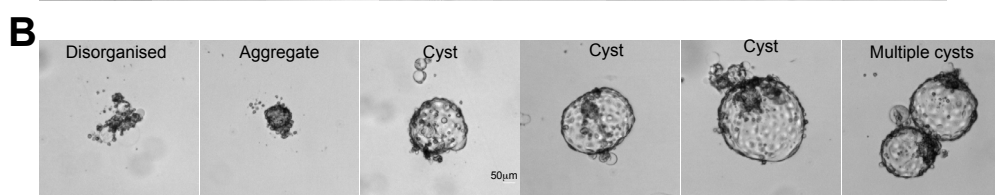
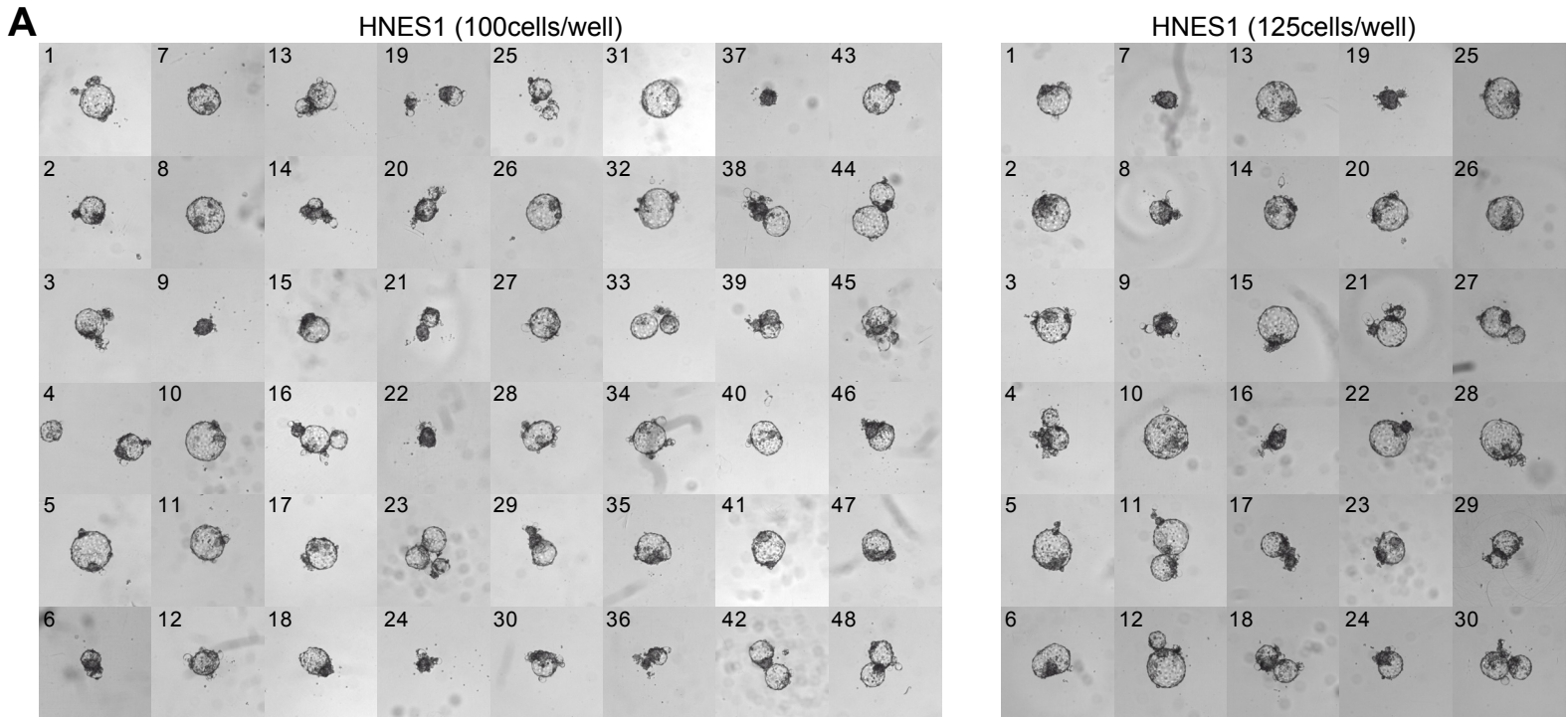
### Figure S2. Single cell transcriptome analysis, relates to Figure 2

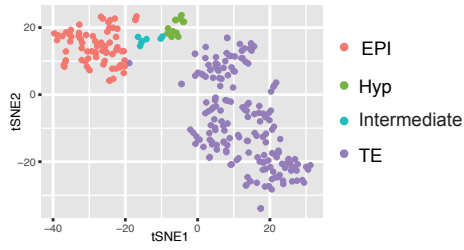
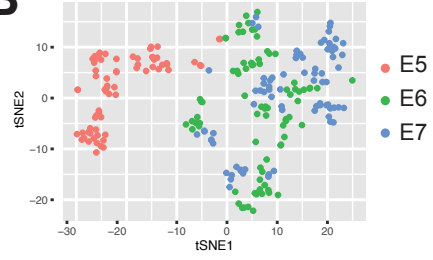
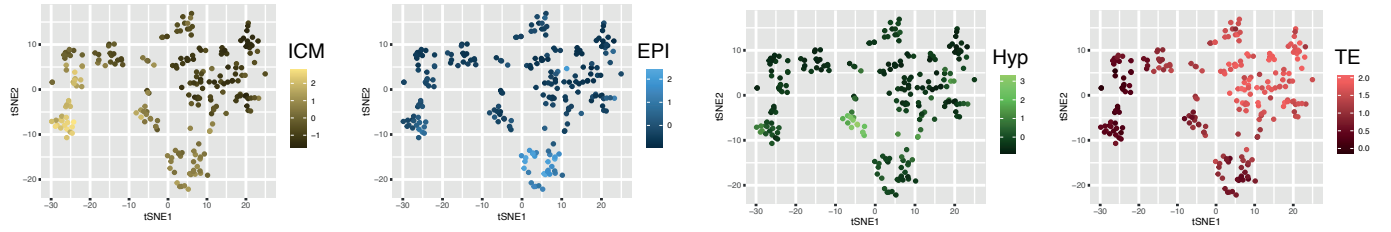
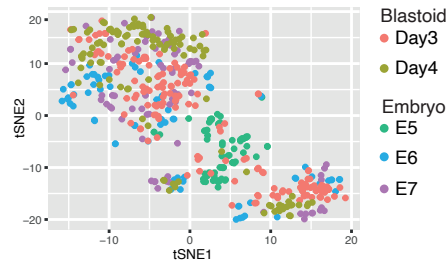
- A. t-distributed stochastic neighbour embedding (tSNE) analysis of blastoids coloured to show marker-based cell lineage assignment. EPI, epiblast; Hyp, hypoblast; TE, trophectoderm; intermediate, cells without clear signature
- B. tSNE analysis of human embryos.
- C. tSNE plot in B coloured to show averaged expression scores of gene sets enriched in human embryo lineages (Petropoulos et al., Cell, 2016): EPI, epiblast (E6/7); Hyp hypoblast (E6/7); TE, trophectoderm (E6/7); ICM (E5).
- D. Integrated tSNE of blastoids and embryos
- E. Expression (FPKM) of selected lineage markers in integrated tSNE plots.

### Figure S3. Single cell transcriptome analysis, relates to Figure 2

scRNA-seq data from (Yu et al., 2021) and (Liu et al., 2021) were downloaded and analysed as in Figure 2 and Figure S2.

- A. tSNE analysis of day 9 blastoid data from Yu et al. (2021) coloured to show averaged expression scores of embryo lineage classifiers.
- B. Expression (FPKM) of selected lineage markers in samples from Yu et al. (2021).
- C. tSNE analysis of induced blastoid data from Liu et al. (2021) coloured to show averaged expression scores of embryo lineage classifiers.
- D. Expression (FPKM) of selected lineage markers in samples from Liu et al. (2021)



**A****B****C****D****E**