Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Revealing cell populations catching the early stages of human embryo development in naive pluripotent stem cell cultures

Marta Moya-Jódar,[1,2,8] Asier Ullate-Agote,[1,3,8] Paula Barlabé,[1,2] Juan Roberto Rodríguez-Madoz,[4,7] Gloria Abizanda,[1,2] Carolina Barreda,[1,2] Xonia Carvajal-Vergara,[1,2] Amaia Vilas-Zornoza,[3,7] Juan Pablo Romero,[3,5] Leire Garate,[4,7] Xabier Agirre,[4,7] Giulia Coppiello,[1,2] Felipe Prósper,[1,2,4,6,7,*] and Xabier L. Aranguren[1,2,*]

[1]Program of Regenerative Medicine, Center for Applied Medical Research (CIMA), University of Navarra, Pamplona 31008, Spain
[2]Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona 31008, Spain
[3]Advanced Genomics Laboratory, Program of Hemato-Oncology, Center for Applied Medical Research (CIMA), University of Navarra, Pamplona, Spain
[4]Hemato-Oncology Program, Center for Applied Medical Research (CIMA), IDISNA, University of Navarra, Pamplona, Spain
[5]10x Genomics, 6230 Stoneridge Mall Road, Pleasanton, CA 94588, USA
[6]Hematology Department, Clínica Universidad de Navarra, University of Navarra, Pamplona, Spain
[7]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Pamplona, Spain
[8]These authors contributed equally
*Correspondence: fprosper@unav.es (F.P.), xlaranguren@unav.es (X.L.A.)
https://doi.org/10.1016/j.stemcr.2022.11.015

SUMMARY

Naive human pluripotent stem cells (hPSCs) are defined as the *in vitro* counterpart of the human preimplantation embryo's epiblast and are used as a model system to study developmental processes. In this study, we report the discovery and characterization of distinct cell populations coexisting with epiblast-like cells in 5iLAF naive human induced PSC (hiPSC) cultures. It is noteworthy that these populations closely resemble different cell types of the human embryo at early developmental stages. While epiblast-like cells represent the main cell population, interestingly we detect a cell population with gene and transposable element expression profile closely resembling the totipotent eight-cell (8C)-stage human embryo, and three cell populations analogous to trophectoderm cells at different stages of their maturation process: transition, early, and mature stages. Moreover, we reveal the presence of cells resembling primitive endoderm. Thus, 5iLAF naive hiPSC cultures provide an excellent opportunity to model the earliest events of human embryogenesis, from the 8C stage to the peri-implantation period.

INTRODUCTION

Embryonic development starts from a single totipotent cell called a zygote, derived from the fertilized egg. Maternally derived RNA is used in the first phases of embryonic development, in particular, until the two-cell (2C) stage in mouse and the eight-cell (8C) stage in human, when zygotic genome activation (ZGA) occurs, and the first wave of embryo genome transcription starts. In humans, the first lineage differentiation events take place in the compact morula at day 4, with the outer cells initiating a trophectoderm (TE)-specific transcriptional program. As a result, 5 days after fertilization, the blastocyst is segregated into the inner cell mass (ICM), which will give rise to the embryo proper, and the TE, which supports uterine implantation. By day 6, the blastocyst's ICM differentiates into the primitive endoderm (PrE), which will be a major constituent of the yolk sac, and the epiblast, which will form the fetus, while the TE will generate the extraembryonic tissues, including the placenta. Therefore, at the late blastocyst stage, before the implantation, these three main cell lineages are already determined (Shahbazi, 2020).

Human pluripotent stem cells (hPSCs) can be obtained and maintained either in primed state, representing the *in vitro* counterpart of the post-implantation epiblast, or in naive state, corresponding to the preimplantation epiblast. Diverse protocols have been established to obtain hPSCs with naive characteristics (Gafni et al., 2013; Guo et al., 2016, 2017; Liu et al., 2017; Takashima et al., 2014; Theunissen et al., 2014). Comparative transcriptional and molecular analyses have demonstrated that, among these protocols, the 5iLA, t2i/L + Gö, and PXGL culture conditions produce cells with the highest resemblance to the *in vivo* preimplantation epiblast (Pastor et al., 2016; Stirparo et al., 2018; Takashima et al., 2014; Theunissen et al., 2014, 2016).

Certain studies have identified different subpopulations of naive hPSCs in 5iLAF cultures based on the expression levels of SSEA4 (Pastor et al., 2016); moreover, expression of the PrE master gene GATA6 has been detected in different naive cultures (Guo et al., 2016; Linneberg-Agerholm et al., 2019) and TE markers, such as GATA3 and CDX2, have been identified in 5iLA cultures (Dong et al., 2020), suggesting the presence of different cell populations representing distinct cell lineages of the human preimplantation embryo. However, a single-cell transcriptomics study did not show significant heterogeneity in human embryonic stem cells (hESCs) in t2i/L + Gö naive conditions (Messmer et al., 2019). Nevertheless, they interestingly identified a

very small intermediate population presenting a gene expression profile separate from both naive and primed states (Messmer et al., 2019). On the other hand, there is significant evidence proving the existence of heterogeneous cell populations in mouse naive PSCs cultures: cells equivalent to ICM (identified by the double expression of PDGFRα and CD31), PrE-precursors (PDGFRα$^+$; CD31$^-$) (Io Nigro et al., 2017) and two-cell-like cells (2CLCs) (Macfarlan et al., 2012) have been described. 2CLCs possess some characteristics proper of the 2C stage mouse embryo as they express genes from the *Zscan4* family and MERVL retrotransposons, while showing downregulation of proteins associated with pluripotency, such as Pou5f1, Sox2, and Nanog, a fingerprint that correlates with the ZGA taking place in the 2C mouse embryo (Macfarlan et al., 2012). In addition, a cell population transitioning from naive to 2CLCs state was identified, being characterized by *Zscan4* expression and lack of MERVL retrotransposons expression (Rodriguez-Terrones et al., 2018). Analogously to what was described in mouse, two recent publications (Mazid et al., 2022; Taubenschmid-Stowers et al., 2022) have reported the presence of a population of eight-cell-like cells (8CLCs) within human naive cultures, which resemble the human embryo at the 8C stage.

In this study, to fill the void of knowledge regarding hPSC heterogeneity, we analyzed by single-cell RNA sequencing (scRNA-seq) naive human induced PSCs (hiPSCs) under 5iLAF culture conditions, finding different cell populations that capture distinct stages of human embryo development, from the 8C stage to the peri-implantation stage.

RESULTS

Identification of cell heterogeneity in naive hiPSC cultures

To determine the cell heterogeneity of our 5iLAF naive culture we performed scRNA-seq and unsupervised clustering analysis of 3,652 cells, obtaining seven clusters (Figure 1A). First, we validated the naive identity of our cells by integrating our data with previously published results (Liu et al., 2020), where the process of cell reprogramming from fibroblasts to iPSCs in primed and in naive (RSeT and t2i/L + Gö) conditions was analyzed at single-cell resolution. Our sample clustered next to the naive cells of this study and presented high score for naive and epiblast gene signatures (Figure S1). These results were confirmed by morphology and protein expression (Figure S2), Consistently, primed markers were not expressed in any of the seven clusters identified in our culture, except for the CD24 marker, which was enriched in cluster 5 (Figure 1B), altogether demonstrating the lack of residual primed cells. Naive and pluripotency markers associated with the human embryo epiblast were homogeneously expressed in clusters 0–3, except for *PRDM14*, that was lower in cluster 2 (Figures 1C and 1D). On the other hand, some of these markers were downregulated in clusters 4–6 (Figures 1C and 1D). In addition, correlation analysis of differentially expressed genes (DEGs), showed high similarity between clusters 0–3, while clusters 4–6 defined distinct cell populations (Figure 1E). Therefore, we considered clusters 0–3, representing 80.5% of the total cells, as a single population which we named epi-cluster 0–3. A comparison of the DEGs from this cluster (Table S1) with publicly available stage-specific gene expression modules of human embryos from zygote to late blastocyst stage (Stirparo et al., 2018), based on data from three different single-cell datasets (Blakeley et al., 2015; Petropoulos et al., 2016; Yan et al., 2013), showed that the majority of DEGs (53.4%) were specific to either late (39.3%), or early (14.1%) blastocyst's ICM (Figure 1F). In addition, we analyzed the expression of transposable elements (Table S2), as their transcription is regulated in a stage-specific manner during human early embryogenesis (Göke et al., 2015). In our epi-cluster 0–3 we observed high expression of SVA family members, as well as LTR5-Hs and HERVK-int elements (Figure 1G), which are characteristic of morula and early blastocyst stage, in line with previously described naive cell-specific transposable elements (Theunissen et al., 2016).

Presence of TE-like cells and PrE-like cells in naive hiPSC cultures

We found that cells in cluster 5, which represent 4.4% of total cells, displayed the highest scores for a previously described TE signature (Xiang et al., 2020) (Figure 2A). In agreement, known TE-associated genes (Petropoulos et al., 2016; Xiang et al., 2020) were enriched in this cluster (Figure 2B). To conclusively demonstrate the TE commitment of cluster 5 we integrated our data with scRNA-seq data from human preimplantation embryos (Petropoulos et al., 2016): this analysis shows that cluster 5 cells locate within the embryo's TE (Figure 2C). Henceforth, we will refer to cluster 5 as TE-like cluster. RNA velocity analysis indicated that TE-like cells arose from naive cells of epi-cluster 0–3 (Figure 2D). This finding is consistent with the described capacity of human naive cells to differentiate into the TE lineage under defined culture conditions (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2021; Io et al., 2021). The presence of TE-like cells in our naive culture was corroborated by the detection of GATA3$^+$ cells by immunofluorescence staining, while we did not observe GATA3$^+$ positive cells in primed condition (Figure 2E).

Interestingly, within the TE-like cluster, we were able to observe a few scattered cells presenting a high score for the PrE signature of the human embryo (Figure S3A)

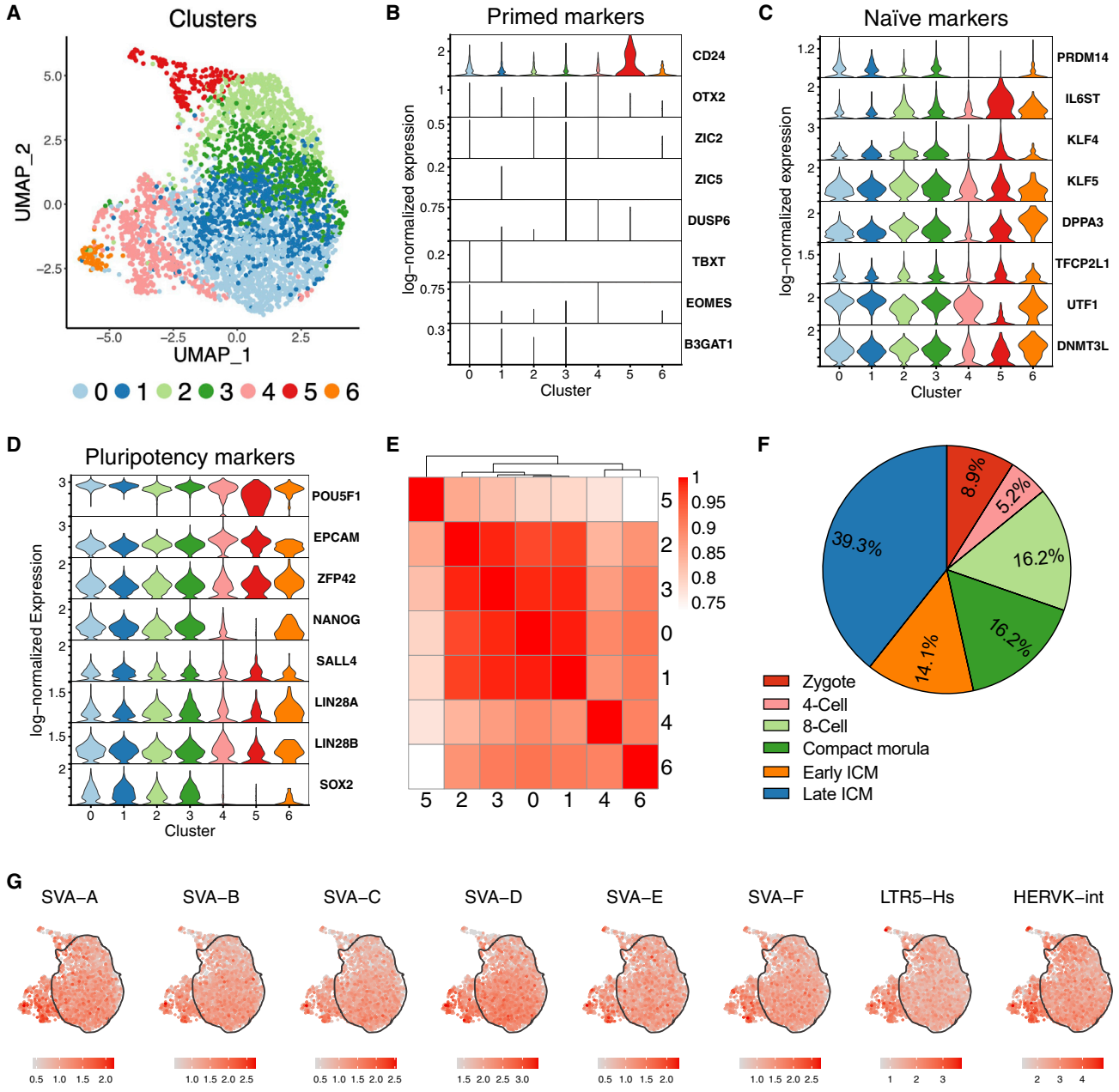


Figure 1. Overall analysis of cell heterogeneity in human naive 5iLAF culture

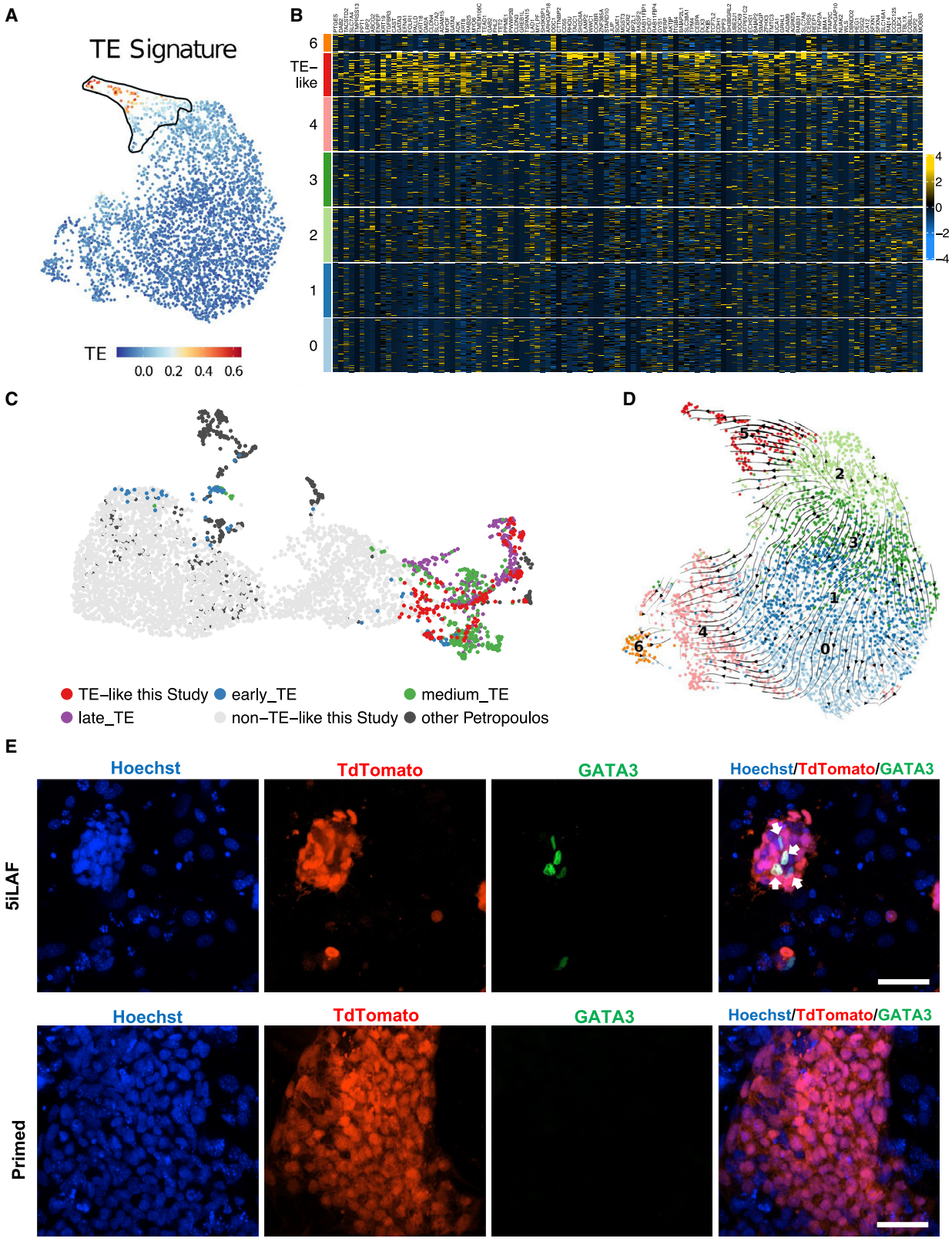
(A) Uniform Manifold Approximation and Projection (UMAP) of the 5iLAF-cultured hiPSCs (3,652 cells). Cells are color coded according to unsupervised clustering analysis.

(B–D) Violin plots showing single-cell log-normalized expression of selected primed, naive, and pluripotency markers in each cluster. Color code as in (A).

(E) Heatmap representing the correlation between the different clusters, considering differentially expressed genes in at least one cluster against the rest.

(F) Proportion of differentially expressed genes in epi-cluster 0–3 corresponding to each stage-specific gene expression module of the human embryo at early developmental stages obtained from (Stirparo et al., 2018) and listed in their Table S6. Genes differentially expressed in epi-cluster 0–3 that were not specific to any developmental stage or were not annotated were removed from the analysis.

(G) UMAP representations of the normalized expression of naive-specific transposable elements. A contour delineates epi-cluster 0–3.



(legend on next page)

described in (Xiang et al., 2020), which cluster next to embryonic PrE cells (Figure S3B). On the other hand, in these cells we did not detect expression of definitive endoderm, visceral endoderm, or parietal endoderm markers (Figures S3D and S3E). These data suggest the presence of PrE-like cells in naive culture. The existence of PrE-like cells in naive 5iLAF cultures is in line with the previously reported expression of GATA6 in human t2iLGÖY naive cultures (Guo et al., 2016, 2017; Linneberg-Agerholm et al., 2019). However, we cannot definitely rule out the possibility that these cells would be definitive endoderm cells, given the fact that in the literature several genes are reported to be expressed in both cell types.

TE-like cells in hiPSC naive cultures mimic the early stages of TE development

Analysis of defined TE-associated markers showed heterogeneous expression within cells of the TE-like cluster (Figures 2A and 2B and 3A). Accordingly, an unsupervised subclustering analysis revealed three subpopulations (5_0, 5_1, and 5_2; Figure 3B). Trajectory analysis using CytoTRACE (Figure 3C) and monocle 3 (Figure 3D) showed that, within the TE-like cluster, the 5_0 subcluster corresponds to less differentiated cells, while the most differentiated subcluster is 5_2. Accordingly, scores for the epiblast signature were lower in cluster 5_0 compared with epi-cluster 0–3, further decreasing in clusters 5_1 and 5_2 (Figure 3E). To determine to which TE developmental stage each subcluster correlates, we scored the cells for the GATA2 and NR2F2 genes modules, recently reported to define different developmental stages of the TE in the human embryo (Meistermann et al., 2021). According to the Meistermann study, early TE cells (B3 stage) are GATA2-module^{+} and NR2F2-module^{-}, while mature TE cells (B4-B6 stage) are GATA2-module^{+} and NR2F2-module^{+} (Meistermann et al., 2021). When applying the GATA2 module to our 5iLAF sample, we found the highest scores throughout the TE-like cluster (Figure 3F). However, cells of cluster 5_0 had relatively lower scores for this module, and mainly had negative scores for the NR2F2 module (Figure 3F), a further proof of their early maturation stage. In addition, this cluster expressed epiblast markers, such as *DPPA3*, *DPPA5*, *DNMT3L*, and *ALPG* at higher levels compared with the other two subclusters (Table S1). Co-expression of epiblast and early TE genes is a hallmark of TE transition in human embryonic development (Niakan and Eggan, 2013; Stirparo et al., 2018), and thus we termed cluster 5_0 as transition TE-like cluster. Cluster 5_1 and cluster 5_2 could be distinguished by differences in scores for the NR2F2 module, being higher in the latter cluster (Figure 3F). Therefore, we named cluster 5_1 as early an TE-like cluster and cluster 5_2 as a mature TE-like cluster. To corroborate these results, we applied signatures for the different steps of the differentiation process of TE cells from pre-cytotrophoblast stage (pre-CTB, corresponding to developmental days 6–7 of the human embryo; D6-7) up to the extravillious trophoblasts stage (EVTs) (D14) derived from 3D-cultured embryos *in vitro* Xiang et al. (2020) (Figures 3G and S4A). Cells with the highest scores for the pre-CTBs (D6-7) signature were found in the transition- and early TE-like clusters (Figure 3G), while cells in the mature TE-like cluster presented the highest scores for the early syncytiotrophoblast (early STB; D8-10) signature (Figure 3G). On the other hand, cells from later stages of placental differentiation (STB, early EVT and EVT) were not present, according to the lower scores obtained for their corresponding signature (Figure S4A) and the absence of placental polypeptide hormone expression (Figure S4B). Importantly, we confirmed these results by integrating our dataset with the embryo data from Xiang and Petropoulos (Petropoulos et al., 2016; Xiang et al., 2020) as shown in Figure 3H. Finally, we did not observe expression of amnion-specific markers (Zhao et al., 2021) (Figures S5A and S5B), and TE-like cells did not cluster with early amnion cells of the embryos (Figure S5C). Moreover, in our 5iLAF cultures we retrieved 38 transposable elements differentially expressed in the TE-like cluster compared with the rest of cells. Between them, we identified repetitive sequences from the LINE1, Alu, and LTR families (Figures 4A and 4B and Table S2). It is noteworthy that 86 transposable

Figure 2. TE-like cells in human naive 5iLAF culture

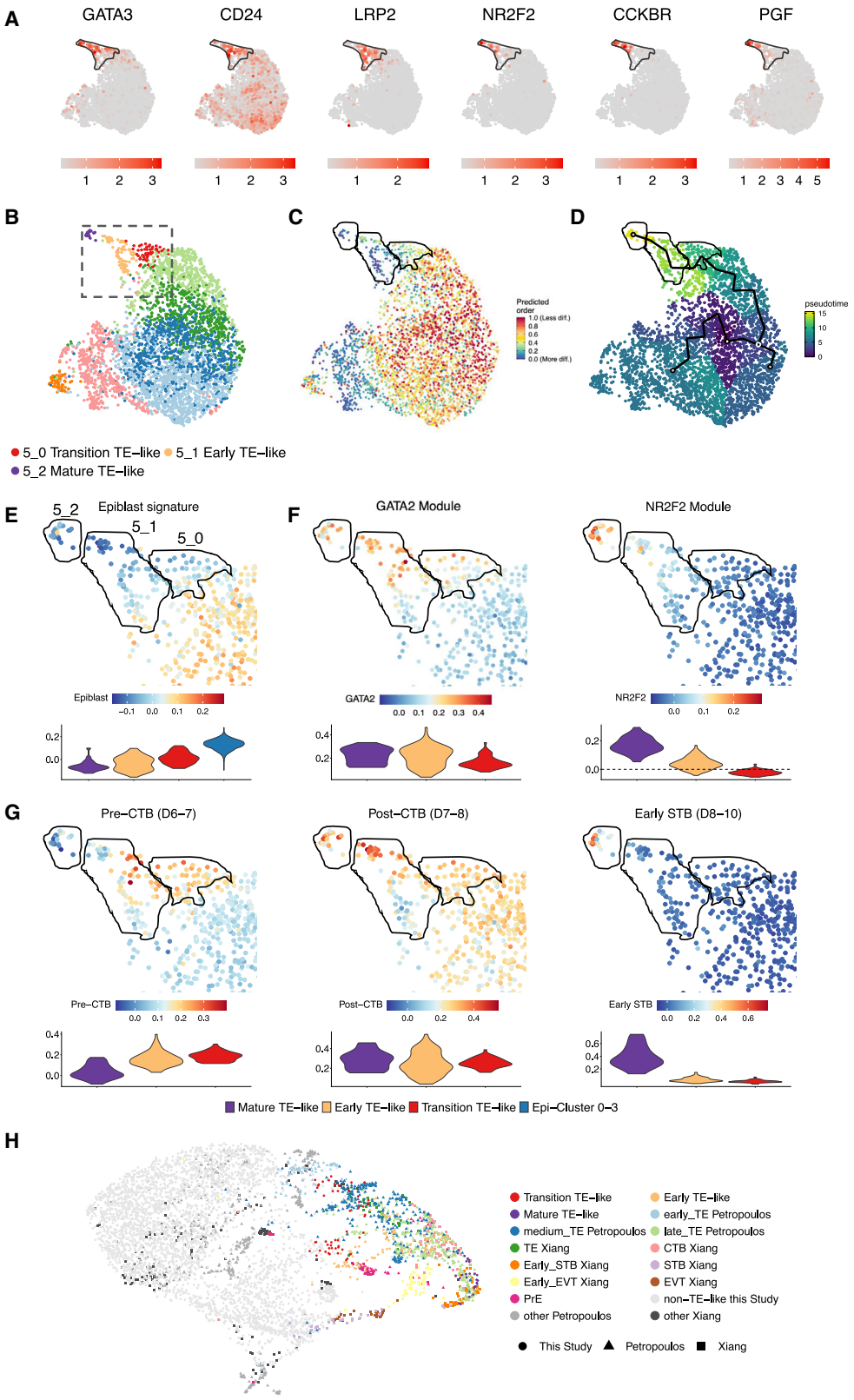
(A) UMAP representation of scores from a TE signature including TE-specific markers that are common between Petropoulos et al. (2016) and Xiang et al. (2020) studies. A contour delineates cluster 5.

(B) Heatmap representing scaled gene expression of the 111 genes from the TE signature. To facilitate visualization of the data, 200 cells from each cluster were randomly selected. Cells are ordered by cluster membership.

(C) UMAP representation of the integration of our sample with the human embryo cells from Petropoulos et al. (2016). TE embryonic cells and TE-like cells are colored, while the rest are shown in different tones of gray. The TE cell-type annotation corresponds to the reannotation in Meistermann et al. (2021).

(D) UMAP plot showing RNA velocities as streamlines.

(E) Immunofluorescence staining of GATA3 (in green) in human naive (upper panels) and human primed (lower panels) cultures. The arrows indicate GATA3^{+} human cells. Human cells are labeled with tdTomato (in red) and Hoechst was used as a nuclear counterstain. Scale bars, 50 μm.



(legend on next page)

elements were differentially expressed in early or mature TE-like subclusters compared with the other clusters, suggesting a specific role of those sequences in the maturation process of TE cells. Between the previously described TE-specific transposable elements, for example, we found that LTR30, LTR10A (Liu et al., 2019), and LTR2B, were enriched in the early TE-like cluster (Figures 4A and 4B). On the other hand, LTR3A (Liu et al., 2019), HERVK3-int, and LTR10B2 were enriched in the mature TE-like cluster (Table S2). Finally, transposable elements known to be involved in placental development (Senft and Macfarlan, 2021), such as endogenous retroviruses ERVW-1 and ERVFRD-1, were enriched in the mature TE-like cluster (Table S1). Thus, the study of transposable elements could help to elucidate the process of TE maturation.

Maturation of TE cells is associated with specific metabolic requirements (Kaneko, 2016; Posfai et al., 2019), so next we analyzed the metabolic transcriptome in the different subclusters. Interestingly, pathways related to energy production were enriched in early TE-like clusters (Figure S4E and Table S1), in agreement with the high ATP consumption of these cells *in vivo*, required to expand the blastocoel cavity (Houghton et al., 2003). Meanwhile, mature TE-like cluster was enriched in pathways related to steroid hormone synthesis/metabolism and regulation of angiogenesis, all functions needed to allow proper implantation of the embryo into the uterine wall (Figure S4E; Table S1). Finally, molecules playing a key role in the embryo attachment to the endometrium (Idelevich and Vilella, 2020) were observed in mature TE-like cells, such as *PGF*, *CGA*, *TGFB1*, and *VEGFA* (Table S1).

In conclusion, TE-like cells present in naive cultures mimic the early stages of TE development, which allows us to study the TE specification and maturation transcriptional program *in vitro*. For instance, the modulation of the transcription factors differentially expressed in our TE-like subclusters (Figures S4C and S4D; Table S1) could elucidate their role in the TE differentiation process.

Identification of 8CLCs in 5iLAF culture

Interestingly, in cluster 6 of our analysis, which corresponds to 1.7% of the total cells, we found cells expressing *ZSCAN4* (Figure 5A), a defining gene for the 8C-stage human embryo (Stirparo et al., 2018). In addition, in this cluster we detected expression of *DUXA*, *PRAMEF1*, *TPRX1*, *LEUTX*, and *KLF17*, and other markers associated with the human 8C-stage embryo (Maeso et al., 2016; Stirparo et al., 2018; Töhönen et al., 2015; Wang et al., 2018) (Figure 5A; Table S1). On the other hand, we observed significantly lower expression levels of stem cell markers, such as *NANOG*, *SOX2*, and *PRDM14*, compared with the epi-cluster 0–3, but not *ZFP42* or *POU5F1* (Figures 1C and 1D). Accordingly, double immunofluorescence staining showed that ZSCAN4^{+} cells presented low or absent signal for SOX2 (Figure 5B). To determine to which stage of human embryo development cluster 6 corresponds to, we integrated our dataset with the scRNA-seq data from human preimplantation embryo (Petropoulos et al., 2016) and found that cluster 6 cells located next to cells of the E3 stage embryo: 8C stage cells undergoing ZGA (Figure 5C). To refine this comparison, we contrasted cluster 6's DEGs to stage-specific gene expression modules (Stirparo et al., 2018) from zygote to blastocyst. From the annotated genes (Stirparo et al., 2018), we observed that ~40% were associated with the 8C stage and ~20% with compact morula (Figure 5D). We next applied to our sample the DUXA module, which consists of genes associated

Figure 3. Analysis of the three different stages of TE maturation present in the TE-like cluster

(A) UMAP representing log-normalized expression of a selected group of TE-associated markers. A contour delineates the TE-like cluster.

(B) UMAP representation after subclustering the TE-like cluster. Cells are color coded according to unsupervised clustering analysis. The dashed box corresponds to the UMAP region shown in (E and G).

(C) UMAP representation of the cellular predicted differentiation ordering from CytoTRACE. A contour delineates each TE-like subcluster.

(D) UMAP-based cellular trajectory reconstruction using monocle 3 (black line) and representation of the inferred pseudotime. A contour delineates each TE-like subcluster.

(E and F) UMAP representation of scores from (E) the epiblast signature (coming from the intersection from Petropoulos et al. (2016) and Xiang et al. (2020), (F) GATA2 gene module and the NR2F2 gene module signatures from Meistermann et al. (2021). A contour delineates each subpopulation from the TE-like cluster. Violin plots of the scores for the different TE-like subpopulations and the epi-cluster 0–3 are depicted underneath, with a dashed line at score zero for the NR2F2 module.

(G) UMAP representation of gene signature scores from different trophoblast lineages obtained from Xiang et al. (2020): pre-CTB (days 6 and 7), post-CTBs (days 7 and 8), and early STBs (days 8–10) after removing genes shared with the epiblast signature. A contour delineates each subpopulation from the TE-like cluster. Violin plots of the scores grouped by TE-like subpopulations are depicted underneath. The color code from the legend is the same for the violin plots in (E and F).

(H) UMAP representation of the integration of our sample with cells of the human preimplantation embryo cells from Petropoulos et al. (2016) and the 3D-cultured embryos from Xiang et al. (2020). TE and primitive endoderm embryonic cells and TE-like cells are colored, while the rest are shown in different tones of gray. The TE cell-type annotation in Petropoulos et al. corresponds to the reannotation in Meistermann et al. (2021) and in Xiang et al. to the reannotation from Rostovskaya et al. (2022). Cells are shaped according to sample.

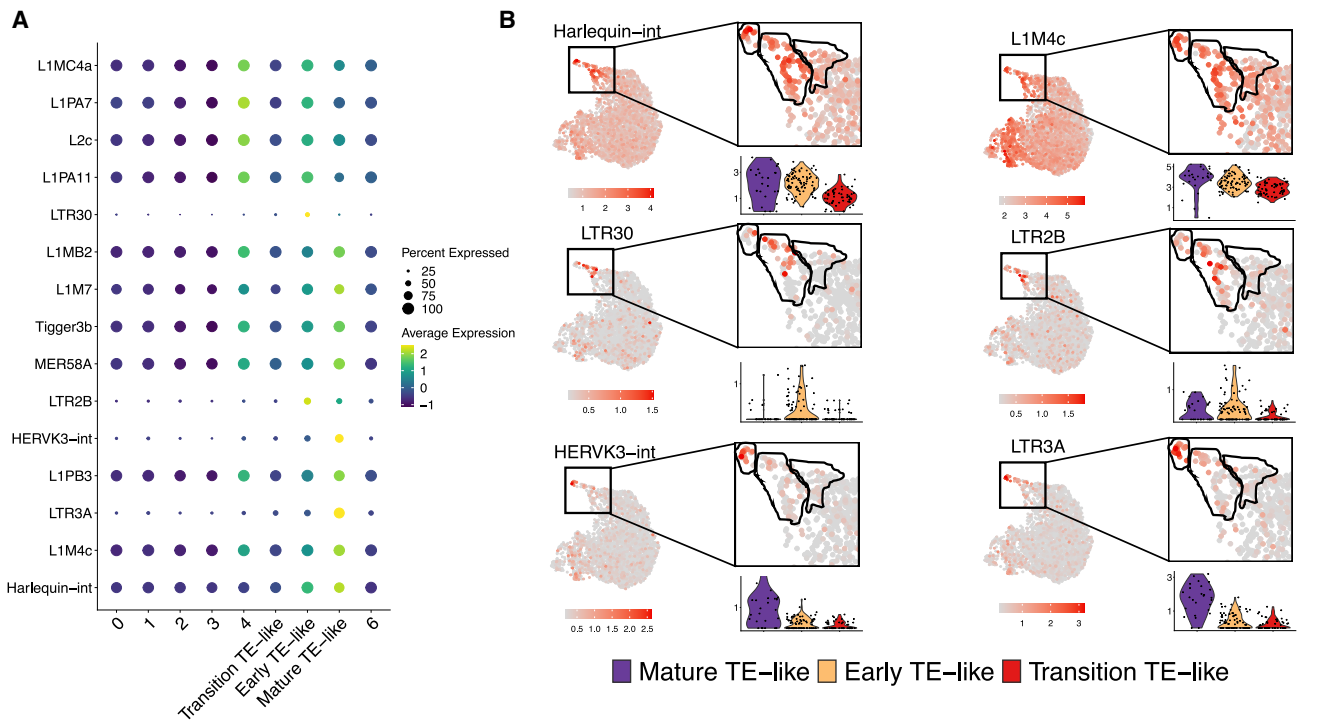


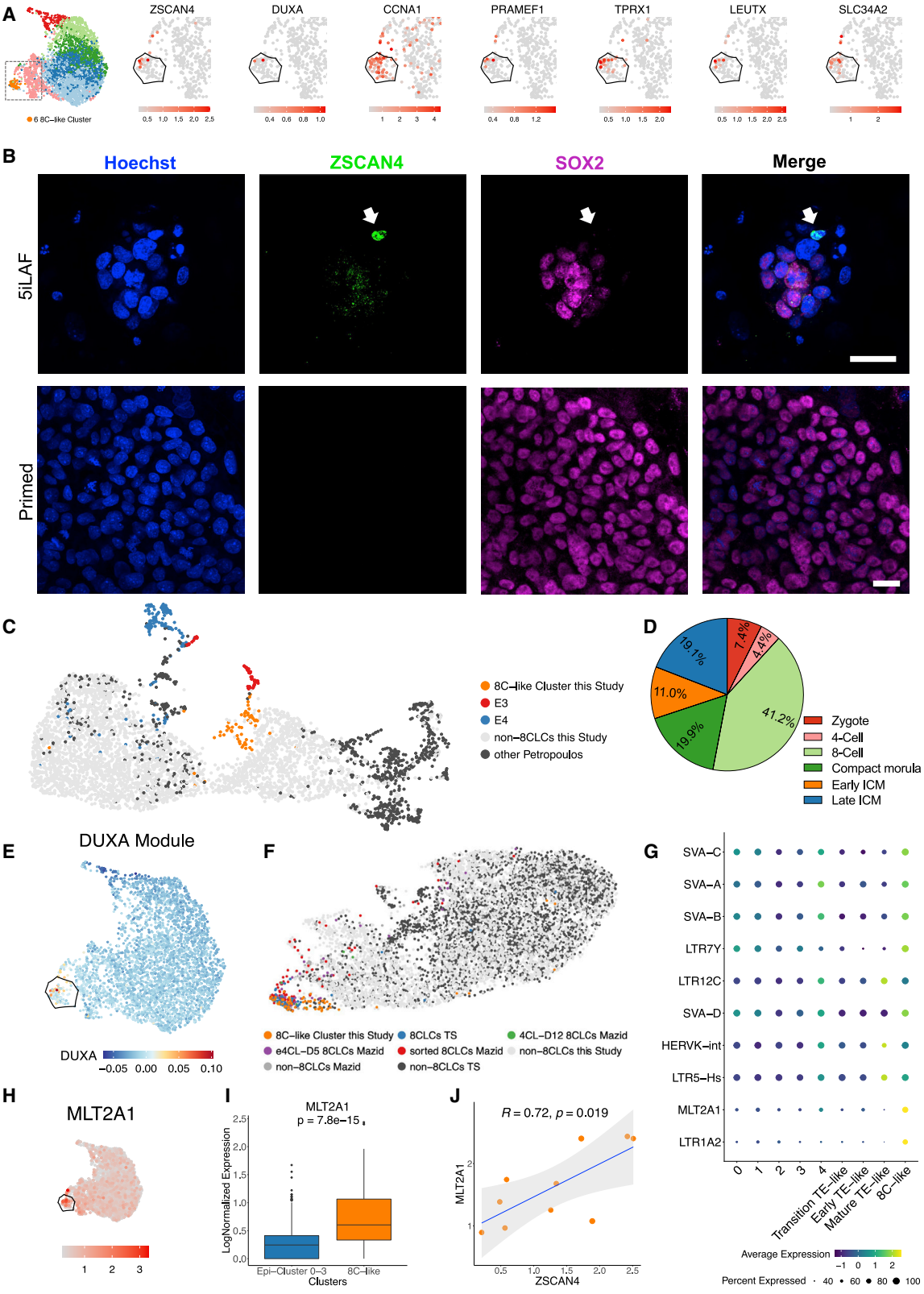
Figure 4. Transposable element analyses in the three subpopulations of the TE-like cluster

(A) Dot plot representing scaled average expression of the top differentially expressed transposable elements in each of the TE-like subclusters. The size of the dot represents the proportion of cells expressing such element in each cluster.

(B) UMAP representing the log-normalized expression in TE-like subclusters of selected transposable elements. A zoom-in of the TE-like cluster is displayed, with a contour delineating each subpopulation. Violin plots grouped by TE-like subpopulations are shown underneath the zoom-ins.

with the ZGA process (Meistermann et al., 2021), finding the highest scores in cluster 6 (Figure 5E). Moreover, as DUX4 has been identified as a key regulator of the ZGA genetic signature both *in vivo* and *in vitro* (de Iaco et al., 2017), we looked for expression of DUX4-induced genes (Hendrickson et al., 2017) in our dataset, and accordingly with our previous analysis we observed that they were enriched in cluster 6 (Figure S6A). These data consistently show that the cluster 6 gene expression profile closely resembles that of the 8C-stage human embryo, therefore we termed it an 8C-like cluster. Recently, the studies from Mazid et al. (2022) and Taubenschmid-Stowers et al. (2022) have also identified 8CLCs in human naive cultures by scRNA-seq. Integration of these datasets with ours shows that our 8C-like cluster cells overlap with the 8CLCs from Mazid et al. and Taubenschmid et al. (Figure 5F), demonstrating their overall similarity. However, e4CL 8CLCs express 8C-associated markers at a higher level (Figure S6B), confirming the improvement of the 8CLC phenotype in e4CL medium (Mazid et al., 2022). On the other hand, the expression profile of pluripotency and naive markers was similar in all conditions (Figures S6C and S6D).

As the expression of MERVL retrotransposons is a hallmark of 2CLCs in mouse (Macfarlan et al., 2012), we next analyzed transposable element expression in the 8C-like cluster. In this cluster we observed a significantly higher proportion of reads corresponding to retroposon, LTR, SINE, and LINE families compared with epi-cluster 0–3 (Figure S6E). In addition, transposable elements known to be expressed between 8C and morula stages in human embryos, such as the SVA retroposons, LTR5-Hs, and LTR7Y (Göke et al., 2015; Liu et al., 2019), were enriched in the 8C-like cluster, as well as 8C-specific transposable elements, such as HERVK-int, LTR12C, and MLT2A1 (Göke et al., 2015; Grow et al., 2015; Liu et al., 2019) (Figure 5G). It is noteworthy that MLT2A1, the most specific repetitive sequence to the 8C human embryo *in vivo* (Göke et al., 2015), was among the most enriched in the 8C-like cluster (Figures 5G–5I; Table S2), in agreement with the expression pattern observed in the 8CLCs from Mazid et al. (2022) and Taubenschmid-Stowers et al. (2022). Importantly, *ZSCAN4* expression correlated with MLT2A1 expression (Figure 5J). Thus, we defined the *ZSCAN4*$^+$/MLT2A1$^+$cells found in the 8C-like cluster as bona fide 8CLCs. In mPSC cultures, *Zscan4*$^+$/MERVL$^+$ 2CLCs range from 0.1% to 0.5% of the total cells (Macfarlan et al.,



(legend on next page)

2012). Accordingly, in our naive 5iLAF sample, we observed a similar proportion of 8CLCs (~0.3%). In conclusion, these data prove the existence of 8CLCs in 5iLAF cultures.

Presence of a cell population characterized by cell-cycle arrest

However, cluster 4, representing 13.4% of the total cells, expressed several epiblast markers, could be distinguished by very low expression of *NANOG*, *SOX2*, *PRDM14*, and *KLF4* (Figures 1C and 1D). The most prominent characteristic found in this cluster was the high percentage of cells in cell-cycle arrest (G1 phase, Figure 6A). In fact, the most significantly enriched gene sets for this cluster are related to response to DNA damage response and cell-cycle arrest (Figure 6B).

When studying the differential expression of transposable elements in cluster 4, we found enriched expression of LINE1 elements (L1) (Figures 6C and 6D). LINE1 are the only family of autonomously active, protein coding retrotransposons in humans (Wallace et al., 2008). L1 elements are retrotransposed during early embryogenesis, contributing to genetic variations among individuals, while they are normally silenced in differentiated tissues, with the exception of the central neural system (Upton et al., 2015). Indeed, uncontrolled LINE1 retrotransposition has been shown to play a role in the initiation and progression of cancers, affecting cancer cells' genomic stability and inducing cell stress via the activation of DNA damage response, as well as cell growth arrest through P53 activation (Wallace et al., 2008). Similarly, the transcriptional profile of cluster 4 shows an enrichment in genes related to DNA damage response (Figure 6B; Table S1) and is consistent with high LINE1 activation. Because detection of L1 RNAs alone is not indicative of active LINE1, as the majority of these transcripts present truncations or mutations, which prevent their translation, we examined the accumulation of L1 protein ORF1p (Sharma et al., 2016), required for retrotransposition. In our 5iLAF cultures we found cytoplasmic foci of ORF1p (Figure 6F), a sign of accumulated LINE1 ribonucleoprotein (Doucet et al., 2010). Because the reduction of the MEK inhibitor (MEKi) PD0325901 (PD) is known to decrease the detrimental effects on genetic and epigenetic stability induced by the naive medium (di Stefano et al., 2018), we next performed ORF1p immunofluorescence staining in naive 5iLAF cultures in reduced PD concentration (0.5 μM). Notably, this assay showed a reduced accumulation of ORF1p (Figures 6F and 6G), thus suggesting a possible link between the genomic instability induced by the MEKi of the naive medium and the expression of retrotransposable LINE1 elements.

Heterogeneous contribution of hiPSCs to human-mouse chimera

As chimera formation is the gold standard assay to determine the developmental potential of stem cells, to functionally prove the existence of cells corresponding to distinct developmental stages in our cultures we microinjected 8–10 5iLAF cells, labeled with tdTomato, into morula-stage mouse embryos and cultured them for 48 h *in vitro* (Figures 7A and 7B). Quantification at 24 h showed that part of the microinjected cells did not engraft in the host embryo, as we observed a decrease in cell number at this time point, whereas we detected cell proliferation at 48 h, as indicated by an increased number of cells (Figure 7C). Strikingly, at this time point, in addition to finding human cells in the ICM, we also observed a fraction of cells contributing to the TE ($GATA3^+$ cells: 11/244 = 4.5%) (Figures 7D and 7E). This percentage of TE contribution is in line with the percentage

Figure 5. 8CLCs in human naive 5iLAF culture

(A) Zoom-ins of the UMAP region including the 8C-like cluster representing the log-normalized expression of selected 8C embryo-associated markers. The zoom-ins correspond to the dashed region shown in the top left UMAP. A contour delineates the 8C-like cluster.

(B) Confocal microscopy images of immunofluorescence staining for ZSCAN4 (in green) and SOX2 (in red) in 5iLAF (on the top panels) and primed (on the lower panels) hiPSCs. The arrow indicates $ZSCAN4^+$ human cell. DNA was counterstained with Hoechst. Scale bars, 20 μm.

(C) UMAP representation of the integration with human embryo cells from Petropoulos et al. (2016). Cells from the earliest developmental stages (E3 and E4) and the 8C-like cluster are colored, while the rest are shown in gray.

(D) Proportion of differentially expressed genes in the 8C-like cluster corresponding to each stage-specific gene expression module of the human embryo at early developmental stages obtained from Stirparo et al. (2018) and listed in their Table S6. Genes differentially expressed in the 8C-like cluster that were not specific to any developmental stage or were not annotated were removed from the analysis.

(E) UMAP representation of scores from the DUXA gene module from Meistermann et al. (2021). A contour delineates the 8C-like cluster.

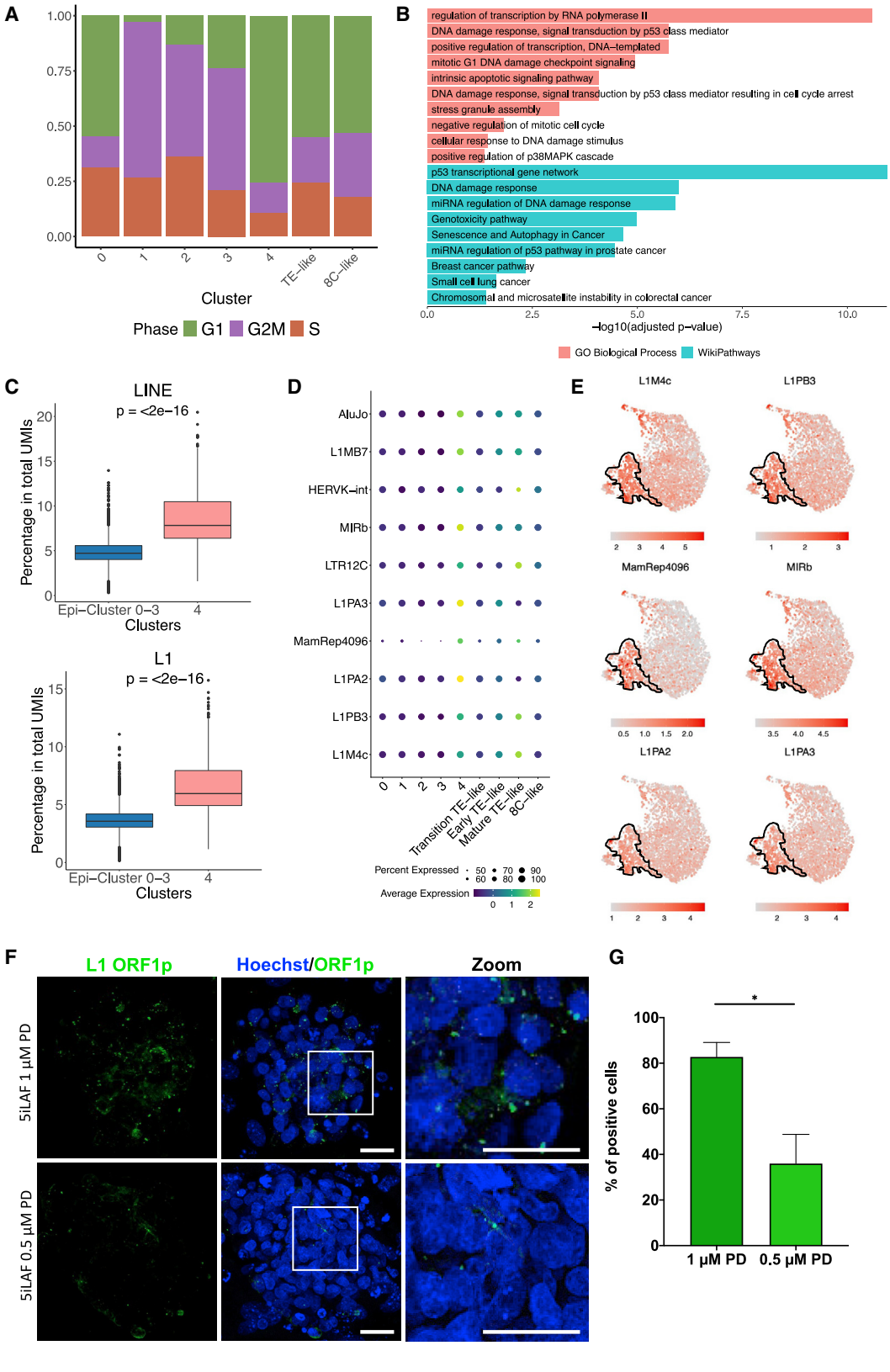
(F) UMAP of the integration of our 5iLAF-cultured sample and samples from Mazid et al. (2022) and Taubenschmid-Stowers et al. (2022). 8CLCs from each study are colored, with non-8CLCs in tones of gray. For Mazid et al., 8CLCs are separated according to the sample of origin.

(G) Dot plot representing scaled average expression of the top 10 differentially expressed transposable elements in the 8C-like cluster. The size of the dot represents the proportion of cells expressing such element in each cluster.

(H) UMAP representation of log-normalized expression of the transposable element MLT2A1. A contour delineates the 8C-like cluster.

(I) Boxplots showing the percentage of total unique molecular identifiers (UMIs) corresponding to the MLT2A1 LTR element in cells from epi-cluster 0–3 and the 8C-like cluster. Two-sided Wilcoxon rank sum tests were performed to obtain statistical significance ($p < 0.05$).

(J) Correlation analysis between *ZSCAN4* and MLT2A1 expression in $ZSCAN4^+$ cells from the 8C-like cluster ($p < 0.05$).



(legend on next page)

of TE-like cells in our culture, suggesting their possible contribution to this lineage *in vivo*.

Thus, we performed new experiments microinjecting a hiPSC population enriched in TE-like cells (FACS sorted $CD24^{high}$ cells). Microinjection of sorted $CD24^{high}$ cells show a significant enrichment in TE contribution: 56.5% (13 $GATA3^{+}$ cells out of 23 engrafted cells Figures 7F and 7G), a frequency 12.5 times higher than bulk naive culture. Moreover, we detected human contribution to TE in 100% of the embryos (n = 8), while after bulk cell microinjection it was 36.4% (n = 22). These data demonstrate the presence of heterogeneity in naive hiPSC cultures and, specifically demonstrate the existence of functional TE-like cells.

DISCUSSION

With this work we reveal and thoroughly describe four main cell populations with distinct characteristics coexisting in human naive 5iLAF cultures. While the expression profile of 80% of the cells, here named epi-cluster 0–3, defined them as the *in vitro* counterpart of the late blastocyst epiblast cells, which corresponds to the gold standard definition of naive, we also revealed the presence of TE-like cells and 8CLCs, as well as a population of epiblast-like cells in cell-cycle arrest.

The TE-like cluster is characterized by a gene expression profile comparable with the early TE of the human embryo (D6–D9) and expresses TE-specific transposable elements at high levels. This population represents 4.4% of the cells. Such a population of cells was not previously found in naive PXGL conditions, possibly as a consequence of the presence of WNT and aPKC inhibitors, which block access to TE differentiation in this medium (Guo et al., 2021). Trajectory analysis showed that cells of the TE-like cluster arise from naive cells. This is in line with the known potential of naive hPSCs to give rise to trophoblast stem cells (hTSCs) in a defined medium; hTSCs being the *in vitro* counterpart of the trophoblasts at D8–D12 (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2021; Io et al., 2021). Interestingly, TE-like cells found in our naive cultures presented a heterogeneous gene and transposable element expression profile, which allowed us to subdivide them into three subclusters corresponding to three early extraembryonic tissue maturation stages: transitional, early, and mature TE. Thus, while hTSCs can be used to study trophoblast lineage differentiation (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2021; Io et al., 2021), TE-like cells in 5iLAF cultures, which resemble the TE of the blastocyst at days D6–D9, may be a suitable model to shape the role of the different genetic determinants of TE formation and its early maturation.

TE cells contribute to the process of implantation by modulating endometrial receptivity (Posfai et al., 2019). In the TE-like cluster, we observed factors previously described to take part in this process and we unraveled new candidates that may play an important role in trophoblast to endometrium communication as, for instance, *SEMA4C* (Table S1), whose receptor, PLEXIN B2, has been reported to participate in endometrial integrity (Singh and Aplin, 2015). Since trophoblast dysfunction leads to complications during pregnancy, such as pre-eclampsia and intrauterine growth restriction, the study of TE-like cells present in patient-derived naive iPSC cultures would facilitate the identification of the molecular players inducing these complications.

It is noteworthy that the 8C-like cluster defines a population of cells enriched with the expression of 8C embryo-specific genes and transposable elements, while showing downregulation of pluripotency-associated genes, such as *PRDM14*, *SOX2*, and *NANOG*. This discovery is in line with the recent description of an analogous 8CLC population in hESC-cultured naive PXGL medium (Taubenschmid-Stowers et al., 2022) and conversion to 8CLCs in e4CL medium (Mazid et al., 2022). In the 8C-like cluster, which consists of 1.7% of the naive culture, we detected the existence of $ZSCAN4^{+}$/$MLT2A1^{+}$ cells, which present a gene and transposable element expression strictly characteristic of the 8C stage of the human embryo *in vivo*. These

Figure 6. Presence of naive cells with arrested cell cycle in cluster 4

(A) Proportion of cells in each cell-cycle phase for the unsupervised clusters.

(B) Selection of significant GO Biological Processes and WikiPathways terms from an overrepresentation analysis using DEGs in cluster 4 against the rest.

(C) Boxplots showing the percentage of total UMIs corresponding to LINEs and more specifically to L1 elements in cells from epi-cluster 0–3 and cluster 4. Two-sided Wilcoxon rank sum tests were performed to obtain statistical significance ($p < 0.05$).

(D) Dot plot representing scaled average expression of the top 10 differentially expressed transposable elements in cluster 4. The size of the dot represents the proportion of cells expressing such element in each cluster.

(E) UMAP representation of log-normalized expression of differentially expressed transposable elements in cluster 4, delineated by a contour.

(F) Confocal microscopy images of immunofluorescence staining for L1 ORF1 protein in 5iLAF 1 μM MEKi PD (above panels) and 0.5 μM MEKi PD (below panels) hiPSCs. DNA was counterstained with Hoechst. Scale bars, 20 μm.

(G) Quantification of ORF1p cytoplasmic aggregates in 5iLAF cells in both MEKi conditions. *p<0.05.

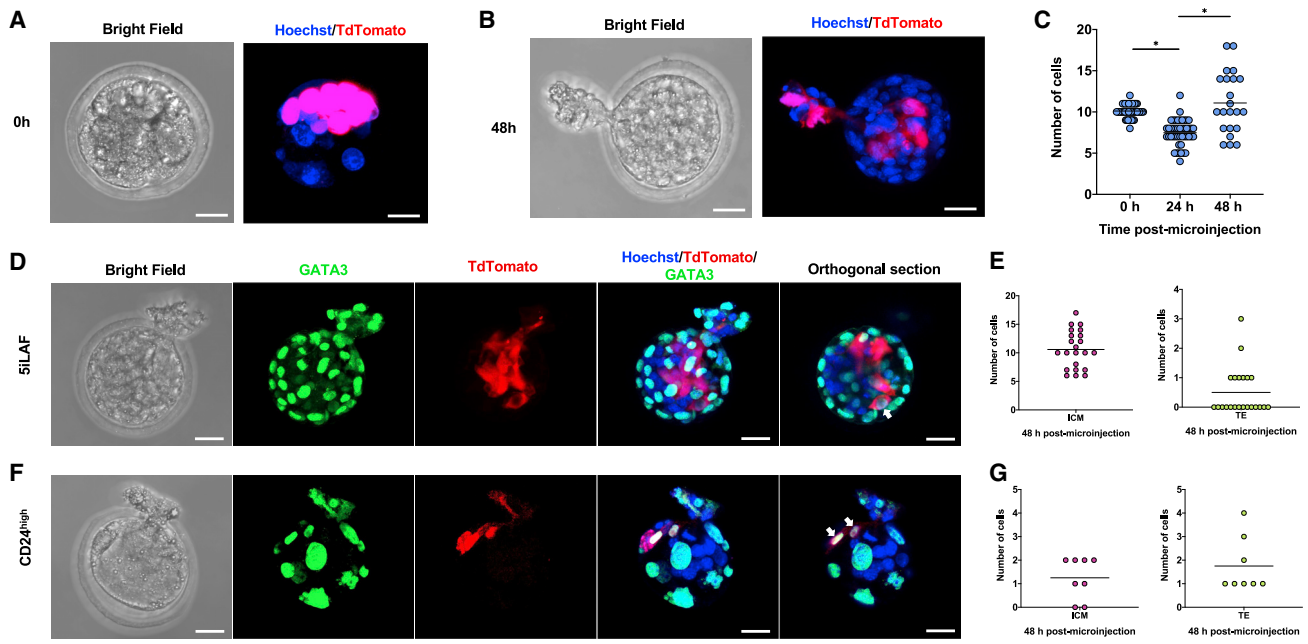


Figure 7. Contribution of human 5iLAF and 5iLAF CD24^{high} cells to the TE lineage in human-mouse chimera embryos

(A) Bright-field and orthogonal projection of a mouse morula freshly microinjected with 10 human 5iLAF tdTomato^{+} cells.

(B) Engraftment of human cells within the mouse embryo after 48 h of *in vitro* culture, shown by orthogonal projection confocal image.

(C) Graphical representation of the human cells present into mouse embryos over time. *$p < 0.05$.

(D–G) (D and F) Immunostaining of GATA3 (in green) in a mouse embryo microinjected with bulk human 5iLAF cells (D) or sorted C24^{high} (F) and cultured *in vitro* for 48 h. The arrows indicate tdTomato^{+} human cells expressing GATA3 in (D and F). (E–G) Graphical representation of human cells contribution to ICM or TE 48 h post-microinjection. Each dot represents a microinjected embryo (n = 22 and n = 8, respectively). Hoechst was used as a nuclear counterstain in (A, B, D, and F). Scale bars, 20 μm.

8CLCs represent ~0.3% of the total cells, consistent with the described proportion of 2CLCs in mouse cultures (Macfarlan et al., 2012). Cells of the 8C-like cluster can be an *in vitro* resource to study the genetic program of the earliest stages of embryo development.

It is well known that human naive PSCs are prone to acquire mutations, chromosomal abnormalities, and epigenetic aberrations (Pastor et al., 2016; Theunissen et al., 2014) and therefore is still necessary to revise the naive culture conditions to improve cell fitness. Accordingly, in our study we have observed that cells of cluster 4, which represent 13.4% of the culture, are in cell-cycle arrest and express genes related to DNA damage response. This observation reinforces the notion that renewed refinement of 5iLAF medium is required to improve cell stability. Stress response-related genes enriched in cluster 4, such as *DDIT3* and *CDKN1A* (Ock et al., 2020) (Table S1), could be used as markers to monitor cell integrity upon modification of the culture conditions. On the other hand, it is noteworthy that, among retrotransposable elements, LINE1 are specifically enriched in this cluster. They have been shown to play an important role on the genomic instability of cancer cells (Grundy et al., 2021), and to induce mechanisms of DNA repair and cell-cycle arrest, which similarly characterize cells in cluster 4. Notably, we have observed accumulation of L1 ORF1p in our cultures, a sign of active LINE1 elements, which is reduced in the presence of titrated concentrations of PD (0.5 μM) in the medium. These observations suggest a possible role for LINE1 elements in the genomic instability related to a strong MEKi in naive cultures. However, further studies will be required to prove their implication in the detrimental effect of MEKi in hPSC cultures, and to elucidate their mechanism of action.

In summary, our study describes the heterogeneity present in 5iLAF naive hiPSCs. Distinct populations corresponding to 8CLCs and TE have been found to coexist with epiblast-like (naive) cells in these cultures. Hence, 5iLAF conditions could be used as a model to investigate human embryo development from the 8C stage, when the zygotic genome activation event occurs, to the peri-implantation period, where the correct maturation of TE cells play a key role, as well as to broaden the

understanding of stem cell biology, thus representing an alternative to the use of human embryos and the associated ethical issues.

EXPERIMENTAL PROCEDURES

Resource availability

Corresponding author

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Xabier L. Aranguren (xlaranguren@unav.es).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The data generated in this study have been deposited in GEO repository: GSE197889.

Animal models

All animals were housed in specific pathogen-free conditions with free access to food and water. All animal procedures were performed according to state and institutional laws, guidelines, and regulations. All studies were approved by the Ethics Committee for Animal Research at the University of Navarra and the Government of Navarra.

hiPSC culture

The G15.AO hiPSC line was used in this study (Zapata-Linares et al., 2016). Cells routinely cultured in primed conditions were converted into naive-like cells in 5iLAF medium as described (Theunissen et al., 2014) with minor modifications (see supplemental experimental procedures). After three passages in 5iLAF medium, the transcriptome of sorted tdTomato + naive hiPSCs was examined using the NEXTGEM Single Cell 3′ Reagent Kits v.3 (10x Genomics) according to the manufacturer's instructions. The median numbers of unique molecular identifiers and genes detected per cell were 20,302 and 4,537, respectively. Methods describing sample preparation for scRNA-seq and computational analyses of RNA-seq data are included in the supplemental information available online.

Cells immunofluorescence staining

Primed and naive hiPSCs were cultured over iMEFs in 8-well chambers (Thermo Scientific Nunc, 177445) until they reached 60%–70% of confluence. Then, they were fixed with 4% paraformaldehyde solution in PBS for 30 min at 4°C, washed 3 times in PBS, permeabilized with PBS 0.1% Triton X-100 (Fisher BioReagents, BP151-500) for 1 h at 4°C, washed again and then blocked in PBS 1% BSA (Sigma-Aldrich, A3912-100G) for 1 h at 4°C. Samples were stained with the following primary antibodies in blocking solution overnight at 4°C: Anti-POU5F1 (1:50, Santa Cruz Biotechnology, sc-9081), Anti-SOX2 (1:50, Sigma-Aldrich, AB5603), Anti-TFE3 (1:200, Merck – Sigma-Aldrich, HPA023881-100UL), Anti-hTFCP2L1 (1:200, R&D Systems, AF5726), Anti-GATA3 (1:100, Abcam, ab199428), Anti-NANOG (1:50, ab80392, Abcam), Anti-ZSCAN4 (1:100, Origene, TA800535), and Anti-LINE-1 ORF1p Antibody (1:100, Merck – Sigma-Aldrich, MABC1152). After rinsing three times with PBS to remove the unbound antibodies, samples were exposed to the corresponding Alexa Fluor 488 goat anti-rabbit IgG secondary antibody (Invitrogen, A11008), Alexa Fluor 488 donkey anti-goat IgG secondary antibody (Life Technologies, A11055), Alexa Fluor 488 goat anti-mouse IgG secondary antibody (Invitrogen, A11029), or Alexa Fluor 647 goat anti-rabbit IgG secondary antibody (Invitrogen, A21245), for 2 h at 4°C in darkness (dilution 1:500 in blocking solution). After three washing steps with PBS, nuclear counterstain was performed with Hoechst 33342 trihydrochloride (Invitrogen, H21492) at 1:800 dilution for 30 min at 4°C in darkness. Images were recorded on a Confocal Scanning Laser Microscope (Zeiss LSM 800) and ZEN 2.3 system software. Composite images were obtained with ZEN 2.3 system software. LINE-1 ORF1p^{+} hiPSCs were quantified manually. Mann-Whitney test was used for statistical analysis. p value < 0.05 was considered significant.

Flow cytometry and FACS sorting

Pluripotency-, primed-, and naive-associated markers were analyzed by FACS. In brief, hiPSCs were cultured up to 70% confluence and washed once with PBS before single-cell dissociation. Then, hiPSCs were dissociated by incubation with TrypLE-EDTA (1:1 TrypLE-0.25 mM EDTA/PBS) for 5 min. HiPSCs were resuspended in 100 μL of 1% BSA-PBS and incubated with the corresponding antibody for 30 min at 4°C in the dark: BV510-conjugated Mouse Anti-human CD24 at 1:100 dilution (BD Horizon/BD Biosciences, 563035), BV421-conjugated Mouse Anti-human CD57 at 1:100 dilution (BD Horizon/BD Biosciences, 563896), PE-conjugated anti-human CD90 at 1:100 dilution (BD Biosciences, 555596), APC-conjugated Mouse Anti-human SSEA4 at 1:100 dilution (R&D Systems, FAB1435A), BV510-conjugated Mouse Anti-human CD77 at 1:100 dilution (BD Horizon/BD Biosciences, 563630), APC-conjugated Mouse Anti-human CD130 at 1:40 dilution (BioLegend, 362006). Isotype-matched IgGs were used as negative controls. Thereafter, samples were washed twice with PBS and finally resuspended in FACS buffer (PBS 1 mM EDTA, 25 mM HEPES, 1% BSA [pH 7]), to be recorded in a BD FACSCanto II with BD FACSDiva software. Flow cytometry data were analyzed using FlowJo software.

Mouse morula microinjection with 5iLAF hiPSCs

B6.DBA2 mouse morulae were microinjected with ten unselected or CD24^{high} sorted tdTomato^{+} 5iLAF hiPSCs using a Leica DMI3000 B microscope and mechanical micromanipulators with a hanging joystick; TransferMan NK2 (Eppendorf). After microinjection, embryos were cultured in 5iLAF medium for 48 h at 37°C in a humidified atmosphere at 5% O_2, 5% CO_2 in air (Drawer Type Incubator, AD-3100), until they developed to the blastocyst stage. Then, tdTomato^{+} hiPSCs inside morulas and blastocysts were photographed and quantified by eye under a fluorescence microscope. Kruskal-Wallis test was used for statistical analysis. $p < 0.05$ was considered significant.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stemcr.2022.11.015.

AUTHOR CONTRIBUTIONS

M.M.-J. generated and characterized naive hiPSCs, performed single-cell experiment, performed mouse-human chimera experiments, analyzed data, and wrote the manuscript. A.U.-A. performed bioinformatic analysis of the data and wrote the manuscript. P.B. provided assistance with experiments and edited the manuscript. A.V.-Z., G.A., C.B., and L.G. provided assistance with experiments. J.P.R.-M. provided assistance with bioinformatic analysis. J.R.R.-M., X.C.-V., and X.A. provided materials and technical advice and edited the manuscript. G.C. discussed the data and wrote the manuscript. F.P. and X.L.A. conceived the research, provided funding, supervised the experiments, analyzed the data, and wrote the manuscript.

ACKNOWLEDGMENTS

This study was supported by PID2021-122589OB-100 financed by MCIN/AEI/10.13039/501100011033/FEDER, UE, the Instituto de Salud Carlos III co-financed by European Regional Development Fund-FEDER "A way to make Europe" Red de Terapia Celular TERCEL (RD16/0011/0005) and Centro de Investigación Biomédica en Red de Cáncer CIBERONC (CB16/12/00489), co-financed by the European Union – NextGenerationEU. Redes de Investigación Cooperativa Orientada a Resultados en Salud RICORS (RD21/0017/0009 and RD21/0017/0019), the Ramón y Cajal State Program (MINECO/FSE, RYC-2015-17233), the Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/FEDER (UE/RTI2018-064485-B-I00), MINECO/FSE (IJCI-2017-33070), and Gobierno de Navarra I + D 2020 co-funded by FEDER funds (0011-1365-2020-000293, 0011-1365-2020-000287, and Bioheart 0011-1411-2022-000092 strategic 2022–2025).

CONFLICT OF INTERESTS

J.P.R. is an employee and shareholder of 10x Genomics.

Received: March 18, 2022
Revised: November 15, 2022
Accepted: November 18, 2022
Published: December 22, 2022

REFERENCES

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 (Cambridge, England) *142*, 3151–3165. https://doi.org/10.1242/DEV.123547.

Castel, G., Meistermann, D., Bretin, B., Firmin, J., Blin, J., Loubersac, S., Bruneau, A., Chevolleau, S., Kilens, S., Chariau, C., et al. (2020). Induction of human trophoblast stem cells from somatic cells and pluripotent stem cells. Cell Rep. *33*, 108419. https://doi.org/10.1016/J.CELREP.2020.108419.

Cinkornpumin, J.K., Kwon, S.Y., Guo, Y., Hossain, I., Sirois, J., Russett, C.S., Tseng, H.W., Okae, H., Arima, T., Duchaine, T.F., et al. (2020). Naive human embryonic stem cells can give rise to cells with a trophoblast-like transcriptome and methylome. Stem Cell Rep. *15*, 198–213. https://doi.org/10.1016/J.STEMCR.2020.06.003.

de Iaco, A., Planet, E., Coluccio, A., Verp, S., Duc, J., and Trono, D. (2017). DUX-family transcription factors regulate zygotic genome activation in placental mammals. Nat. Genet. *49*, 941–945. https://doi.org/10.1038/ng.3858.

di Stefano, B., Ueda, M., Sabri, S., Brumbaugh, J., Huebner, A.J., Sahakyan, A., Clement, K., Clowers, K.J., Erickson, A.R., Shioda, K., et al. (2018). Reduced MEK inhibition preserves genomic stability in naive human embryonic stem cells. Nat. Methods *15* (9), 732–740. https://doi.org/10.1038/s41592-018-0104-1.

Dong, C., Beltcheva, M., Gontarz, P., Zhang, B., Popli, P., Fischer, L.A., Khan, S.A., Park, K.M., Yoon, E.J., Xing, X., et al. (2020). Derivation of trophoblast stem cells from naïve human pluripotent stem cells. Elife *9*, e52504. https://doi.org/10.7554/ELIFE.52504.

Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnaoui, M., Bucheton, A., Moran, J.v., and Gilbert, N. (2010). Characterization of LINE-1 ribonucleoprotein particles. PLoS Genet. *6*, 10011500–e1001219. https://doi.org/10.1371/JOURNAL.PGEN.1001150.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., and Zviran, A. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282–286. https://doi.org/10.1038/NATURE12745.

Göke, J., Lu, X., Chan, Y.S., Ng, H.H., Ly, L.H., Sachs, F., and Szczerbinska, I. (2015). Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. Cell Stem Cell *16*, 135–141. https://doi.org/10.1016/J.STEM.2015.01.005.

Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y., et al. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature *522*, 221–225. https://doi.org/10.1038/NATURE14308.

Grundy, E.E., Diab, N., and Chiappinelli, K.B. (2022). Transposable element regulation and expression in cancer. FEBS J. *289*, 1160–1179. https://doi.org/10.1111/FEBS.15722.

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 *28*, 1040–1056.e6. https://doi.org/10.1016/J.STEM.2021.02.025.

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. https://doi.org/10.1242/DEV.146811.

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 Rep. *6*, 437–446. https://doi.org/10.1016/J.STEMCR.2016.02.005.

Hendrickson, P.G., Doráis, J.A., Grow, E.J., Whiddon, J.L., Lim, J.W., Wike, C.L., Weaver, B.D., Pflueger, C., Emery, B.R., Wilcox, A.L., et al. (2017). Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. Nat. Genet. *49*, 925–934. https://doi.org/10.1038/NG.3844.

Houghton, F.D., Humpherson, P.G., Hawkhead, J.A., Hall, C.J., and Leese, H.J. (2003). Na+, K+, ATPase activity in the human and bovine preimplantation embryo. Dev. Biol. *263*, 360–366. https://doi.org/10.1016/J.YDBIO.2003.07.014.

Idelevich, A., and Vilella, F. (2020). Mother and embryo cross-communication. Genes *11*, 376. https://doi.org/10.3390/GENES11040376.

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 *28*, 1023–1039.e13. https://doi.org/10.1016/J.STEM.2021.03.013.

Kaneko, K.J. (2016). Metabolism of preimplantation embryo development: a bystander or an active participant? Curr. Top. Dev. Biol. *120*, 259–310. https://doi.org/10.1016/BS.CTDB.2016.04.010.

Linneberg-Agerholm, M., Wong, Y.F., Herrera, J.A.R., Monteiro, R.S., Anderson, K.G.V., and Brickman, J.M. (2019). Naïve human pluripotent stem cells respond to wnt, nodal and lif signalling to produce expandable naïve extra-embryonic endoderm. Development *146*. https://doi.org/10.1242/DEV.180620.

Liu, L., Leng, L., Liu, C., Lu, C., Yuan, Y., Wu, L., Gong, F., Zhang, S., Wei, X., Wang, M., et al. (2019). An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. Nat. Commun. *10*, 364. https://doi.org/10.1038/S41467-018-08244-0.

Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., et al. (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. Nat. Methods *14*, 1055–1062. https://doi.org/10.1038/NMETH.4436.

Liu, X., Ouyang, J.F., Rossello, F.J., Tan, J.P., Davidson, K.C., Valdes, D.S., Schröder, J., Sun, Y.B.Y., Chen, J., Knaupp, A.S., et al. (2020). Reprogramming roadmap reveals route to human induced trophoblast stem cells. Nature *586*, 101–107. https://doi.org/10.1038/S41586-020-2734-6.

lo Nigro, A., de Jaime-Soguero, A., Khoueiry, R., Cho, D.S., Ferlazzo, G.M., Perini, I., Abon Escalona, V., Aranguren, X.L., Chuva de Sousa Lopes, S.M., Koh, K.P., et al. (2017). PDGFRα + cells in embryonic stem cell cultures represent the in vitro equivalent of the preimplantation primitive endoderm precursors. Stem Cell Rep. *8*, 318–333. https://doi.org/10.1016/J.STEMCR.2016.12.010.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57–63. https://doi.org/10.1038/NATURE11244.

Maeso, I., Dunwell, T.L., Wyatt, C.D.R., Marlétaz, F., Vető, B., Bernal, J.A., Quah, S., Irimia, M., and Holland, P.W.H. (2016). Evolutionary origin and functional divergence of totipotent cell homeobox genes in eutherian mammals. BMC Biol. *14*, 45. https://doi.org/10.1186/S12915-016-0267-0.

Mazid, M.A., Ward, C., Luo, Z., Liu, C., Li, Y., Lai, Y., Wu, L., Li, J., Jia, W., Jiang, Y., et al. (2022). Rolling back human pluripotent stem cells to an eight-cell embryo-like stage. Nature *605*, 315–324. https://doi.org/10.1038/S41586-022-04625-0.

Meistermann, D., Bruneau, A., Loubersac, S., Reignier, A., Firmin, J., François-Campion, V., Kilens, S., Lelièvre, Y., Lammers, J., Feyeux, M., et al. (2021). Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. Cell Stem Cell *28*, 1625–1640.e6. https://doi.org/10.1016/J.STEM.2021.04.027.

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. https://doi.org/10.1016/J.YDBIO.2012.12.008.

Ock, S.A., Knott, J.G., and Choi, I. (2020). Involvement of CDKN1A (p21) in cellular senescence in response to heat and irradiation stress during preimplantation development. Cell Stress Chaperones *25*, 503–508. https://doi.org/10.1007/S12192-020-01090-4.

Pastor, W.A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K., Jacobsen, S.E., and Clark, A.T. (2016). Naïve human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory. Cell Stem Cell *18*, 323–329. https://doi.org/10.1016/J.STEM.2016.01.019.

Petropoulos, S., Edsgärd, 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. https://doi.org/10.1016/J.CELL.2016.03.023.

Posfai, E., Rovic, I., and Jurisicova, A. (2019). The mammalian embryo's first agenda: making trophectoderm. Int. J. Dev. Biol. *63*, 157–170. https://doi.org/10.1387/IJDB.180404EP.

Rodriguez-Terrones, D., Gaume, X., Ishiuchi, T., Weiss, A., Kopp, A., Kruse, K., Penning, A., Vaquerizas, J.M., Brino, L., and Torres-Padilla, M.E. (2018). A molecular roadmap for the emergence of early-embryonic-like cells in culture. Nat. Genet. *50*, 106–119. https://doi.org/10.1038/S41588-017-0016-5.

Rostovskaya, M., Andrews, S., Reik, W., and Rugg-Gunn, P.J. (2022). Amniogenesis occurs in two independent waves in primates. Cell Stem Cell *29*, 744–759.e6. https://doi.org/10.1016/j.stem.2022.03.014.

Senft, A.D., and Macfarlan, T.S. (2021). Transposable elements shape the evolution of mammalian development. Nat. Rev. Genet. *22*, 691–711. https://doi.org/10.1038/S41576-021-00385-1.

Shahbazi, M.N. (2020). Mechanisms of human embryo development: from cell fate to tissue shape and back. Development *147*, dev190629. https://doi.org/10.1242/DEV.190629.

Sharma, R., Rodić, N., Burns, K.H., and Taylor, M.S. (2016). Immunodetection of human LINE-1 expression in cultured cells and human tissues. Methods Mol. Biol. *1400*, 261–280. https://doi.org/10.1007/978-1-4939-3372-3_17.

Singh, H., and Aplin, J.D. (2015). Endometrial apical glycoproteomic analysis reveals roles for cadherin 6, desmoglein-2 and plexin b2 in epithelial integrity. Mol. Hum. Reprod. *21*, 81–94. https://doi.org/10.1093/MOLEHR/GAU087.

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. https://doi.org/10.1242/DEV.158501.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, 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. https://doi.org/10.1016/J.CELL.2014.08.029.

Taubenschmid-Stowers, J., Rostovskaya, M., Santos, F., Ljung, S., Argelaguet, R., Krueger, F., Nichols, J., and Reik, W. (2022). 8C-like cells capture the human zygotic genome activation program in vitro. Cell Stem Cell *29*, 449–459.e6. https://doi.org/10.1016/J.STEM.2022.01.014.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., Pontis, J., Wang, H., Iouranova, A., Imbeault, M., et al. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. Cell Stem Cell *19*, 502–515. https://doi.org/10.1016/J.STEM.2016.06.011.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. Cell Stem Cell *15*, 471–487. https://doi.org/10.1016/J.STEM.2014.07.002.

Töhönen, V., Katayama, S., Vesterlund, L., Jouhilahti, E.M., Sheikhi, M., Madissoon, E., Filippini-Cattaneo, G., Jaconi, M., Johnsson, A., Bürglin, T.R., et al. (2015). Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. Nat. Commun. *6*, 8207. https://doi.org/10.1038/NCOMMS9207.

Upton, K.R., Gerhardt, D.J., Jesuadian, J.S., Richardson, S.R., Sánchez-Luque, F.J., Bodea, G.O., Ewing, A.D., Salvador-Palomeque, C., van der Knaap, M.S., Brennan, P.M., et al. (2015). Ubiquitous L1 mosaicism in hippocampal neurons. Cell *161*, 228–239. https://doi.org/10.1016/J.CELL.2015.03.026.

Wallace, N.A., Belancio, V.P., and Deininger, P.L. (2008). L1 mobile element expression causes multiple types of toxicity. Gene *419*, 75–81. https://doi.org/10.1016/J.GENE.2008.04.013.

Wang, Y., Zhao, C., Hou, Z., Yang, Y., Bi, Y., Wang, H., Zhang, Y., and Gao, S. (2018). Unique molecular events during reprogramming of human somatic cells to induced pluripotent stem cells (iPSCs) at naïve state. Elife *7*, e29518. https://doi.org/10.7554/ELIFE.29518.

Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., Guo, J., Ai, Z., Niu, Y., Duan, K., et al. (2020). A developmental landscape of 3D-cultured human pre-gastrulation embryos. Nature *577*, 537–542. https://doi.org/10.1038/S41586-019-1875-Y.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1131–1139. https://doi.org/10.1038/NSMB.2660.

Zhao, C., Reyes, A.P., Schell, J.P., Weltner, J., Ortega, N.M., Zheng, Y., Björklund, Å.K., Rossant, J., Fu, J., Petropoulos, S., and Lanner, F. (2021). Reprogrammed blastoids contain amnion-like cells but not trophectoderm. Preprint at bioRxiv. https://doi.org/10.1101/2021.05.07.442980.

Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A.T.L., Marioni, J.C., and Reik, W. (2019). Transcriptional heterogeneity in naive and primed human pluripotent stem cells at single-cell resolution. Cell Rep. *26*, 815–824.e4. https://doi.org/10.1016/J.CELREP.2018.12.099.