Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification

Dimitri Meistermann[1,3,*], Alexandre Bruneau[1,*], Sophie Loubersac[1,2,*], Arnaud Reignier[1,2], Julie Firmin[1,2], Valentin François - Campion[1], Stéphanie Kilens[1], Yohann Lelièvre[3], Jenna Lammers[2], Magalie Feyeux[1,4], Phillipe Hulin[4], Steven Nedellec[4], Betty Bretin[1], Gaël Castel[1], Nicolas Allègre[9], Simon Covin[1], Audrey Bihouée[4,5], Magali Soumillon[6,7], Tarjei Mikkelsen[6,8], Paul Barrière[1,2], Claire Chazaud[9], Joel Chappell[10], Vincent Pasque[10], Jérémie Bourdon[3], Thomas Fréour[1,2,†], Laurent David[1,4,11†].

1. Université de Nantes, CHU Nantes, Inserm, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ITUN, F-44000 Nantes, France;
2. CHU Nantes, Université de Nantes, Service de Biologie de la Reproduction, Nantes, France;
3. LS2N, UNIV Nantes, CNRS, Nantes, France;
4. Université de Nantes, CHU Nantes, Inserm, CNRS, SFR Santé, FED 4203, Inserm UMS 016, CNRS UMS 3556, Nantes, France;
5. Institut du thorax, UNIV Nantes, INSERM, CNRS, Nantes, France;
6. Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA; Broad Institute, Cambridge, MA 02142, USA; Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA;
7. Present address: Flexomics LLC; 38 Wareham Street, Fl 3; Boston, MA 02118
8. Present address: 10x Genomics, 7068 Koll Center Pkwy #401, Pleasanton, CA 94566, USA.
9. GReD laboratory, Université Clermont Auvergne, CNRS, Inserm, Faculté de Médecine, CRBC, F-63000 Clermont-Ferrand, France
10. KU Leuven - University of Leuven, Department of Development and Regeneration, Leuven Stem Cell Institute, Herestraat 49, B-3000 Leuven, Belgium
*Authors contributed equally
[†]co-corresponding authors
11. Lead Contact

Correspondence: thomas.freour@chu-nantes.fr, laurent.david@univ-nantes.fr
Lead Contact: laurent.david@univ-nantes.fr

Summary

Understanding lineage specification during human pre-implantation development is a key requirement to improve assisted reproductive technologies and stem cell research. Single-cell RNA-seq has allowed insights into the processes orchestrating human pre-implantation development. However, the exact sequence of molecular events leading to cell fate specification remains to be defined. Here, we employ pseudotime analysis of scRNA-seq data to reconstruct the developmental progress of early mouse and human embryo development.

Using time-lapse annotated embryos, we provide an integrated analysis allowing precise ordering of continuous transcriptomic changes throughout human development. Our analysis revealed that human trophectoderm/inner cell mass transcriptomes diverge at the transition from the B2 to B3 blastocyst stage, just before blastocyst expansion. Moreover, we explore the dynamics of fate markers IFI16 and GATA4 and show that they gradually become mutually exclusive upon establishment of epiblast and primitive endoderm fates, respectively. We also provide evidence that NR2F2 marks trophectoderm maturation initiating from the polar side, and subsequently spreads to all cells after implantation. Altogether, our study pinpoints the precise timing of lineage specification events in the human embryo and identifies transcriptomic hallmarks and cell fate markers.

Keywords
epiblast; trophectoderm; pluripotency; cell fate; lineage specification; blastocyst; preimplantation development; scRNA-Seq; pseudotime; Human and mouse embryos

Introduction

A key question in biology is how cell fate specification takes place in the early human embryo. The main information we have from this period of development is the changes in morphology of the embryo. Concomitantly with morphological changes and growth of the embryo, the pluripotent epiblast and the trophectoderm cells are established. Since human pluripotency is established during human preimplantation development, a better understanding of early human embryogenesis has a strong impact on the field of pluripotent stem cells. Human preimplantation development is also important for fertility treatment, and in particular in vitro fertilization.

The main morphological events occurring during pre-implantation development are compaction and cavitation. Compaction is characterized by the tight interactions that form between cells after the 8-cell stage; cavitation is the formation of a single cavity within the embryo, through fluid pumping. Compaction and cavitation must be coordinated with lineage specification for normal development of the embryo. Indeed, work in mice has revealed that establishment of polarity in the outer cells of the morula, subsequent to compaction, is necessary for lineage specification (Hirate et al., 2013), and cavitation has been shown to be necessary for proper TE fate progression (Chan et al., 2019; Dumortier et al., 2019). Whether the regulation of morphogenesis is similar in human remains unclear.

Lineage specification in mice is considered to be driven by two sequential events. The first cell fate decision in the morula segregates the outer trophectoderm (TE) cells from inner cell mass (ICM) cells (White et al., 2018). Subsequently, in the blastocyst, ICM cells undergo a second cell fate decision, to form the pluripotent epiblast (EPI) and primitive endoderm (PrE),

precursor of the embryo proper and yolk sac, respectively (Chazaud et al., 2006). Whether this sequence is conserved in human embryos has been the subject of debate. On one hand, recent single-cell transcriptomic analysis of human pre-implantation embryos suggested the simultaneous establishment of TE, EPI and PrE lineages rather than a two-step model like in the mouse (Petropoulos et al., 2016). On the other hand, integrated analysis of several single-cell transcriptome studies of early human embryos supported the existence of an early human ICM molecular state, distinct from EPI and PrE, implying a two-step model, but could not definitively exclude the concurrent cell lineage specification model (Singh et al., 2019; Stirparo et al., 2018). Therefore, the sequence of events driving lineage specification in human remains unclear.

In mouse, the first cell fate decision, segregating TE from ICM, occurs at the morula stage and is driven by a YAP/TEAD4/CDX2 axis (Chazaud and Yamanaka, 2016). Initially, outer cells are polarized and express aPKC and Amot, regulators of the Hippo pathway, at their apical membrane. In these outer polarized cells, the Hippo pathway permits Yap1 to translocate into the nucleus where it binds to Tead4. The Yap/Tead4 complex activates TE fate drivers such as Cdx2, while non-polarized inner cells engage in an ICM fate.

Little is known about TE fate specification in human. CDX2 expression appears after cavitation in human embryos (Niakan and Eggan, 2013) and is therefore unlikely to be driving the TE fate as seen in mouse. Only few TE markers have been discovered in humans: GATA3, LAMA1, LAMA3 and KRT7 in hatched/post-implantation blastocysts; GATA2 in blastocysts after the B2/B3 stage and CCR7 in polar TE cells (Aberkane et al., 2018; Deglincerti et al., 2016; Hannan et al., 2010; Kilens et al., 2018; Niakan and Eggan, 2013; Petropoulos et al., 2016; Shahbazi et al., 2016). Moreover, the molecular events that enable competent human blastocysts to interact with receptive luminal endometrium have remained largely elusive. An important observation has been that human blastocysts attach to the endometrial surface epithelium from the polar side ((Aberkane et al., 2018; Grewal et al., 2008; Lindenberg, 1991). This suggests specific molecular events occurring on the polar TE side of the blastocyst, however the molecular events and underlying mechanisms remain to be proposed.

The second cell fate decision, separating ICM cells into EPI and PrE, occurs at the blastocyst stage in mouse. This second lineage specification is driven by a Nanog/Fgf4/Gata6 axis (Chazaud and Yamanaka, 2016). In EPI cells, Nanog induces Fgf4 secretion; Fgf4 will induce Gata6 expression and in turn, Gata6 inhibits Nanog in those cells, yielding PrE fated cells. In human embryos, GATA6 specific expression in PrE occurs upon implantation, while it occurs at 90 cells blastocyst stage in mouse (Deglincerti et al., 2016). Additionally, FGF signaling does not seem to regulate EPI/PrE specification in humans (Kuijk et al., 2012; Roode et al., 2012). Altogether, those results demonstrate differences with the mouse model. scRNA-seq studies have provided atlases of gene expression during human pre-implantation development but it is still not clear if and when cell fate decisions take place (Blakeley et al., 2015; Petropoulos et

al., 2016; Singh et al., 2019; Yan et al., 2013). Altogether, how distinct molecular regulators relate to specification of TE, EPI and PrE lineages in humans remains to be further elucidated.

In this study, we define the precise kinetics of gene expression and cell fate decisions during mouse and human early embryo development. We coupled morphological staging with molecular events, highlight differences and similarities between mouse and human embryos. Lastly, we identify proteins that mark EPI, TE and PrE fates progression. This study provides a resource to study cell fate decisions during early human development.

Results

Designing an experimental approach coupling fine-staging with scRNA-seq analysis

One of the main difficulties when comparing mouse and human development is the differences in timing of development and the specificity of each developmental stage per species. In mouse, embryonic days are accurately linked to a morphology and molecular events (**Figure S1A**). In human, however, *in vitro* culture of embryos shows discrepancies between timing of culture and morphology. Therefore, reproductive biologists rely on a morphology-based annotation in human, with precise criteria (**Figure 1A**)(Alpha Scientists in Reproductive and Embryology, 2011). The numerous morphological events occurring during the fifth day post fertilization (dpf) illustrate this: at 5 dpf, human embryos progress from early blastocyst (B1 and B2) to blastocyst (B3) and expanded blastocyst (B4) (**Figure 1A**). Molecular analysis must be linked to morphological staging in order to better understand human pre-implantation development.

Pseudotime analysis of scRNAseq data appeared to be a good approach to study lineage specification and hierarchize molecular events. We used Monocle2 to generate developmental cell fate trajectories and identify the likeliest path that temporally orders transcriptomic signatures of single-cells (Qiu et al., 2017)(**Figure 1A**). In order to identify lineage signatures, which is an essential benchmark for stem cells models, we clustered cells based on gene correlations. Finally, as we used embryos donated to research, carefully staged using time-lapse monitoring and *zona pellucida* thickness measurement (Feyeux et al., 2020), and graded according to clinical criteria, we were able to link embryo morphology with molecular events.

Transcriptomic pseudotime analysis of mouse and human pre-implantation development

We first computed the pseudotime of mouse pre-implantation embryo development using two existing datasets spanning early 16-cell morula to 90 cell stage blastocysts (Allègre et al., 2019; Posfai et al., 2017). Pseudotime analysis identified a common pool of progenitor cells

that subsequently separates in two branches (**Figure 1B**). The pseudotime value indicates the level of transcriptomic change from the root cell (**Figure S2B**). Projecting the developmental stage annotations onto the pseudotime identified that the initial unspecified branch is composed mostly of cells from early 16-cell morula. Late 16-cell morula cells occupy the bifurcation point at which cells make their first fate decision. Future cells, from early 32 and late 32-cell morula, subsequently occupy one of the two branches, toward the ICM and TE fates. Cells from 64 and 90 cell blastocysts are all located at the tips of specified branches of the pseudotime (**Figure 1B**). Enrichment of Sox2, Cdx2 and Sox17 expression allow identification of EPI, TE and PrE branches, respectively (Frum et al., 2018)(**Figure 1C**). Altogether, pseudotime analysis allowed us to reconstruct the developmental paths taken during early mouse embryo development, including the expected TE/ICM branching point as well as EPI/PrE bifurcation.

Having established the developmental pathways of early embryo cells in mouse, we turned our focus to human embryos. We first generated scRNA-seq data of 25 human embryos from morula to B5 blastocyst (150 cells). We integrated our scRNA-seq dataset with available datasets (Fogarty et al., 2017; Petropoulos et al., 2016; Yan et al., 2013), resulting in an atlas of 1751 cells from 128 embryos. Pseudotime analysis revealed 5 branches, with each branch contains samples from at least 2 datasets (**Figure 1D, Figure S2A**). As our embryos were followed by time-lapse microscopy prior to scRNA-seq, we could univocally associate developmental stages for each embryo. Moreover, as blastocysts were laser-dissected to physically separate the mural TE from the polar TE/ICM, we could also overlay positional information on the pseudotime (**Figure 1D, Figure S1B-D**). This revealed the positioning of morula cells in the middle of the first branch of the pseudotime and the TE branch as the longest branch extending on the right (**Figure 1D**). Projection of the expression levels of transcription factors suggested branches predominantly associated with EPI, PrE and TE, enriched for KLF17 (EPI), SOX17 (PrE) and GATA2 (TE), respectively (**Figure 1E**) (Blakeley et al., 2015; Kilens et al., 2018; Niakan and Eggan, 2013; Xiang et al., 2020).

To link developmental stage and molecular events, we performed a per-embryo analysis in which cells from the same embryo were analysed separately. We then regrouped embryos based on their position on the pseudotime (**Figure S2B**). Sample images are shown for blastocyst ID=13 (B4 stage upon scRNA-seq analysis) (**Figure 1F**). We projected all cells from our cohort of embryos on the pseudotime linking embryo annotation to the pseudotime (**Figure 1G, Figure S2C**). Our precise annotation showed that in human, TE and ICM cells become transcriptionally distinct between the B2 and B3 stages, after the beginning of blastocyst cavitation (**Figure 1F, Figure S2C)**. These results were in stark contrast with mouse embryos, where branching between ICM and TE fate occurred earlier, between 16 and 32 cells morula stage (**Figure 1B**). Our results suggest that in human embryos, ICM cells are transcriptionally distinct from TE cells shortly after the formation of the blastocyst, unlike in the mouse where this event takes place earlier.

Defining hierarchy of gene expression and lineage signature during human preimplantation development

The next step was to define lineages associated with each cell. Pseudotime analysis is particularly suited to study pre-implantation development as a dynamic process. We used WGCNA (weighted gene co-expression network analysis) (Langfelder and Horvath, 2008) to identify gene expression signatures associated with distinct developmental stages and lineage. Gene module activation or repression can be summarized by a linear combination of gene expression values called module eigengenes. To produce a precise mapping of transcriptome identity we used the module eigengenes from gene modules reduced in two dimensions by a UMAP approach (Uniform Manifold Approximation and Projection) (Leland McInnes, 2018). Finally, we clustered cells based on their relative position on the UMAP (Martin Esterk and 1996). This featured 8 major clusters indicative of specific lineages and stages; we qualified cells located in-between clusters as undefined or "intermediates" (**Figure 2A, Figure S3**).

Module eigengenes classify EPI, PrE and TE cells in an unsupervised manner. In contrast with Petropoulos et al. (2016) who used successive rounds of PCA to identify the different cell types and the loadings to identify the associated markers, our method gives a more confined identification of PrE, a refined association of fates with developmental stages and also identifies cells that are undefined or "intermediates", limiting the noise in the fate-associated signatures (**Figure S4A**). Stirparo et al. (2018) classified E5 non-TE cells as ICM. The majority of the cells they identified as ICM are within our EPI cluster (**Figure S4B**). We did not identify a stringent population of unspecified cells that could we considered as a proper ICM cluster. However, our analysis highlighted 3 clusters of unspecified cells and subdivided the TE cells in 3 sub-clusters based on their progression on the pseudotime (**Figure 2B**). Our unsupervised analysis therefore identified developmental molecular states during pre-implantation development and provides an alternative, refined cellular annotation of cell fate lineages in the early human embryo.

To further analyze the temporal dynamics of gene expression during human pre-implantation development, we matched the gene module expression with each cluster of cells obtained with the UMAP (**Figure 2C**). This revealed the unique behavior of 8 modules of genes, named after a representative gene (**Figure S3**). The 3 clusters of unspecified cells are distinguished based on the expression of (1) the ZSCAN4 module, exclusively within the 8 cell stage embryos, (2) the DUXA module and onset of the POU5F1B and DNMT3L modules in morula cells and (3) the absence of the ZSCAN4 module, the fading of the DUXA modules and the presence of the POU5F1B/DNMT3L modules, in the early blastocysts (**Figure 2C, Figure S4C**). Regarding TE cells, 3 states emerge from the pseudotime analysis: (1) a cluster expressing the GATA2, POU5F1B and DNMT3L modules that we named early TE, (2) a cluster expressing the GATA2 module but not POU5F1B and (3) a cluster expressing GATA2 and NR2F2 modules (**Figure 2C,**

Figure S4C). These analyses defined distinct modules of genes whose dynamic expression characterizes early human embryo development.

The dynamic expression pattern of modules within clusters and pseudotime inform on the potential role of genes present in each module. Some modules are not associated with fates, but rather reflect overall changes in embryo development. For example, the ZSCAN4 and DUXA modules are associated with zygotic genome activation between the 8-cell and morula stages; while the DNMT3L module is emerging in all cells at the blastocyst stage and contains epigenetic regulators, in line with methylation dynamics in human pre-implantation development and metabolic pathways (**Figure S4D**) (Greenberg and Bourc'his, 2019). On the other hand, some modules are associated with fates and are expressed sequentially. The POU5F1B module, containing pluripotency genes, is first detected in all cells of the morula, B1 and B2 blastocysts, and is then maintained in the epiblast but reduced in the PrE and TE. Consequently, the POU5F1B module is the first module which goes through a process of global to restricted lineage specific expression in more developed embryos (**Figure 2C**). Other modules are related to specific lineages, for example, the IFI16 module for developed/stable-EPI, GATA4 module for PrE, GATA2 module for early-TE/medium-TE/late-TE, and NR2F2 module for late-TE. Indeed, expression of several genes within lineage-specific modules align with the current understanding of developmental expression dynamics. For example, the IFI16 module contains NANOG and overlaps with expression of the POU5F1B module in B4 blastocysts (**Figure 2C**). The GATA4 module contains characterized regulators of PrE fate, SOX17 and FOXA2. Our dual analysis by pseudotime and gene correlation allowed us to hierarchize states emerging during pre-implantation development, linking them with developmental stages (**Figure S4C**). This highlights how our analysis can be used to refine human pre-implantation studies by serving as a reference for developmental stages and the progression of molecular signatures which lead towards specification. All annotations and expression profiles can be found in a user-friendly online browser that we generated to ease access to human and mouse pre-implantation datasets (**Figure S4E**).

Human morula and B1/B2 blastocyst cells show heterogeneity

Our pseudotime analysis combined with time-lapse imaging of embryos suggest that embryo cells are transcriptionally different after human cavitation. To refine this conclusion, we studied transcriptional heterogeneity within each embryo by clustering embryos based on their individual position on the pseudotime. Cells belonging to each embryo can only belong to one cluster. We then measured the pseudotime spreading of each cluster, as it reflects transcriptional variability, and reported annotations associated with embryos from each cluster. For example, cluster 1 spans a relatively short pseudotime distance and contains only 3 dpf embryos (**Figure 3A**). Clusters K3 and K4 have the highest pseudotime distance variation (**Figure 3B**). Interestingly, cluster K3 contains mostly 4 dpf embryos, suggesting that within human morula, some cells are maturing faster (**Figure 3C**).

To test whether morula cells, located just prior to the first pseudotime bifurcation, have already initiated a transcriptional program biased towards a fate, we used RNA velocity, a measurement of non-spliced mRNA that infers cellular fate (La Manno et al., 2018). In mouse, RNA velocity analysis of transcriptionally similar cells showed that cells had a velocity trending toward ICM or TE (**Figure 3D**). Thus, at the 16-cell stage, mouse cells are biased towards TE or ICM, but the transcriptomic (mRNA) differences are mostly visible at the 32-cell morula stage (**Figure 3D**)(Jedrusik et al., 2008). Similarly, in human, RNA velocity showed that despite having similar mRNA profiles, unspecified cells localized on the peudotime prior to branching are already biased towards ICM or TE fates, in agreement with mouse work (Graham and Zernicka-Goetz, 2016) (**Figure 3E**). This can be further observed when projecting the velocity vectors on the UMAP cluster B1/B2 (**Figure 3F-G**). Altogether, heterogeneity in human embryos seemed to be established at the morula stage, but inner and outer cells are not transcriptionally distinct until B3 blastocyst stage.

Identification of markers of compaction and cavitation during human pre-implantation development

Compaction of cells at the morula stage requires establishment of apico-basal polarity. In mouse, compaction is initiated or characterized by the sequential establishment of adherent junctions, marked by CDH1, followed by the establishment of tight-junctions, marked by ZO1 (Barcroft et al., 1998; Barcroft et al., 2003; White et al., 2018). To refine our pseudotime, we investigated establishment of apico-basal polarity in human embryos. Immunostaining of ZO1 and CDH1 during human pre-implantation development showed basolateral expression of CDH1 from the morula stage, with subsequent apical recruitment of ZO1 (**Figure 4A**). This supports that apico-basal polarity in human embryos is linked to molecular events similar to other mammalian species (Reijo Pera and Prezzoto, 2016).

After acquisition of apico-basal polarity and compaction, an important event in mammalian pre-implantation development is the formation of a blastocoel cavity (White et al., 2018). This is mediated by the expression of aquaporins at the morula/early blastocyst stage (Barcroft et al., 2003; Xiong et al., 2013). AQP3 is the main water channel expressed during human pre-implantation development and is one of the few genes that are initially expressed in all cells but then restricted to TE. Immunostaining for AQP3 revealed cytoplasmic membrane-localization at the morula stage, before being restricted apically at the B2 stage (**Figure 4B**). Loss of AQP3 in the ICM at the B2 blastocyst stage could therefore indicate the onset of a program linked to the ICM has started, consistent with our pseudotime analysis.

Altogether, these results allowed us to place the major morphogenesis events, compaction and cavitation, on the unspecified branch of the pseudotime (**Figure 2D**). CDH1, ZO1 and AQP3 can therefore be used as markers to identify molecular progression of human embryos.

IFI16 and GATA4 highlight progression of EPI and PrE specification

Currently, the number of markers used to study EPI/PrE lineage specification in human is limited. In particular, we need markers that would allow us to track progression of cells within EPI or PrE fates. Using our integrated transcriptomic analysis, we sought to identify markers of developed-EPI in B5 hatching blastocyst. Among transcription factors, we focused on IFI16 as a candidate for developed-EPI. We analyzed protein expression of IFI16 in B5 embryos which showed IFI16 positive staining restricted to ICM cells in contrast to GATA2 expression in TE cells (**Figure 5A**). IFI16 was not detected before B2 and was robustly detected after B4 in ICM cells (**Figure 5B**). These results are consistent with previous transcriptional analyses suggesting a link between IFI16 and EPI (Boroviak et al., 2018).

To further understand IFI16 dynamics within the human ICM, we performed co-staining with NANOG and GATA4, which are mutually exclusive in expanded blastocysts (**Figure 5C**)(Roode et al., 2012). In B4 and B5 embryos, IFI16 co-localized either with NANOG (EPI) positive cells or with GATA4 (PrE) positive cells. To clarify whether IFI16 was restricted to EPI and/or PrE upon their maturation, we extended our analysis to later developmental time points. To do so, we used a recently established *in vitro* culture protocol to simulate implantation (Deglincerti et al., 2016; Shahbazi et al., 2016). Here, B3 or B4 embryos were cultured for 72h *in vitro*, which are considered equivalent to "day 8" human embryos. We observed NANOG-positive cells which clearly express IFI16, while GATA4-positive cells are either IFI16-negative or express IFI16 weakly (**Figure 5D**). Quantification of immunofluorescence in pre-implantation embryos and B3 or B4 embryos cultured 72h *in vitro* suggests that before implantation, ICM cells broadly express IFI16, whereas after implantation, cells expressing high levels of IFI16 are negative for GATA4 (**Figure 5E-F**).

We found that IFI16 is restricted to EPI post-implantation in human embryos. Loss of IFI16 might therefore indicate advancement towards the PrE fate. This observation raised an interesting hypothesis: PrE could be specified later than EPI in human (EPI and PrE are differentially linked to dpf, linear regression, Pvalue = 1.51e-12, **Figure S5A**). The difficulty to identify a distinct ICM molecular signature that would correspond to progenitors of EPI and PrE like in mouse might indicate that EPI and ICM molecular fates in human are highly similar, or that PrE in human differentiates from EPI. To better understand PrE fate, we analyzed a recent scRNA-seq dataset of human embryos cultured up to 120h ("day 5"), then switched to IVC media for up to 216h ("day 14") (Zhou et al., 2019). We focused our analysis on 24h and 72h prolonged culture. UMAP showed that EPI and PrE cells were more related to each other (closer) at day 5 + 24h than at day 5 + 72h (**Figure S5E, S5G**). Interestingly, day 5 + 24h embryos had cells co-expressing both *IFI16* and *GATA4*, while this was not the case in day 5 + 72h embryos (**Figure S5F, S5H**). We completed this analysis by projected velocity vectors on the UMAP from our dataset, or the post implantation dataset (**Figure 5G-H** and **Figure S5E-F**). The

vectors show that at late blastocyst stage, the EPI cells are attracted by the PrE cells. This further supports our hypothesis that PrE specification initiates at B4 stage and that PrE might differentiate from cells that are transcriptionally close to EPI in human. Further analysis is required at B3 stage to identify the molecular states of cells that would eventually become PrE.

To further validate our observations, we investigated the expression of another marker: OTX2. OTX2 is an EPI-associated lineage marker in mouse embryos (Acampora et al., 2016; Buecker et al., 2014; Shahbazi et al., 2017) but has been linked to PrE in human (Boroviak et al., 2018; Zhou et al., 2019). In pre- and post-implantation embryos, we observed that OTX2 only co-localized with GATA4 positive cells and is not expressed in NANOG positive cells (**Figure S5B-D**). Interestingly, in "day 8 embryos", OTX2 was expressed in the GATA4 positive cells closer to NANOG-positive cells (**Figure S5C-D**). It remains to be defined whether OTX2 indicates a more advanced PrE development or a different subset of PrE.

Altogether, our transcriptomic analyses validated by IF support that EPI/PrE specification in human occurs in B4, late, blastocysts, which is different from mouse, where EPI and PrE fates are established in 90 cell, early, blastocysts (Schrode et al., 2015).

NR2F2 marks TE maturation at the polar side of human embryos

We next sought to decipher the maturation events in the TE branch. A transcription factor of interest was NR2F2, due to its expression in the later part of the TE branch, which corresponds to developmental stages of expanded/hatched blastocysts (B4 and after, **Figure S3B**). In B5 blastocysts, immunofluorescence showed that nuclear localization of NR2F2 was restricted to polar TE cells juxtaposed to the EPI cells, expressing NANOG (**Figure 6A**). We then investigated NR2F2 expression in B4 blastocysts. Immunostaining revealed GATA3 and GATA2 expression throughout TE, while NR2F2 was only expressed in the subset of polar TE cells (**Figure 6B**). The pseudotime analysis indicates that NR2F2 expressing cells are the most mature, as they are at the end of the trajectory. However, immunofluorescence analysis showed expression of NR2F2 in B4 embryos and B4 TE cells are generally localized in the middle of the TE branch of the pseudotime (**Figure S2C**). This, therefore, suggest that maturation of TE is initiated on the polar side by an EPI-TE dialog.

To further characterize onset of NR2F2 positive cells, we co-stained NR2F2 with markers of the EPI (IFI16) and PrE specification (GATA4). This revealed that NR2F2 expression is concomitant of IFI16 and GATA4 expression (**Figure 6C** and **Figure S6**). Staining of NR2F2 in B3/B4 +72h *in vitro* culture embryos confirmed the spreading of NR2F2 to all TE cells (**Figure 6D**). We therefore propose that maturation of TE is initiated by contact with EPI then spreads upon implantation.

The observation that maturation of TE cells is initiated at the polar side is consistent with the observation that the majority of human blastocysts attached to endometrial cells by the polar side with subsequent spreading of TE development to mural cells (Aberkane et al., 2018; Bentin-Ley et al., 2000; Grewal et al., 2008; Lindenberg, 1991). We grouped cells as NR2F2-positive and NR2F2-negative to highlight pathways specific to polar TE. This revealed pathways for which the receptor is on NR2F2+ polar TE cells (I.e. GM-CSFR), and the ligand is secreted by the endometrium (i.e. GM-CSF) (Giacomini et al., 1995; Ziebe et al., 2013). Analysis of the most enriched signaling pathways with fGSEA (Korotkevich et al., 2019) identified prime candidates potentially driving the molecular dialog from EPI toward TE such as TGFβ, IGF1, BMP2, FGF4, and from TE toward EPI : PDGF, IL6 and WNT (**Figure S7 and Figure 6E**). This suggests that FGF4 might still have a role in preimplantation development, despite being ruled out from regulating PrE specification (Kuijk et al., 2012; Roode et al., 2012). BMP signaling has recently been shown to inhibit cavitation (De Paepe et al., 2019), but its function after the first lineage specification remains to be studied. The role of IL6 has recently been studied in pig: IL6 KO or inhibition of JAK signaling perturbed pre-implantation development, supporting the hypothesis that in human and pigs, in absence of LIF, JAK/STAT signaling is triggered by IL6 (Ramos-Ibeas et al., 2019). Polar TE-EPI interactions in human is reminiscent of the ICM-polar TE molecular dialog observed in mouse blastoids (Rivron et al., 2018). Indeed, mouse blastoids showed that EPI is necessary for proper TE maturation. However, human specific events might be involved as some genes critical for mouse TE maturation are not restricted to TE in human, such as KLF6 (Rivron et al., 2018).

In summary, we found that NR2F2 marks mature TE in B4 expanded blastocyst and spreads to all TE cells. Identification of NR2F2 positive cells will give rise to numerous hypothesis regarding the dual molecular dialog between polar TE and both the EPI and the endometrium.

Discussion

We generated continuous mouse and human pre-implantation embryo transcriptome atlases, by implementing lineage reconstruction approaches from scRNA-seq data. The pseudotime ordered transcriptomic mouse trajectory corroborates existing functional experiments (Chazaud and Yamanaka, 2016; Posfai et al., 2017). For example, cells become transcriptionally distinct at the early 32 cells stage, immediately after differential Cdx2 protein expression, at the late 16-cell morula stage. Experimentally though, mouse TE cells at the 32 cell stage are committed while ICM cells may be specified but not yet committed (Grabarek et al., 2012; Handyside, 1978; Rossant and Lis, 1979). The human pseudotime analysis shows the precise timing of transcriptionally distinct TE, at the transition from B2 to B3 stage. Our embryo per embryo analysis strategy revealed that transcriptomic heterogeneity could be found in human morula, prior to complete segregation of cells into either of the two lineages. Coupling pseudotime with gene correlation analysis revealed the dynamics of several gene

modules which could be associated with specific states within each branch of the pseudotime. This led us to reveal by immunofluorescence that IFI16 is expressed in a majority of ICM cells at the B4 stage, and progressively fades in the developing PrE cells. Moreover, we discovered that maturation of TE was initiated on the polar side, adjacent to the EPI, which yielded a molecular dialog. Altogether, our study improves our understanding of the molecular events occurring during human lineage specification and provides new markers of specific developmental steps (**Figure 7**).

Human in vitro pre-implantation development is highly heterogeneous from one embryo to another. The best way to annotate embryo stage to date is to record development of human embryos with time-lapse devices (technology et al., 2020). The combination of our scRNA-seq analysis of time-lapse annotated human embryos with previous datasets resolved annotation overlap between 4, 5 and 6 dpf, which previously limited the translation of those datasets into a sequence of events, pacing human pre-implantation development. Indeed, binning embryos at 5 dpf results in the averaging of transcriptome information from cavitating B1 blastocyst to B4 expanded blastocysts. Since the TE, PrE and EPI transcriptomes are emerging in that time frame, binning it all together creates a lot of noise in subsequent analyses. Our analysis, combining time-lapse annotation and scRNA-seq, therefore improves our ability to generate hypotheses regarding the molecular events regulating human pre-implantation development.

Two morphogenic events are particularly noticeable during mammalian pre-implantation development: compaction and cavitation. Compaction seems to involve rapid epithelialization in both mouse and human (White et al., 2018). We found that transcriptional variability was greatly increased between cells from the morula, but without achieving fate-associated transcriptomic signatures. The delay in the onset of distinct transcriptomic profiles of TE/ICM in human compared to mouse is rather surprising. This observation suggests that in addition to compaction, other parameters are likely required to induce/complete lineage specification. One of these could be cell division. Human morulae compact at 10 cells and cavitate at 20 cells, resulting in a B1 stage that rapidly evolves to B2 stage blastocyst, generally composed of 40 cells. This therefore suggests that two rounds of cellular division, post-compaction, are required before cells can initiate/complete distinct transcriptomic programs relating to the ICM/TE. This occurs similarly, yet slightly earlier, in mouse development in which cell compaction takes place at the 8-cell stage and the emergence of distinct cellular profiles is first seen at the 32 cell stage. Regarding cavitation, our analysis also revealed the expression of AQP3 which restricted to TE cells at the B2 stage. However, it remains to be demonstrated whether AQP3 expression depends on TE fate regulators or if TE fate regulators depend on AQP3.

EPI/PrE specification is difficult to resolve in human, regardless of the approaches used (Blakeley et al., 2015; Petropoulos et al., 2016; Stirparo et al., 2018). Our analysis suggests that within the ICM, the EPI fate is the first to be achieved. Computational modeling in mouse came

to the same conclusion: EPI cells need to secrete Fgf4 for the PrE cells to specify (De Mot et al., 2016). Regarding PrE fate, we found that PrE signature is better defined when analyses include prolonged cultured embryos: day 6 IVC media-cultured embryos seem to have more developed PrE cells than day 6 IVF media-cultured embryos (Petropoulos et al., 2016; Zhou et al., 2019). This could be because IVF media are not currently designed to support extended culture of embryos. Indeed, It has been determined that day 7 blastocyst were less likely to implant (Shoukir et al., 1998) and are only transferred when no other embryos are available (Tiegs et al., 2019). Another possibility is that IVC media favors rapid growth/development of human embryos. This observation shows how important it is to understand pre-implantation development in human to improve embryo culture media.

We consider two main hypotheses for the EPI/PrE specification in human: 1/ early ICM cells are transcriptionally closer to EPI than to PrE or TE, resulting in ICM cells clustering with EPI cells; 2/ PrE differentiates from EPI. More experiments are necessary to solve those hypotheses, but our study provides a framework: we have pinpointed the morphology associated with specific marker expression and we provide novel markers informative of EPI/PrE specification, such as IFI16 and GATA4. In particular, IFI16 characterizes the transition between molecular states during EPI and PrE peri-implantation development. IFI16 expression suggests that IFNg signaling might be important for EPI specification. Furthermore, IFI16 has been shown to induce epigenetic silencing of viral promoters in immune cells (Roy et al., 2019). Therefore, one hypothesis is that IFI16 might inhibit the burst of transposable elements expressed at the morula stage (Giocanti-Auregan et al., 2016; Theunissen et al., 2016). Shutting down transposable elements might be necessary for EPI/PrE specification.

Transcriptomically, TE acquires a distinct expression profile at the transition from the B2 to B3 stage with cells progressively losing expression of pluripotency-associated genes, correlating with loss of plasticity of TE cells (De Paepe et al., 2013). We provide evidences showing that an EPI/TE molecular dialog could be involved in TE maturation. Interestingly, it is the TE cells adjacent to the EPI, on the polar side, which typically initiate implantation prior to maturation of the remaining mural TE cells. However, in mouse embryos the secretion of Fgf4 and nodal by EPI inhibits differentiation of polar TE, and mouse embryo attachment is mediated by, the more mature, mural TE cells, highlighting differences between the two species (Guzman-Ayala et al., 2004; Yamagata and Yamazaki, 1991). It will be interesting to understand what triggers maturation of TE and whether NR2F2 plays a central role in this process in both mouse and human.

Studies have demonstrated the importance of TE morphological quality for prediction of successful IVF (Ahlstrom et al., 2011; Chen et al., 2014; Hill et al., 2013; Thompson et al., 2013). Therefore, understanding this aspect of human pre-implantation development is of utmost importance for improvement of IVF. Moreover, we will now be able to focus on specific secreted markers associated with NR2F2-positive cells that could inform clinicians about the

implantation potential of the blastocyst before transfer. Indeed, secreted markers that correlate with the expression of NR2F2 could be predictive of implantation capacity of embryos. This discovery will therefore support much needed hypothesis on human implantation, to complement knowledge coming from the study of murine reproduction (Aplin and Ruane, 2017).

Our study also provides key knowledge for the modeling of pre-implantation development with cellular models. Embryonic signature comparison with pluripotent stem cells residing in various states of pluripotency showed established hierarchy of pluripotent stem cells to mimic EPI: naïve pluripotent cells are the closest to pre-implantation EPI (5 dpf to 8 dpf) while primed pluripotent cells, commonly used, are closer to post-implantation EPI (after 10 dpf) (Castel et al., 2020; Kilens et al., 2018; Liu et al., 2017; Takashima et al., 2014; Theunissen et al., 2014). We and others also discovered that naïve pluripotent stem cells can convert to hTSCs, which is reminiscent of late specification in human blastocysts (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2020). Recently, we proposed that human trophoblasts stem cells correspond to human day 8 post-implantation cytotrophoblast (Castel et al., 2020). Our study therefore highlights the importance of knowledge about human peri-implantation development to understand pluripotent and trophoblast stem cell biology.

In conclusion, our study clarifies the timing of molecular events and signaling pathways involved in human pre-implantation development. Fine tuning those pathways will lead to more efficient media formulation for embryo culture during IVF procedures. Improvement of IVF procedures is necessary since the procedural average efficiency is below 27% (Embryology, 2015). Ultimately, our study contributes to understanding human pre-implantation, a gateway to improve IVF success rates.

Limitations of the Study

This work provides a blueprint of human development, but numerous hypotheses that it unveils remain to be validated. For example, we would need to show that an embryo with two ICM would have two maturating “polar” TE. Another important follow-up will be to validate signaling pathways involved in human peri-implantation development.

Acknowledgements

We thank our colleagues F. Lanner and K. Niakan for sharing data and providing feedback. We thank J. Jullien for critical review of the manuscript. DM is supported by FINOX and MSD contributed to the project. We thank core facilities: BIRD, PFiPSC and MicroPicell. This work was supported by “Paris Scientifique region Pays de la Loire: HUMPLURI” and IHU CESTI.

Author contributions

TF and LD designed the study. DM, ABr, SL and LD wrote the manuscript with input from all authors. SL, AR, JF, JL and ABr performed embryo manipulation. PH, SN, SL, ABr, AR and LD

performed IF analysis. AR and SK prepared scRNAseq samples; MS and TM performed single-cell RNAseq. DM, VFC, YL, performed bioinformatics analyses under supervision of ABi and JB and input from BB and GC. TF and PB supervised human embryo donation. All authors approved the final version of the manuscript.

Declaration of interests

Competing interest declaration: DM is supported by FINOX Biotech FORWARD initiative 2016.

Correspondence to:
L. David (laurent.david@univ-nantes.fr), T. Fréour (thomas.freour@chu-nantes.fr).

Figure legends

Figure 1. Pseudotime analysis of human and mouse pre-implantation embryo development

(A) Top-panel: Schematic representation of human pre-implantation development highlighting the link between developmental stages, morphological phases, days post fertilization and average number of cells observed. Mid-panel: summary of datasets used. Lower-panel: description of our experimental design. We used 2 mouse single-cell datasets and 4 human single-cells datasets to identify cell fate trajectories, resulting in a continuous transcriptomic pseudotime. We also performed co-expressed gene module identification with WGCNA and a UMAP was performed on activation score of WGCNA module (module eigengenes) to map cell identities during pre-implantation development. Human embryo used for this study were imaged by a time-lapse microscope

(B) Projection from the mouse scRNA-seq and scqPCR samples (Posfai et al, 2017, Allègre et al, 2021) from the reversed graph embedding method (Monocle2). Developmental stages are indicated as number of cells in the embryo.

(C) Projection of lineage marker expression levels on the mouse model: Sox2 (epiblast), Cdx2 (trophectoderm), Sox17 (primitive endoderm).

(D) Projection of stage and positional information for the samples we analyzed for this study, on the pseudotime. Morula cells (grey) and blastocyst dissection origin of cells are indicated (mural trophectoderm: black, polar trophectoderm or inner cell mass: brown).

(E) Projection of lineage marker expression levels on the human model: KLF17 (epiblast), SOX17 (primitive endoderm), and GATA2 (trophectoderm).

(F) Frames from time-lapse microscopy for embryo “B13”. For each embryo sequenced in this study, their morphokinetics were acquired by time-lapse microscopy. Developmental events include morula compaction (tM), blastulation (tSB) leading to B1 stage, full blastocyst (tB) at B3 stage and expanded blastocyst once the zona pellucida thickness is halved (tEB) at B4 stage. tM is used as t0 to compare thawed embryos.

(G) Projection of cells from 4 embryos, including embryo “B13” (F), on the pseudotime.

(H) Projection of developmental day (3 to 7 dpf) for all samples combined for this study, and the result of our refined staging.

Figure 2. **Gene modules define cell fate transcriptional signatures and pseudotime indicates the hierarchy of events**

(A) Dimension reduction (UMAP) of WGCNA module eigengene, colorized by density-based clusters. Small dots indicate a minor cluster composed by intermediate transcriptome. Medium TE and Late TE were segregated by an additional k-means clustering based on NR2F2 and GATA2 module eigengenes.

(B) Projection of density-based clusters from (A) on the pseudotime.

(C) Expression heatmap of all genes related to WGCNA modules. Each row is a WGCNA module, each column is a set of 50 cells drawn from a UMAP cluster (A). Height of each row represents the number of gene in the corresponding WGCNA module following a log scale.

For each module, a set of genes belonging to the module is indicated on the right side of the heatmap.

Figure 3. Early blastocysts contain a heterogeneous cell population that are preparing to specify despite having similar transcriptomes

(A) Projection of each embryo cluster (k-means) according to the position of their cells on the pseudotime. For each embryo cluster, the corresponding number of embryos is indicated, as well as a representative embryo is indicated. Clusters are ordered by their mean pseudotime.
(B) Range of pseudotime values per embryo cluster.
(C) Distribution of number of cells per embryo cluster subdivided by day post fertilization.
(D) Transcriptome pseudotime and RNA velocity comparison of cells before and after specification in mouse.
(E) Transcriptome pseudotime and RNA velocity comparison of cells before and after specification in human.
(F) RNA velocity vectors projected on the UMAP of WGCNA module eigengene for the cell clusters that are central in the first specification, restricted to the dataset of Petropoulos et al. Cells are colorized by cell clusters.
(G) Summarization of the direction and magnitude of RNA velocity vector of the *B1 and B2* cell cluster by a circular diagram. Each represented vector has an opacity of 20% to represent an estimation of vector density. 114 vectors out 171 are in a sector of 45° of the diagram that is facing the *Early TE* cluster, and 42 are facing the EPI cluster. The sectors are represented by dashed orange line.

Figure 4. Polarization of outer cells of morula precedes restriction of TE marker AQP3

(A) Immunofluorescence of CDH1 (cyan), ZO1 (purple), Actin (yellow) and nuclear counterstaining (white) at indicated stages. Arrowhead points to the apical part of the cell, with ZO1 staining next to CDH1 staining.
(B) Immunofluorescence of AQP3 (yellow) and DAPI (white) as nuclear counterstain at indicated developmental stages.
Scale bar = 47 µm.

Figure 5. NANOG/IFI16/GATA4 dynamics highlight progression of EPI / PrE specification

(A) Immunofluorescence of IFI16 (red) and GATA2 (cyan) at B5 blastocyst stage. z-section indicates a z-cutting plane.
(B) Immunofluorescence of IFI16 at indicated stages. Nuclear counterstaining is in white.
(C) Immunofluorescence of NANOG (cyan), GATA4 (green) and IFI16 (red) at indicated stages. Nuclear counterstaining is in white.
(D) Immunofluorescence of NANOG (cyan), GATA4 (green) and IFI16 (red) in a B3 + 72h attached embryo. Arrow: GATA4+/IFI16+ cell. Arrowhead: GATA4+/IFI16- cell.
(E) Manual counting of GATA4 high/IFI16 negative, GATA4 high/IFI16 low, GATA4 low/IFI16 high, and IFI16 high/GATA4 negative cells in pre-implantation and post-implantation embryos.

(F) Average percentage of total IFI16/GATA4 expressing cells in pre- and post-implantation embryos.
Scale bar = 47 µm.
(G) RNA velocity vectors projected on the UMAP of WGCNA module eigengene for the cell clusters that are central for the second specification, restricted to the dataset of Petropoulos et al. Cells are colorized by cell clusters.
(H) Summarization of the direction and magnitude of RNA velocity vector of the *EPI.PrE* cell cluster by a circular diagram. Each represented vector has an opacity of 45% to represent an estimation of vector density. 18 vectors out of 28 are oriented within an angle of 45° in the direction of the PrE cluster (p-value≈2.21.10-6, obtained from the binomial distribution). The sectors are represented by dashed orange line.

Figure 6. NR2F2 reveals the differential maturation between the polar and mural TE
(A-B) Immunofluorescence of NR2F2 (cyan) and NANOG (red) at B5 stage (A) or with GATA2 (yellow) and GATA3 (purple) at B4 stage (B). Asterisk marks a cell that is negative for NR2F2 and positive for GATA3 / GATA2.
(C) Immunofluorescence of NR2F2 (cyan), GATA4 (green) and IFI16 (red). Nuclear counter-staining (dapi) is in white.
(D) Immunofluorescence of NR2F2 (cyan), GATA4 (green) and IFI16 (red) after 72h of prolonged culture of a B3 blastocyst.
(E) Schematic representing cytokine-receptors loops identified.

Figure 7. Human pre-implantation development model
Each annotation is represented per morphology-based stages of human pre-implantation development. Upper panel: Schematic representation of gene module global expression. Functional enriched functions are presented on the right. Complete enrichment data are presented in Figure S4D**.** Lower panel: expression of lineage specific stage per stage and per cell fate.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact

Further information and requests should be directed to the Lead Contact, Laurent DAVID (laurent.david@univ-nantes.fr)

Materials Availability

This study did not generate new unique reageants

Data and code availability

All original data have been deposited on European Nucleotide Archive under accession number PRJEB30442.

The source code can be retrieved by following the links below.

scRNA-Seq alignment pipeline:

https://gitlab.univ-nantes.fr/E114424Z/SingleCell_Align

scRNA-Seq preprocessing and normalization:

https://gitlab.univ-nantes.fr/E114424Z/singlecellnormalize

WGCNA workflow:

https://gitlab.univ-nantes.fr/E114424Z/WGCNA

Monocle2 workflow, UMAP and cell clustering:

https://gitlab.univ-nantes.fr/E114424Z/monocle2_workflow

Pseudotime User Interface source code:

https://gitlab.univ-nantes.fr/E114424Z/ptuihost

All other parts of the code are available upon request

Experimental model and subject details

Human pre-implantation embryos.

The use of human embryo donated to research as surplus of IVF treatment was allowed by the French embryo research oversight committee: Agence de la Biomédecine, under approval number RE13-010 and RE18-010. All human pre-implantation embryos used in this study were obtained from and cultured at the Assisted Reproductive Technology unit of the University Hospital of Nantes, France, which are authorized to collect embryos for research under approval number AG110126AMP of the Agence de la Biomédecine. Embryos used were initially created in the context of an assisted reproductive cycle with a clear reproductive aim

and then voluntarily donated for research once the patients have fulfilled their reproductive needs or tested positive for the presence of monogenic diseases. Informed written consent was obtained from both parents of all couples that donated spare embryos following IVF treatment. Before giving consent, people donating embryos were provided with all of the necessary information about the research project and opportunity to receive counselling. No financial inducements are offered for donation. Molecular analysis of the embryos was performed in compliance with the embryo research oversight committee and The International Society for Stem Cell Research (ISSCR) guidelines (Kimmelman et al., 2016).

METHODS DETAILS

Human pre-implantation embryos culture.

Human embryos were thawed following the manufacturer's instructions (Cook Medical: Sydney IVF Thawing kit for slow freezing and Vitrolife: RapidWarmCleave or RapidWarmBlast for vitrification). Human embryos frozen at 8-cell stage were loaded in a 12-well dish (Vitrolife: Embryoslide Ibidi) with non-sequential culture media (Vitrolife G2 plus) under mineral oil (Origio: Liquid Paraffin), at 37°C, in 5% O_2/6% CO_2.

Human embryo prolonged culture.

Embryos cultured for 5 days (11xB3 and 5xB4 blastocysts) were thawed as described above. The *zona pellucida* (ZP) of each embryo was removed by brief laser impulse followed by manual extrusion or aspiration of the embryo. Zona pellucida free embryos were washed in GTL medium and immediately transferred in 8-well IbiTreat µ-plates (IB-80826; Ibidi GmbH) with 300µL IVC1 medium (M11-6, Cell Guidance systems)(Deglincerti et al. 2016) and cultured at 37°C, in 21% O2/5% CO2. After 48h of in vitro culture (day 2), 250µL of IVC1 was replaced with 250µL of IVC2 medium (M12-6, Cell Guidance systems). At B3 or B4 + 72h *in vitro* culture, 11/16 embryos (69%) were attached, 3 (19%) were floating and 2 (12%) disagregated.

Human embryo time-lapse imaging.

Embryos were loaded into the Embryoscope® (Vitrolife®), a tri-gas incubator with a built-in microscope allowing time-lapse monitoring of embryo development. Images were captured on seven focal plans every 15-min intervals using Hoffman modulation contrast (HMC) optical setup[1] and a 635 nm LED as light source as provided in the Embryoscope®. The resolution of the camera is 1280x1024 pixels. The development of each embryo was prospectively annotated as described by Ciray et al., by two trained embryologists undergoing regular internal quality control in order to keep inter-operator variability as low as possible (Ciray et al., 2014). ZP thickness was measured by our analysis pipeline (Feyeux et al, 2020). The term tM refers to a fully compacted morula. At the blastocyst stage, tSB is used to describe the onset of a cavity formation, tB is used for full blastocyst i.e the last frame before the ZP starts to thin, tEB for expanded blastocyst, i.e. when the ZP is 50% thinned. Blastocyst contractions and the beginning of herniation were also recorded.

Immunofluorescence of human embryos.

Embryos were fixed at the morula, B1, B2, B3, B4, B5 or B6 stages according to the grading system proposed by Gardner and Schoolcraft (Gardner et al., 2000) or at B3 or B4 + 72h *in vitro* culture. Embryos were fixed with 4% paraformaldehyde for 5 min at room temperature and washed in PBS/BSA. Embryos were permeabilized and blocked in IF Buffer (PBS–0.2% Triton, 10% FBS) at room temperature for 60 min. Samples were incubated with primary antibodies over-night at 4°C. Incubation with secondary antibodies was performed for 2 hours at room temperature along with DAPI counterstaining. Primary and secondary antibodies with dilutions used in this study are listed in **Table S1**.

Imaging.

Confocal immunofluorescence pictures were taken with a Nikon confocal microscope and a 20× Mim objective. Optical sections of 1µm-thick were collected. The images were processed using Fiji (http://fiji.sc) and Volocity 6.3 visualization softwares. Volocity software was used to detect and count nuclei. Two experienced operators graded nuclei signal for each IF channel. Grading was perform independently and counts were then averaged. A nucleus is considered positive for staining if signal is clearly higher than background. “High” or “low” positive staining is up to each operator according to the general signal intensity for each channel. One embryo was photobleached in order to restain it (Figure 6C, Figure S6A). Under Nikon A1 confocal, lasers at 568nm and 647nm were set to 100% power for 10min. The NR2F2 (Rabbit / 488) – SOX17 (Goat / 647) – IFI16 (Mouse / 568) stained embryo was restained with NR2F2 (Rabbit / 488) – GATA4 (Rat / 568) – IFI16 (Mouse / 647).

Single-cell isolation

PGD-trained people performed brief laser impulses to separate the polar TE and ICM and mural TE of blastocysts. Embryos were then incubated 10min in G-PGD/HSA (Vitrolife®) and transferred in accutase. Embryos were triturated with gentle pipette flush under binocular magnifier for 3-4min. Single-cells/clumps were washed in GTL (Vitrolife®) and triturated again if necessary. Single-cells were picked one by one with IVF-lab capillaries.

Single-cell RNA sequencing.

Single-cell isolation and overall dataset are presented in **Figure S1** and **Table S5**. Single-cell RNA-seq libraries were prepared according to the SmartSeq2 protocol with some modifications (Trombetta et al., 2014). Briefly, total RNA was purified using RNA-SPRI beads. Poly(A)+ mRNA was reverse-transcribed to cDNA which was then amplified. cDNA was subject to transposon-based fragmentation that used dual-indexing to barcode each fragment of each converted transcript with a combination of barcodes specific to each sample. In the case of single cell sequencing, each cell was given its own combination of barcodes. Barcoded cDNA fragments were then pooled prior to sequencing. Sequencing was carried out as paired end 2x25bp with an additional 8 cycles for each index. The FASTQ files were mapped with Hisat2 on GRChH38 genome version, downloaded from ensembl.org. HTSeq (Anders et al., 2015) was

used to generate raw counts tables from BAM files, using the matching GTF for the reference genome.

Raw count table treatments.

Samples were filtered with the use of the R function *isOutlier* from SCRAN (Lun et al., 2016) library. This function tags samples as outliers with a threshold based on median derivation away from the median of the metric. We filtered samples with two metrics: the number of expressed genes with a threshold of 2 median away derivation from median, and the total number of counts in the sample with a threshold of 3 median away derivation from median. Both metrics were used to discard samples in a two-sided way, below and above the median. This two-sided filter was applied to remove samples carrying too little (93876 counts - 5558 genes) or too much (80431229 counts – 18711 genes) information. Indeed, these are considered as potential doublet events. Genes that were expressed in less than two cells and with an average expression less than 0.1 were removed.

The four datasets were then normalized together using the *computeSumFactor* function from SCRAN. Logged and non-logged data were collected using the *normalize* function from scater R library. To compute batch effect free expression, we normalized the data as described above but per dataset. We used mutual nearest neighbors correction implemented by the function *mnnCorrect* to achieve the batch correction from the log-normalized data. The reference dataset that were used for *mnnCorrect* is from Petropoulos et al. (2016).

Computation of pseudotime.

Monocle2 needs a subset of genes to make the dimensionality reduction and pseudotimes trajectories. To choose the best set of ordering genes we took samples that passed the quality control from Petropoulos et al. to avoid batch effects. We used SCRAN for processing size factors (normalization factor). We then created an R object with the *newCellDataSet* object used by Monocle2 with the raw expression and the expressionFamily parameter set as "negbinomial.size()". Size factors were attributed according to the SCRAN results. The next step consisted of estimating empirical dispersion of each gene in the negative binomial model with the *estimateDispersions* function. We used the dispersion table function to gather the empirical dispersion and the fitted theoretical dispersion for each gene. We made a ratio of empirical dispersion on the theoretical dispersion for each gene. This ratio describes an over-dispersion score.

For a given gene *i* the over-dispersion score S is calculated as follows: $S_i = \frac{\sigma^2_{emp\ i}}{\sigma^2_{theo\ \mu_i}}$

Genes with an average log expression < 0.5 across samples were filtered out. Remaining genes were ranked based on their over-dispersion score.

Pseudotimes were generated using a range of top ranked ordering genes from the top 500 to the top 5500. This led to 5000 pseudotimes. For each pseudotime, A new R object was created with the newCellDataSet function, with the batch corrected expression from the four datasets as input and the expressionFamily parameter set as "gaussianff()". Selected ordering genes

were then set as input of Monocle2 algorithm, with the number of resulting dimensions set to three dimensions. An automatic classification of pseudotimes was set up following three criteria based on their topology:

- Number of branches populated by mural TE cells.
- Succession of developmental stage.
- Position and number of branching points.

The most common topologies were: (i) all mural TE cells within one branch, (ii) developmental stages succeeding one other, i.e. morula between 8-cell and blastocysts, (iii) two branching points, (iv) first branching point at E5. The chosen pseudotime belonged systematically to the most abundant topologies and is calculated from 4484 ordering genes. The resulting 3-dimension pseudotime was rotated to obtain a 2-dimensional projection.

WGCNA.

WGCNA (Langfelder and Horvath, 2008) was performed on batch corrected data using a soft power of 10 with signed Pearson correlation. Resulting module were manually curated to choose a set of 8 modules that were well represented in data and that have distinct behaviors. For each module we use the module eigengene metric that is given by WGCNA to infer the global module expression across the samples. A loess regression of eigengene by pseudotime was used for **Figure S4A**.

UMAP and cell clustering

UMAP was computed with the R library uwot. Module eigengenes from the 8 studied WGCNA modules were used as features for the UMAP algorithm. The n_neighbors parameter was set to 1751 (size of the dataset) and min_dist = 0.01. Density clustering of cells was performed with dbscan with eps = 0.21 and minPts = 5. 52 outliers were attributed manually to existing clusters. An additional kmeans (k=2) was performed on the main TE cluster from NR2F2 and GATA2 module eigengenes to separate the medium and the late TE. (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Sergushichev, 2016) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Korotkevich et al., 2019) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015) (Liu et al., 2015)

Loess regressed expression by pseudotime.

We used a Locally Weighted Regression (LOESS) to fit expression in the pseudotime by cell fate with a neighbor impact of 0.75. Expression profiles of common segments were fitted to extract global tendencies. A last LOESS was computed with a low neighbor impact to merge segments to obtain continuous expression curves.

Subdivision of pseudotime branches.

The original pseudotime was constituted by five states, as Monocle2 separates states only by branching point. We subdivided pre-specification and trophectoderm branch samples by using WGCNA module eigengene. For both branches, WGCNA modules with a Pearson correlation higher than 0.75 with the pseudotime were selected. A loess regression of these module eigengene by pseudotime was performed, followed by a hierarchical clustering of regressed module eigengenes. The clustering was then partitioned. For each branch the best partition was determined at three clusters with the greatest relative loss of inertia method.

Data representation used in each figure.

Raw expression:

- estimating gene dispersion and select ordering genes

Normalized expression:

- projection of gene expression on pseudotime or UMAP (**Figure 1C, Figure 1E**, **Figure S3B**))
- Heatmap (**Figure 2C**)
- Pseudotime User Interface (**Figure S4E)**

Loess regressed expression by pseudotime:

- expression profile curves (**Figure S7**)

Batch corrected logged expression:

- computation of pseudotime
- computation of WGCNA modules

ComplexHeatmap(Gu et al., 2016), ggplot2 and d3.js. were used for graphical representation. Hierarchical clustering was done using the Ward criteria and from a correlation distance for the gene/pathway eigengenes, or from the euclidean distance for other metrics.

Mouse single-cell RNA-Seq analysis.

Mouse dataset were analyzed in a similar way to human datasets, without batch correction. Alignment step was done from the mm10 version of the genome. Timing of blastulation is corroborated by time-lapse.

RNA velocity

RNA velocity was performed from BAM of samples that have passed all quality control in the final counts table. First, we used velocyto.py using the command *velocyto run*, with the parameter --logic as “SmartSeq2”, and the parameter -m (RepeatMasker annotations) as a GTF downloaded from the UCSC genome browser. The global GTF was the same that were used for the computation of raw counts table. Resulting loom files were merged using loompy.combine from lompy python package. We used velocyto.R for computing Velocity matrix. Loom files were read with the function *read.loom.matrices*. Then we separated spliced reads matrix, unspliced reads matrix and spanning reads matrix. For each of the matrices gene filtering was performed with the function *filter.genes.by.cluster.expression*. The min.max.cluster.average parameters were set for the corresponding matrix as:

- spliced reads matrix: 5

- unspliced reads matrix: 1
- spanning reads matrix: 0.5

Then RNA velocity was estimated using *gene.relative.velocity.estimates*, with the following parameters: fit.quantile = 0.05, deltaT = 1, kCells = 5.
PCA of **Figure 3D-E** were calculated with the function *pca.velocity.plot*.
RNA velocity vectors were projected on an isometric representation of the UMAP (**Figure 3F-G** & **Figure 5G-H**) with the function with the function *show.velocity.on.embedding.cor*. Only the cells from the Petrolopoulos et al. dataset are projected to avoid batch effects during the computation of RNA velocities. The sectors of circular diagrams were picked to face the centroid of different cluster from the centroid of the cluster represented in the diagram.

QUANTIFICATION AND STATISTICAL ANALYSIS

Enrichment analysis.

Module enrichment analysis was performed with FGSEA (Sergushichev, 2016). Ranking metric for f GSEA was set as WGCNA gene module membership (Korotkevich et al., 2019). This score is processed by Pearson correlation of module eigengene and gene expression. Enrichment was made on five databases: Gene ontology (Cellular Component, Molecular Function and Biological Process), Reactom and KEGG. All retained term were enriched below an adjusted Benjamini-Hochberg p-value of 0.05. A list of transcriptional factors (TF) was downloaded from the RegNetwork database (Liu et al., 2015). Pathway eigengene metric (**Figure S4B**) was processed by taking the first component of a principal component analysis of genes from each enriched term.

Methods bibliography.

Ciray, H.N., Campbell, A., Agerholm, I.E., Aguilar, J., Chamayou, S., Esbert, M., Sayed, S., and Time-Lapse User, G. (2014). Proposed guidelines on the nomenclature and annotation of dynamic human embryo monitoring by a time-lapse user group. Human reproduction *29*, 2650-2660.

Gardner, D.K., Lane, M., Stevens, J., Schlenker, T., and Schoolcraft, W.B. (2000). Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. Fertility and sterility 73, 1155-1158.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847-2849.

Kimmelman, J., Heslop, H.E., Sugarman, J., Studer, L., Benvenisty, N., Caulfield, T., Hyun, I., Murry, C.E., Sipp, D., and Daley, G.Q. (2016). New ISSCR guidelines: clinical translation of stem cell research. Lancet 387, 1979-1981.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lonnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature 560, 494-498.

Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., Chy, H.S., et al. (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. Nature methods 14, 1055-1062.

Sergushichev, A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. biorxiv.

Trombetta, J.J., Gennert, D., Lu, D., Satija, R., Shalek, A.K., and Regev, A. (2014). Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. Curr Protoc Mol Biol 107, 4 22 21-17.

Table Legends

Table S1. Antibodies

Details of primary and secondary antibodies used for this study.

Table S2. Cells annotations for human dataset

Annotation for each cell used in the study of human pre-implantation development. Annotations beginning with "Author" are referring to annotation of author's dataset. "Author.lineage": lineage attribution of authors. "Author.TEside": Trophectoderm side attribution from Petropoulos et al. (2016). "Author.pseudoTime": pseudotime of each cell according to Petropoulos et al. (2016). "StirparotEtAl.lineage": lineage attribution of dataset from Petropoulos et al. (2016) by Stirparo et al. (2018).

Table S3. Cell annotations for mouse dataset

Annotation for each cell used in the study of human pre-implantation development. Annotations beginning with "Author.lineage" is referring to the lineage attribution of Posfai et al. (2017).

Table S4. Gene modules

Detailed gene list for each WGCNA modules.

Table S5 Detailed sampling of single-cell RNA-Seq dataset from this study

For each embryo, the embryo ID, sex, number of cells sequenced, stage and grade are indicated.

References

Aberkane, A., Essahib, W., Spits, C., De Paepe, C., Sermon, K., Adriaenssens, T., Mackens, S., Tournaye, H., Brosens, J.J., and Van de Velde, H. (2018). Expression of adhesion and extracellular matrix genes in human blastocysts upon attachment in a 2D co-culture system. Molecular human reproduction *24*, 375-387.

Acampora, D., Omodei, D., Petrosino, G., Garofalo, A., Savarese, M., Nigro, V., Di Giovannantonio, L.G., Mercadante, V., and Simeone, A. (2016). Loss of the Otx2-Binding Site in the Nanog Promoter Affects the Integrity of Embryonic Stem Cell Subtypes and Specification of Inner Cell Mass-Derived Epiblast. Cell reports *15*, 2651-2664.

Ahlstrom, A., Westin, C., Reismer, E., Wikland, M., and Hardarson, T. (2011). Trophectoderm morphology: an important parameter for predicting live birth after single blastocyst transfer. Human reproduction *26*, 3289-3296.

Allègre, N., Chauveau, S., Dennis, C., Renaud, Y., Estrella, L.V., Pouchin, P., Cohen-Tannoudji, M., and Chazaud, C. (2019). A Nanog-dependent gene cluster initiates the specification of the pluripotent epiblast. BioRxiv.

Alpha Scientists in Reproductive, M., and Embryology, E.S.I.G.o. (2011). The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Human reproduction *26*, 1270-1283.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Aplin, J.D., and Ruane, P.T. (2017). Embryo-epithelium interactions during implantation at a glance. Journal of cell science *130*, 15-22.

Barcroft, L.C., Hay-Schmidt, A., Caveney, A., Gilfoyle, E., Overstrom, E.W., Hyttel, P., and Watson, A.J. (1998). Trophectoderm differentiation in the bovine embryo: characterization of a polarized epithelium. Journal of reproduction and fertility *114*, 327-339.

Barcroft, L.C., Offenberg, H., Thomsen, P., and Watson, A.J. (2003). Aquaporin proteins in murine trophectoderm mediate transepithelial water movements during cavitation. Developmental biology *256*, 342-354.

Bentin-Ley, U., Horn, T., Sjogren, A., Sorensen, S., Falck Larsen, J., and Hamberger, L. (2000). Ultrastructure of human blastocyst-endometrial interactions in vitro. Journal of reproduction and fertility *120*, 337-350.

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., 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*.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell stem cell *14*, 838-853.

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 reports *33*, 108419.

Chan, C.J., Costanzo, M., Ruiz-Herrero, T., Monke, G., Petrie, R.J., Bergert, M., Diz-Munoz, A., Mahadevan, L., and Hiiragi, T. (2019). Hydraulic control of mammalian embryo size and cell fate. Nature *571*, 112-116.

Chazaud, C., and Yamanaka, Y. (2016). Lineage specification in the mouse preimplantation embryo. Development *143*, 1063-1074.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Developmental cell *10*, 615-624.
Chen, X., Zhang, J., Wu, X., Cao, S., Zhou, L., Wang, Y., Chen, X., Lu, J., Zhao, C., Chen, M.*, et al.* (2014). Trophectoderm morphology predicts outcomes of pregnancy in vitrified-warmed single-blastocyst transfer cycle in a Chinese population. Journal of assisted reproduction and genetics *31*, 1475-1481.
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 Reports.
De Mot, L., Gonze, D., Bessonnard, S., Chazaud, C., Goldbeter, A., and Dupont, G. (2016). Cell Fate Specification Based on Tristability in the Inner Cell Mass of Mouse Blastocysts. Biophys J *110*, 710-722.
De Paepe, C., Aberkane, A., Dewandre, D., Essahib, W., Sermon, K., Geens, M., Verheyen, G., Tournaye, H., and Van de Velde, H. (2019). BMP4 plays a role in apoptosis during human preimplantation development. Molecular reproduction and development *86*, 53-62.
De Paepe, C., Cauffman, G., Verloes, A., Sterckx, J., Devroey, P., Tournaye, H., Liebaers, I., and Van de Velde, H. (2013). Human trophectoderm cells are not yet committed. Human reproduction *28*, 740-749.
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.
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*.
Dumortier, J.G., Le Verge-Serandour, M., Tortorelli, A.F., Mielke, A., de Plater, L., Turlier, H., and Maitre, J.L. (2019). Hydraulic fracturing and active coarsening position the lumen of the mouse blastocyst. Science *365*, 465-468.
Embryology, E.S.o.H.R.a. (2015). Data collection and research (https://www.eshre.eu/).
Feyeux, M., Reignier, A., Mocaer, M., Lammers, J., Meistermann, D., Barriere, P., Paul-Gilloteaux, P., David, L., and Freour, T. (2020). Development of automated annotation software for human embryo morphokinetics. Human reproduction *35*, 557-564.
Fogarty, N.M.E., McCarthy, A., Snijders, K.E., Powell, B.E., Kubikova, N., Blakeley, P., Lea, R., Elder, K., Wamaitha, S.E., Kim, D.*, et al.* (2017). Genome editing reveals a role for OCT4 in human embryogenesis. Nature *550*, 67-73.
Frum, T., Murphy, T.M., and Ralston, A. (2018). HIPPO signaling resolves embryonic cell fate conflicts during establishment of pluripotency in vivo. Elife *7*.
Gardner, D.K., Lane, M., Stevens, J., Schlenker, T., and Schoolcraft, W.B. (2000). Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. Fertility and sterility *73*, 1155-1158.
Giacomini, G., Tabibzadeh, S.S., Satyaswaroop, P.G., Bonsi, L., Vitale, L., Bagnara, G.P., Strippoli, P., and Jasonni, V.M. (1995). Epithelial cells are the major source of biologically active granulocyte macrophage colony-stimulating factor in human endometrium. Human reproduction *10*, 3259-3263.
Giocanti-Auregan, A., Vacca, O., Benard, R., Cao, S., Siqueiros, L., Montanez, C., Paques, M., Sahel, J.A., Sennlaub, F., Guillonneau, X.*, et al.* (2016). Altered astrocyte morphology and vascular development in dystrophin-Dp71-null mice. Glia *64*, 716-729.
Grabarek, J.B., Zyzynska, K., Saiz, N., Piliszek, A., Frankenberg, S., Nichols, J., Hadjantonakis, A.K., and Plusa, B. (2012). Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. Development *139*, 129-139.
Graham, S.J., and Zernicka-Goetz, M. (2016). The Acquisition of Cell Fate in Mouse Development: How Do Cells First Become Heterogeneous? Current topics in developmental biology *117*, 671-695.
Greenberg, M.V.C., and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. Nature reviews Molecular cell biology *20*, 590-607.

Grewal, S., Carver, J.G., Ridley, A.J., and Mardon, H.J. (2008). Implantation of the human embryo requires Rac1-dependent endometrial stromal cell migration. Proceedings of the National Academy of Sciences of the United States of America *105*, 16189-16194.
Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847-2849.
Guo, G., Stirparo, G.G., Strawbridge, S., Yang, J., Clarke, J., Li, M.A., Myers, S., Özel, B.N., Nichols, J., and Smith, A. (2020).
Guzman-Ayala, M., Ben-Haim, N., Beck, S., and Constam, D.B. (2004). Nodal protein processing and fibroblast growth factor 4 synergize to maintain a trophoblast stem cell microenvironment. Proceedings of the National Academy of Sciences of the United States of America *101*, 15656-15660.
Handyside, A.H. (1978). Time of commitment of inside cells isolated from preimplantation mouse embryos. J Embryol Exp Morphol *45*, 37-53.
Hannan, N.J., Paiva, P., Dimitriadis, E., and Salamonsen, L.A. (2010). Models for study of human embryo implantation: choice of cell lines? Biology of reproduction *82*, 235-245.
Hill, M.J., Richter, K.S., Heitmann, R.J., Graham, J.R., Tucker, M.J., DeCherney, A.H., Browne, P.E., and Levens, E.D. (2013). Trophectoderm grade predicts outcomes of single-blastocyst transfers. Fertility and sterility *99*, 1283-1289 e1281.
Hirate, Y., Hirahara, S., Inoue, K., Suzuki, A., Alarcon, V.B., Akimoto, K., Hirai, T., Hara, T., Adachi, M., Chida, K*., et al.* (2013). Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos. Current biology : CB *23*, 1181-1194.
Jedrusik, A., Parfitt, D.E., Guo, G., Skamagki, M., Grabarek, J.B., Johnson, M.H., Robson, P., and Zernicka-Goetz, M. (2008). Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. Genes & development *22*, 2692-2706.
Kilens, S., Meistermann, D., Moreno, D., Chariau, C., Gaignerie, A., Reignier, A., Lelievre, Y., Casanova, M., Vallot, C., Nedellec, S*., et al.* (2018). Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. Nat Commun *9*, 360.
Kimmelman, J., Heslop, H.E., Sugarman, J., Studer, L., Benvenisty, N., Caulfield, T., Hyun, I., Murry, C.E., Sipp, D., and Daley, G.Q. (2016). New ISSCR guidelines: clinical translation of stem cell research. Lancet *387*, 1979-1981.
Korotkevich, G., Sukhov, V., and Sergushichev, A. (2019). Fast gene set enrichment analysis. BioRxiv.
Kuijk, E.W., van Tol, L.T., Van de Velde, H., Wubbolts, R., Welling, M., Geijsen, N., and Roelen, B.A. (2012). The roles of FGF and MAP kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos. Development *139*, 871-882.
La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lonnerberg, P., Furlan, A*., et al.* (2018). RNA velocity of single cells. Nature *560*, 494-498.
Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics *9*, 559.
Leland McInnes, J.H., James Melville (2018). UMAP: Uniform Manifold

Approximation and Projection for

Dimension Reduction. arXiv.
Lindenberg, S. (1991). Experimental studies on the initial trophoblast endometrial interaction. Dan Med Bull *38*, 371-380.
Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., Chy, H.S*., et al.* (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. Nature methods *14*, 1055-1062.
Liu, Z.P., Wu, C., Miao, H., and Wu, H. (2015). RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. Database : the journal of biological databases and curation *2015*.
Lun, A.T., McCarthy, D.J., and Marioni, J.C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. F1000Res *5*, 2122.

Martin Esterk, H.-P.K., Jörg Sander, Xiaowei Xu, and (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, 226-231.
Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. Developmental biology *375*, 54-64.
Petropoulos, S., Edsgard, 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.
Posfai, E., Petropoulos, S., de Barros, F.R.O., Schell, J.P., Jurisica, I., Sandberg, R., Lanner, F., and Rossant, J. (2017). Position- and Hippo signaling-dependent plasticity during lineage segregation in the early mouse embryo. Elife *6*.
Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nature methods *14*, 979-982.
Ramos-Ibeas, P., Sang, F., Zhu, Q., Tang, W.W.C., Withey, S., Klisch, D., Wood, L., Loose, M., Surani, M.A., and Alberio, R. (2019). Pluripotency and X chromosome dynamics revealed in pig pre-gastrulating embryos by single cell analysis. Nat Commun *10*, 500.
Reijo Pera, R.A., and Prezzoto, L. (2016). Species-Specific Variation Among Mammals. Current topics in developmental biology *120*, 401-420.
Rivron, N.C., Frias-Aldeguer, J., Vrij, E.J., Boisset, J.C., Korving, J., Vivie, J., Truckenmuller, 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. Developmental biology *361*, 358-363.
Rossant, J., and Lis, W.T. (1979). Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. Developmental biology *70*, 255-261.
Roy, A., Ghosh, A., Kumar, B., and Chandran, B. (2019). IFI16, a nuclear innate immune DNA sensor, mediates epigenetic silencing of herpesvirus genomes by its association with H3K9 methyltransferases SUV39H1 and GLP. Elife *8*.
Sergushichev, A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. biorxiv.
Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.M., Campbell, A., Devito, L.G., Ilic, D.*, et al.* (2016). Self-organization of the human embryo in the absence of maternal tissues. Nature cell biology *18*, 700-708.
Shahbazi, M.N., Scialdone, A., Skorupska, N., Weberling, A., Recher, G., Zhu, M., Jedrusik, A., Devito, L.G., Noli, L., Macaulay, I.C.*, et al.* (2017). Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. Nature *552*, 239-243.
Shoukir, Y., Chardonnens, D., Campana, A., Bischof, P., and Sakkas, D. (1998). The rate of development and time of transfer play different roles in influencing the viability of human blastocysts. Human reproduction *13*, 676-681.
Singh, M., Widmann, T.J., Bansal, V., Cortes, J.L., Schumann, G.G., Wunderlich, S., Martin, U., Garcia-Perez, J.L., Hurst, L.D., and Izsvak, Z. (2019).
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*.
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.
technology, E.W.g.o.T.-l., Apter, S., Ebner, T., Freour, T., Guns, Y., Kovacic, B., Le Clef, N., Marques, M., Meseguer, M., Montjean, D.*, et al.* (2020). Good practice recommendations for the use of time-lapse technology(dagger). Hum Reprod Open *2020*, hoaa008.

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.

Theunissen, Thorold W., Powell, Benjamin E., Wang, H., Mitalipova, M., Faddah, Dina A., Reddy, J., Fan, Zi 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.

Thompson, S.M., Onwubalili, N., Brown, K., Jindal, S.K., and McGovern, P.G. (2013). Blastocyst expansion score and trophectoderm morphology strongly predict successful clinical pregnancy and live birth following elective single embryo blastocyst transfer (eSET): a national study. Journal of assisted reproduction and genetics *30*, 1577-1581.

Tiegs, A.W., Sun, L., Patounakis, G., and Scott, R.T. (2019). Worth the wait? Day 7 blastocysts have lower euploidy rates but similar sustained implantation rates as Day 5 and Day 6 blastocysts. Human reproduction *34*, 1632-1639.

Trombetta, J.J., Gennert, D., Lu, D., Satija, R., Shalek, A.K., and Regev, A. (2014). Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. Curr Protoc Mol Biol *107*, 4 22 21-17.

White, M.D., Zenker, J., Bissiere, S., and Plachta, N. (2018). Instructions for Assembling the Early Mammalian Embryo. Developmental cell *45*, 667-679.

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.

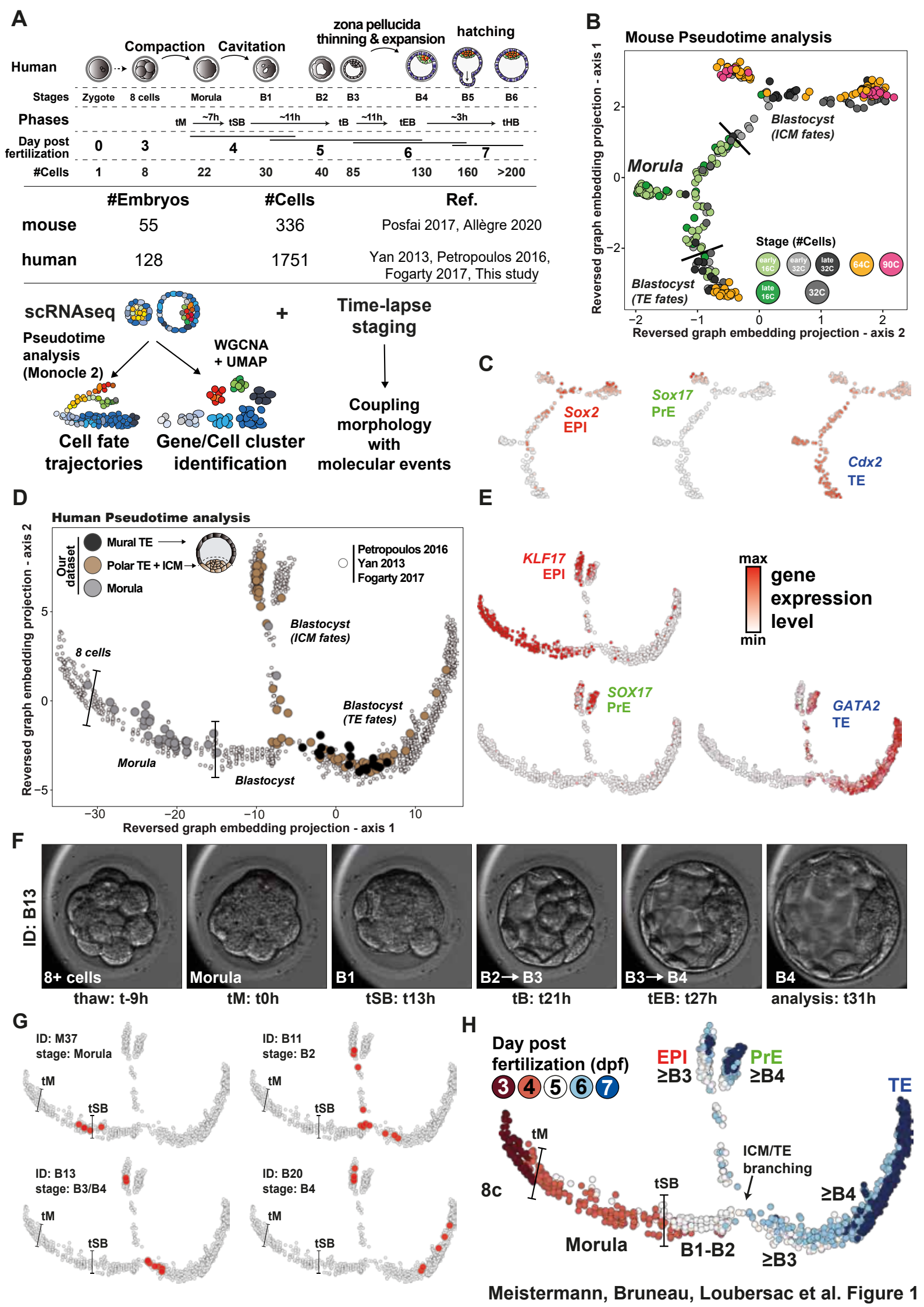
Xiong, Y., Tan, Y.J., Xiong, Y.M., Huang, Y.T., Hu, X.L., Lu, Y.C., Ye, Y.H., Wang, T.T., Zhang, D., Jin, F.*, et al.* (2013). Expression of aquaporins in human embryos and potential role of AQP3 and AQP7 in preimplantation mouse embryo development. Cell Physiol Biochem *31*, 649-658.

Yamagata, T., and Yamazaki, K. (1991). Implanting mouse embryo stain with a LNF-I bearing fluorescent probe at their mural trophectodermal side. Biochemical and biophysical research communications *181*, 1004-1009.

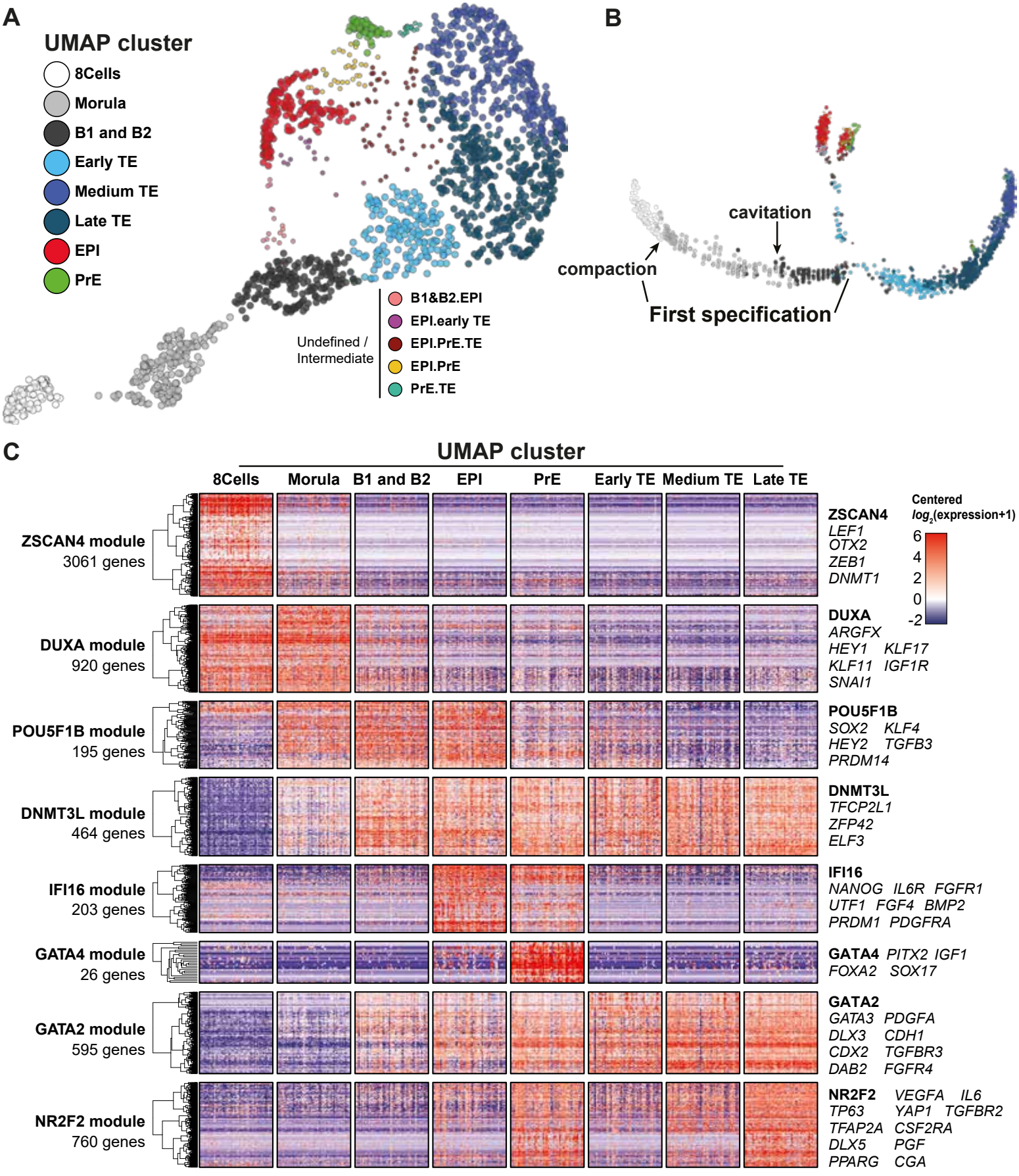
Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J.*, et al.* (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nature structural & molecular biology *20*, 1131-1139.

Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., Lian, Y., Li, J., Wen, L., Yan, L.*, et al.* (2019). Reconstituting the transcriptome and DNA methylome landscapes of human implantation. Nature *572*, 660-664.

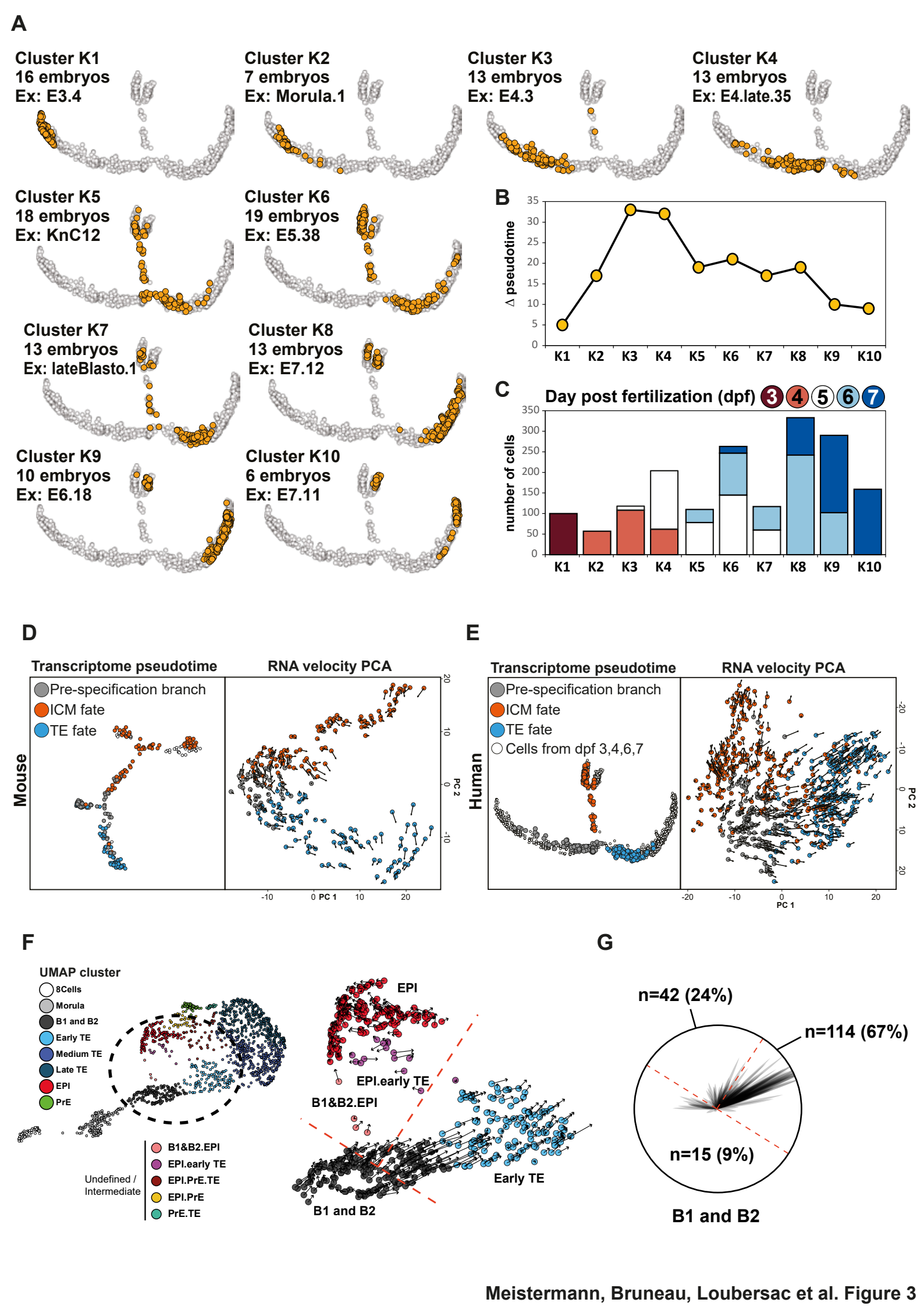
Ziebe, S., Loft, A., Povlsen, B.B., Erb, K., Agerholm, I., Aasted, M., Gabrielsen, A., Hnida, C., Zobel, D.P., Munding, B.*, et al.* (2013). A randomized clinical trial to evaluate the effect of granulocyte-macrophage colony-stimulating factor (GM-CSF) in embryo culture medium for in vitro fertilization. Fertility and sterility *99*, 1600-1609.



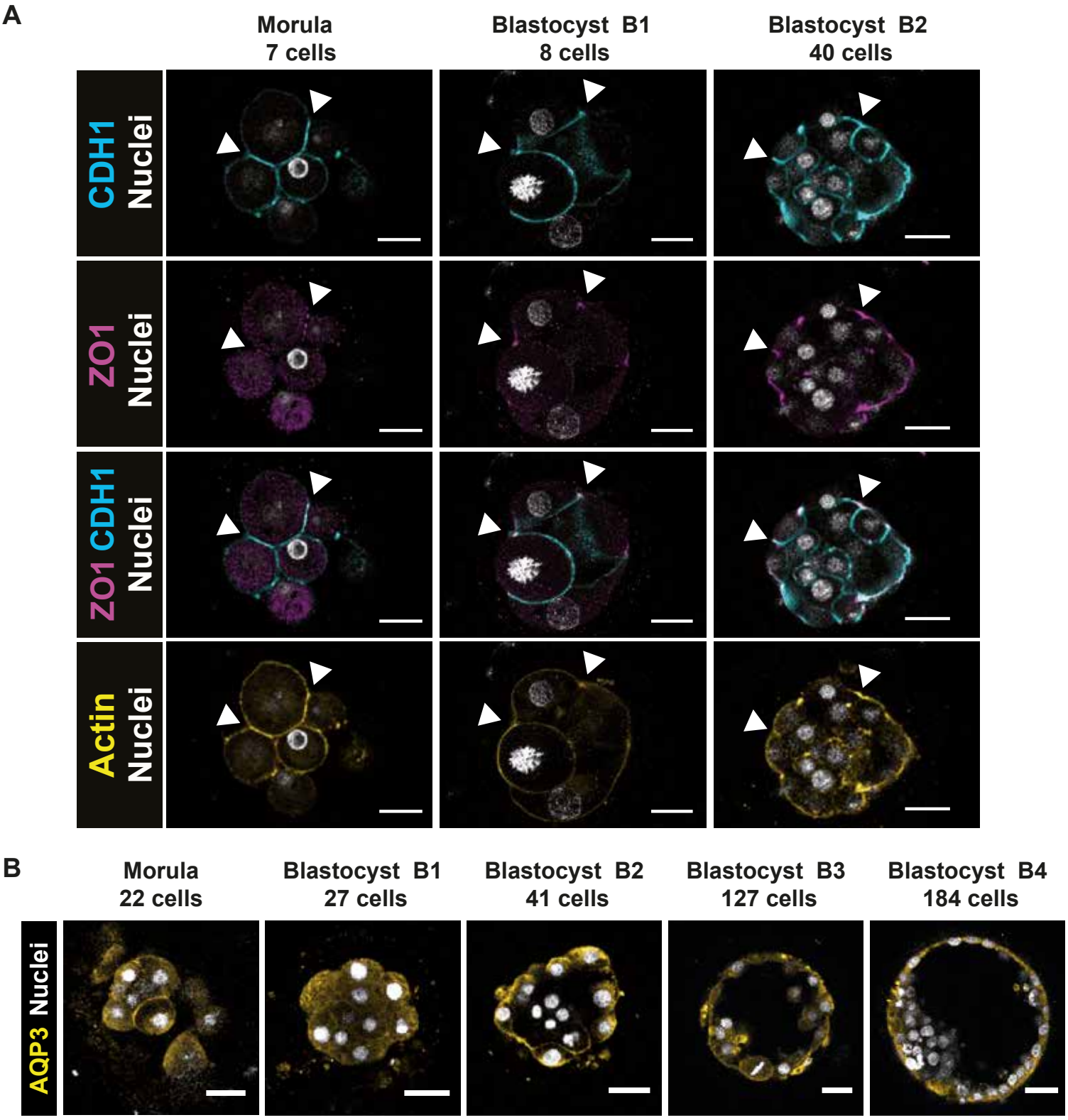
Meistermann, Bruneau, Loubersac et al. Figure 1



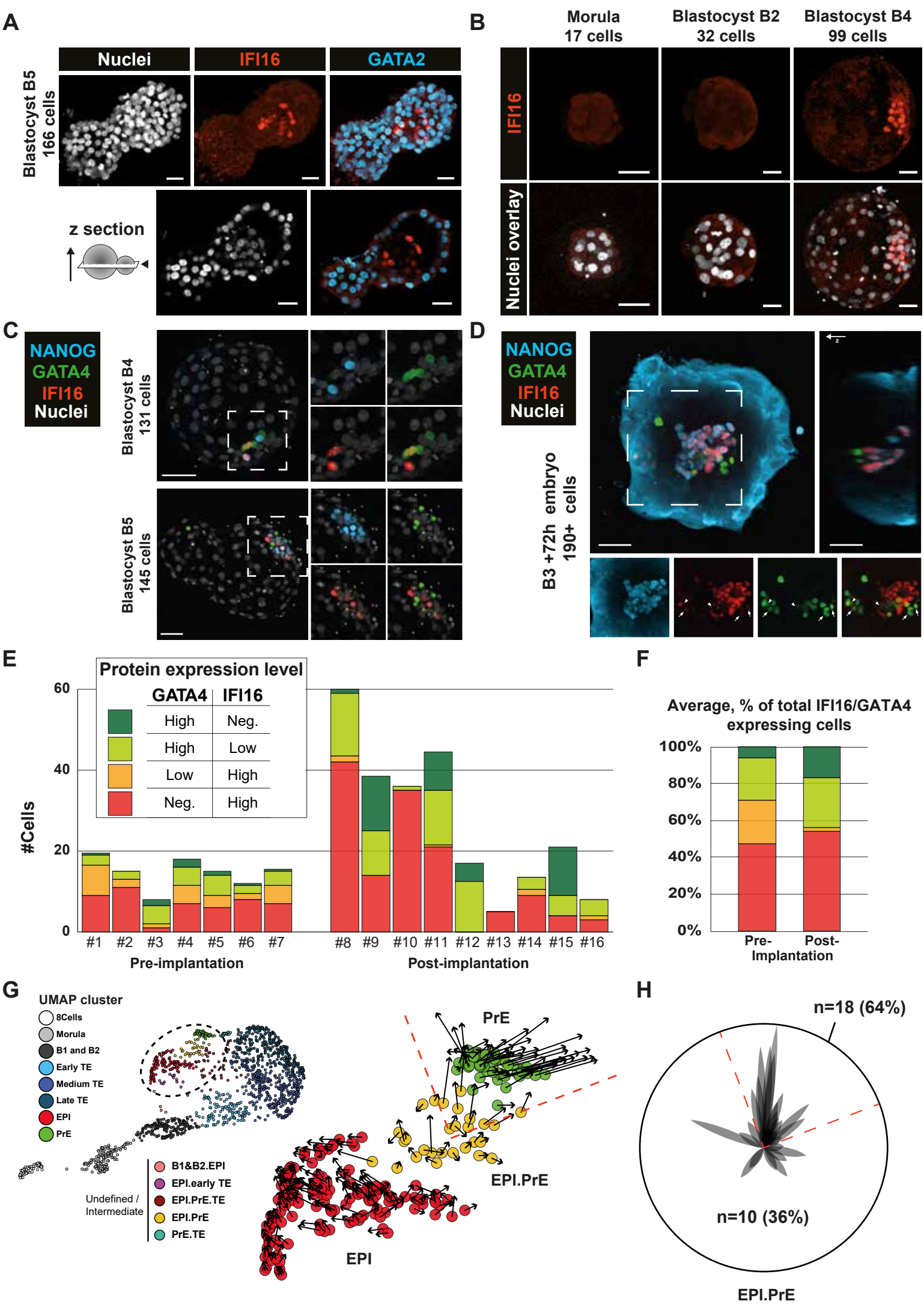
Meistermann, Bruneau, Loubersac et al. Figure 2



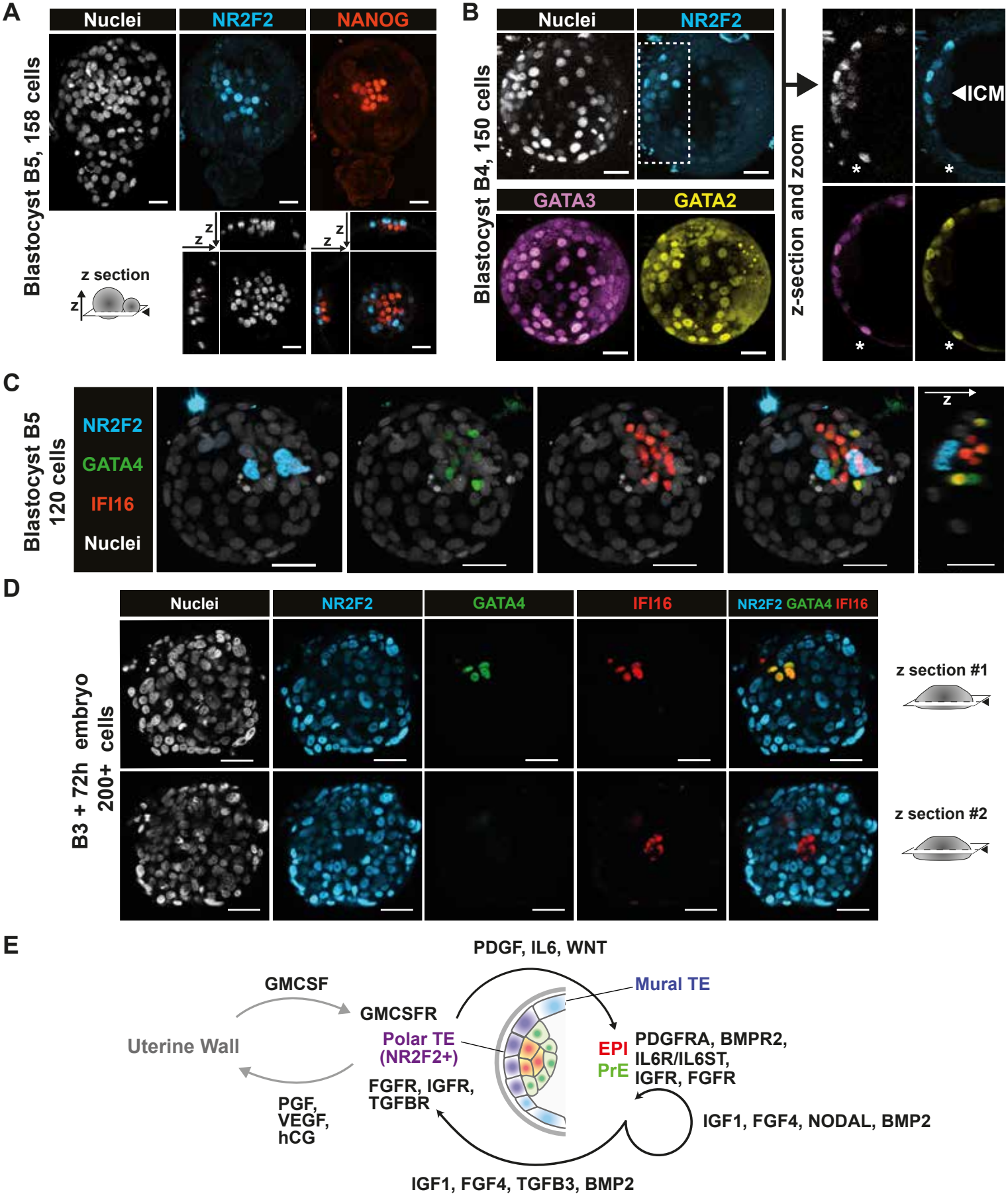
Meistermann, Bruneau, Loubersac et al. Figure 3



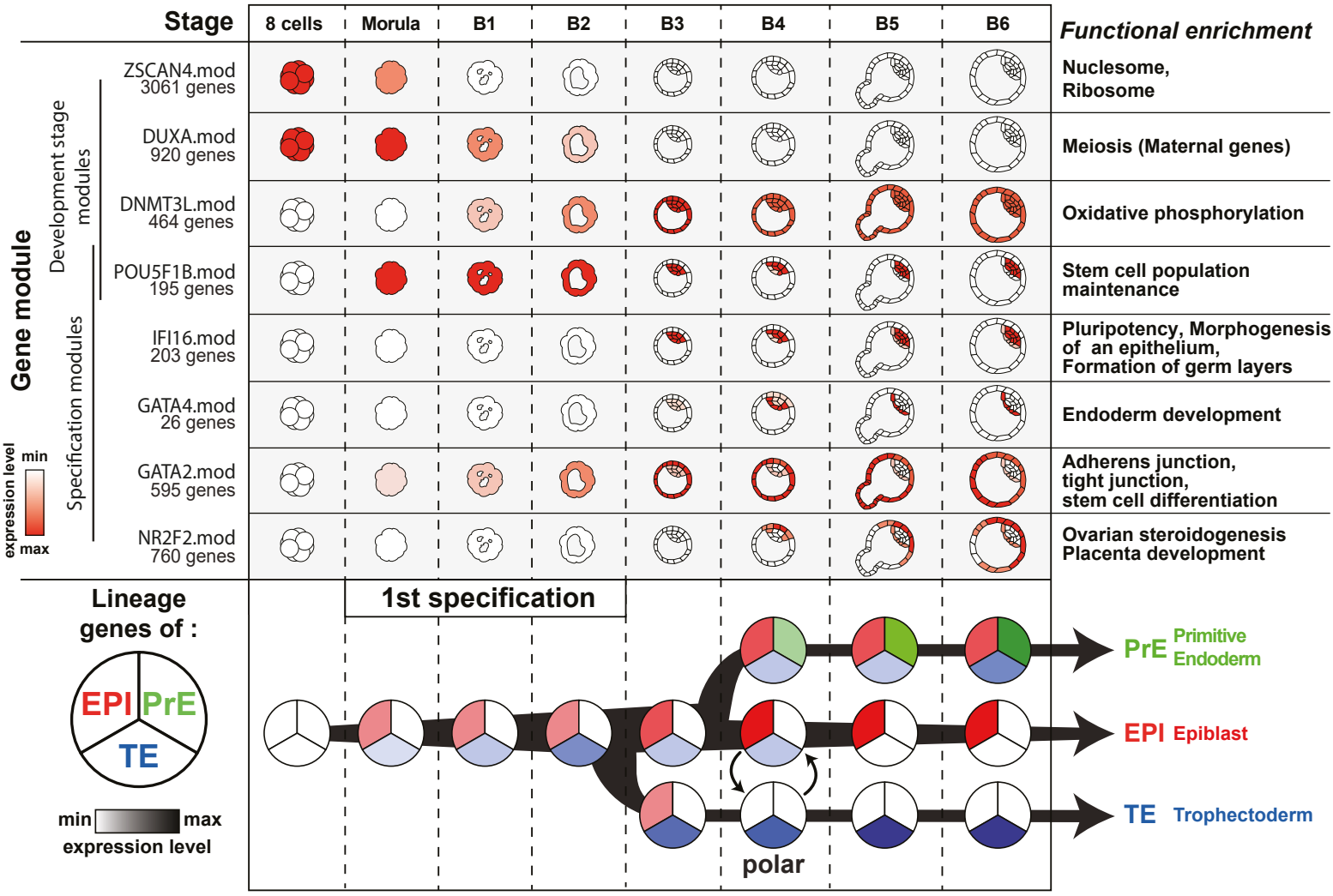
Meistermann, Bruneau, Loubersac et al. Figure 4



Meistermann, Bruneau, Loubersac et al. Figure 5

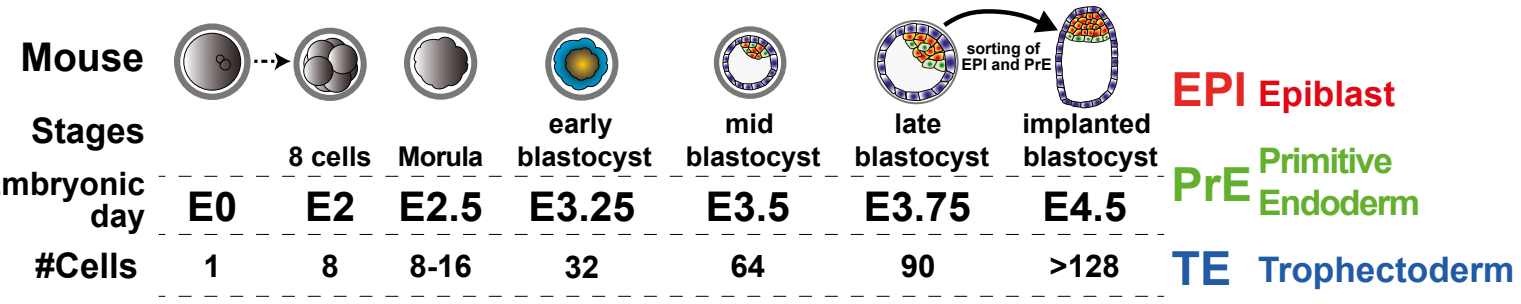


Meistermann, Bruneau, Loubersac et al. Figure 6

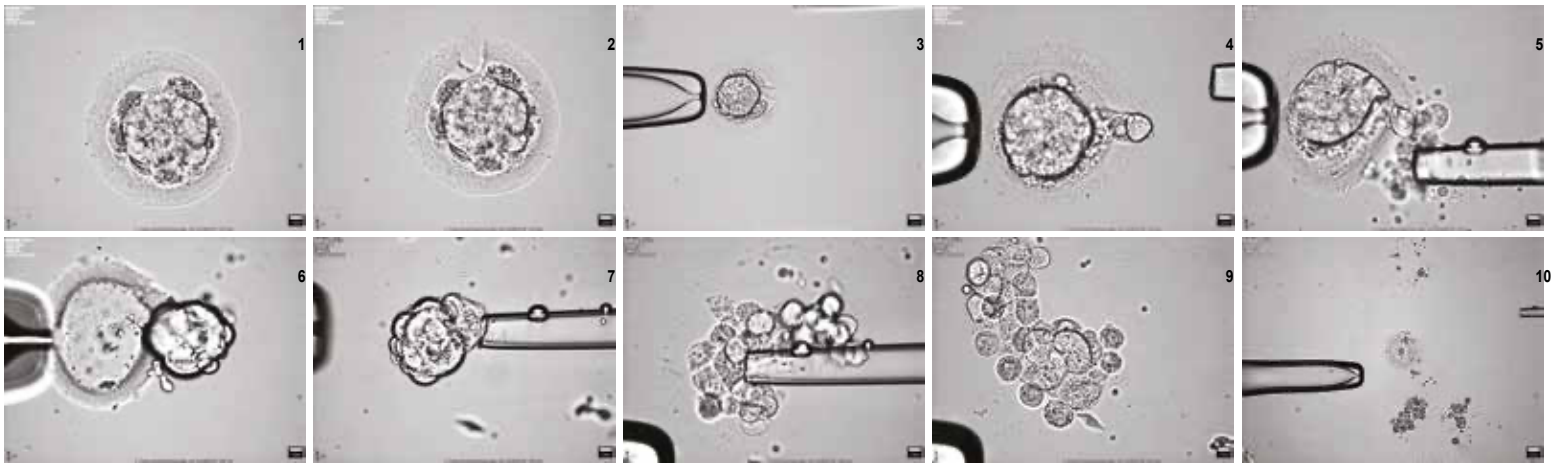


Meistermann, Bruneau, Loubersac et al. Figure 7

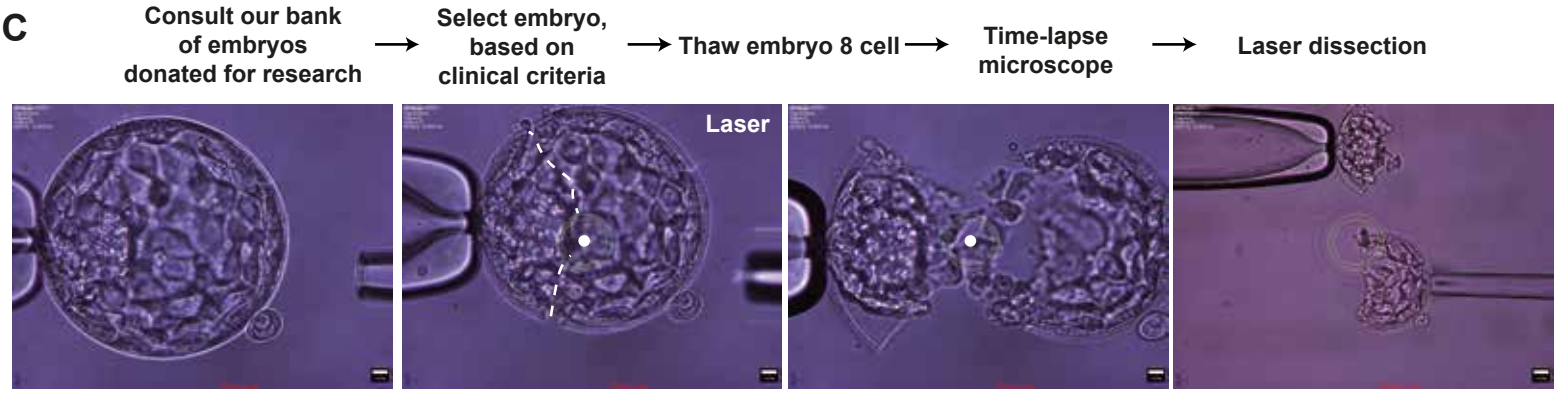
A



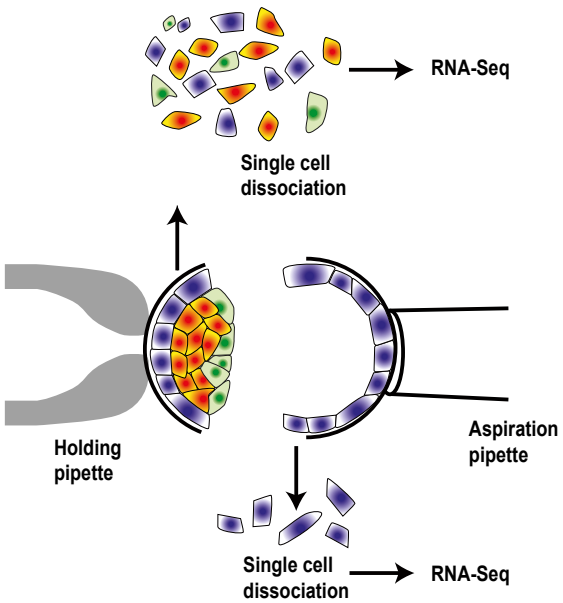
B



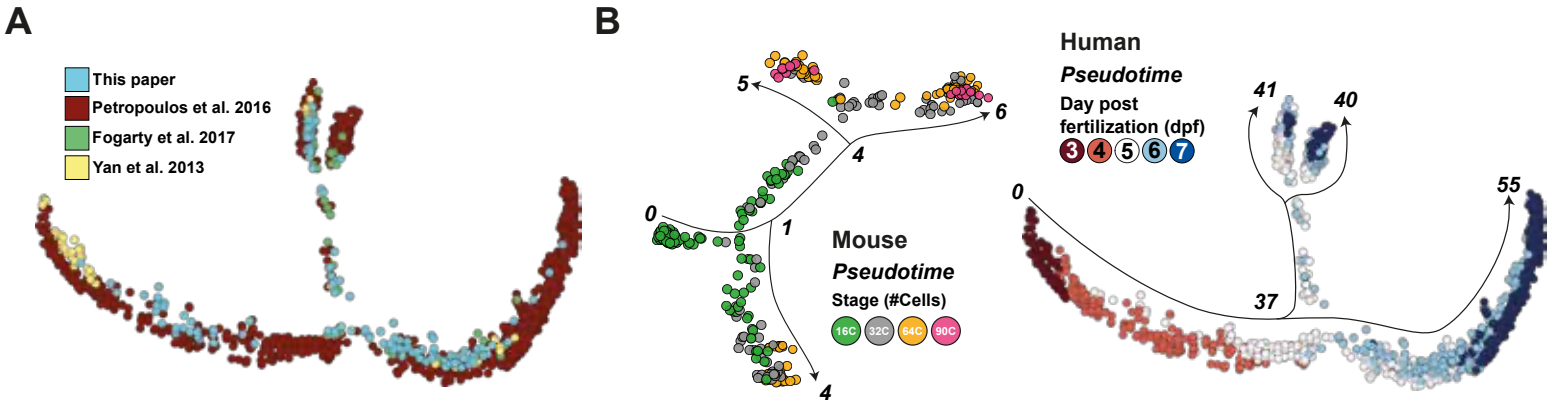
C

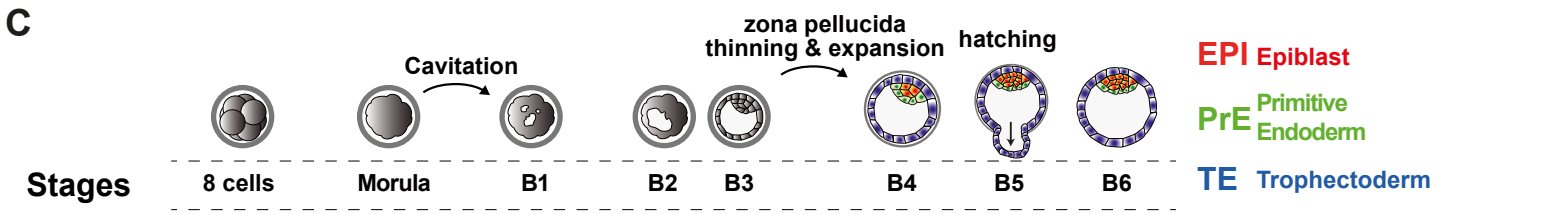


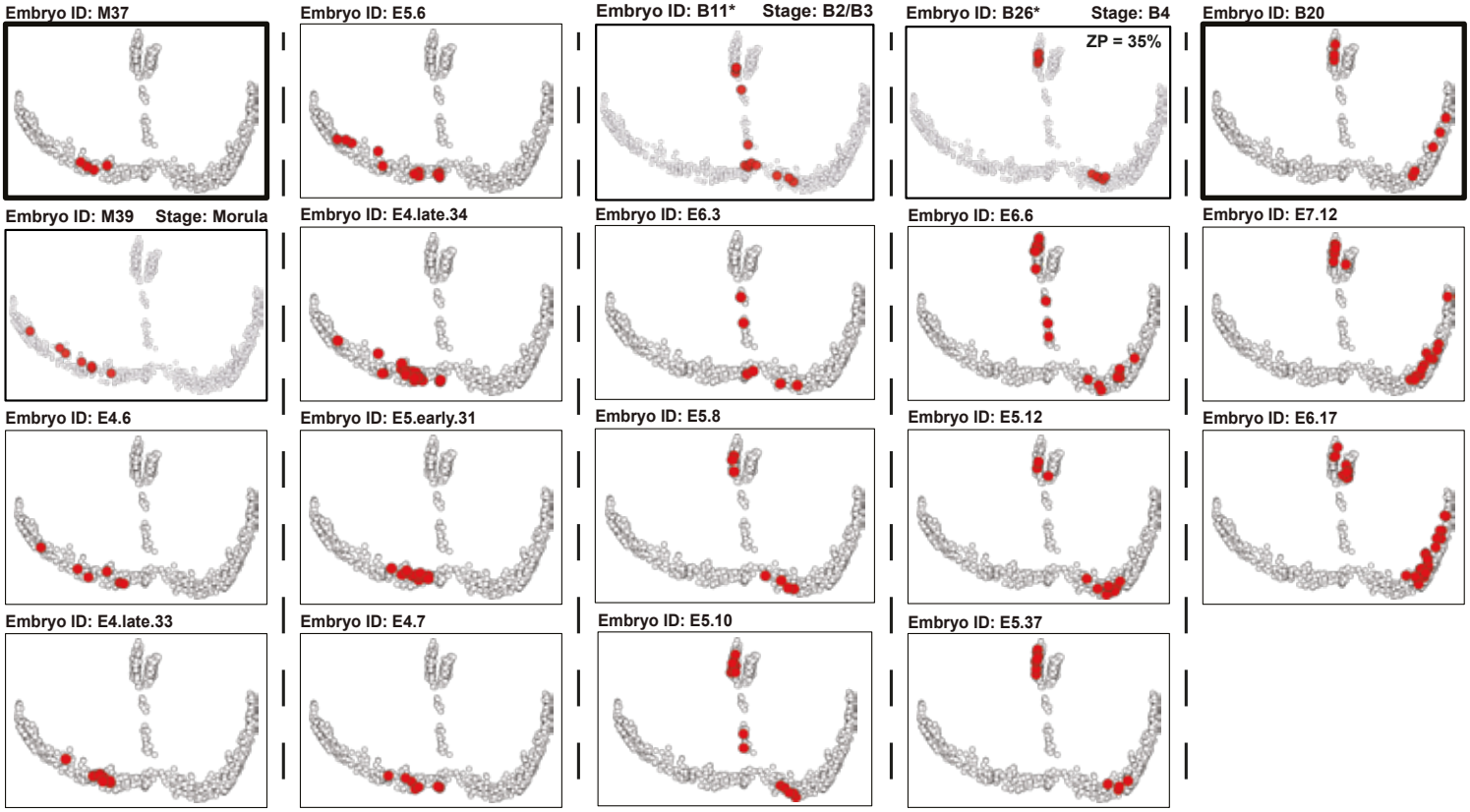
D



Meistermann, Bruneau, Loubersac et al. Suplementary Figure 1







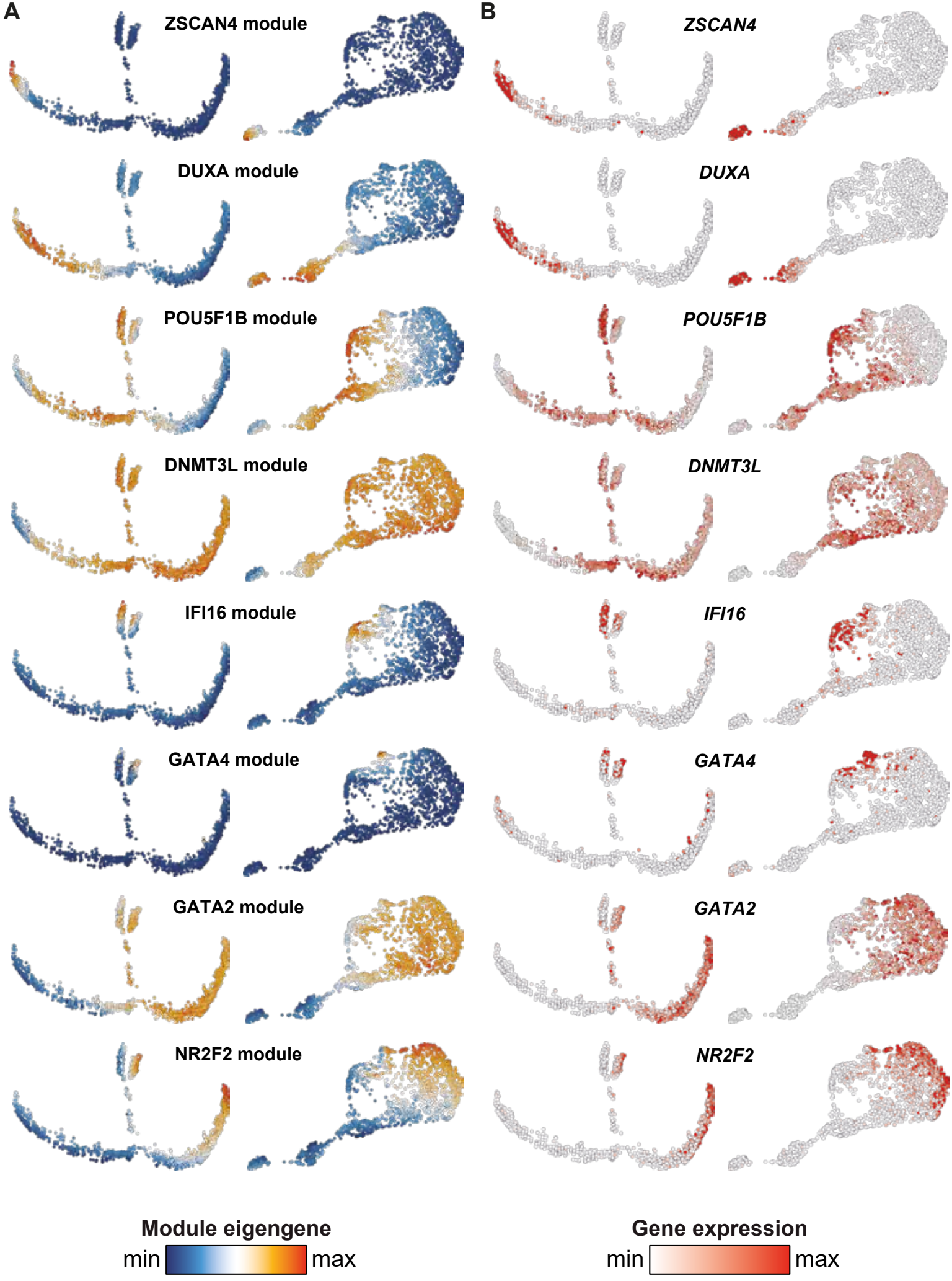
M34	Morula
M57	
M39	
M37	
M42	
M30	
M31	

ID	Stage	ZP thickness
B11	B2/B3	69%
B41*	B3/4	72%
B28*	B4	70%
B22	B3	N.A.
B38	B4	55%
B13*	B4	43%
B50*	B4	46%

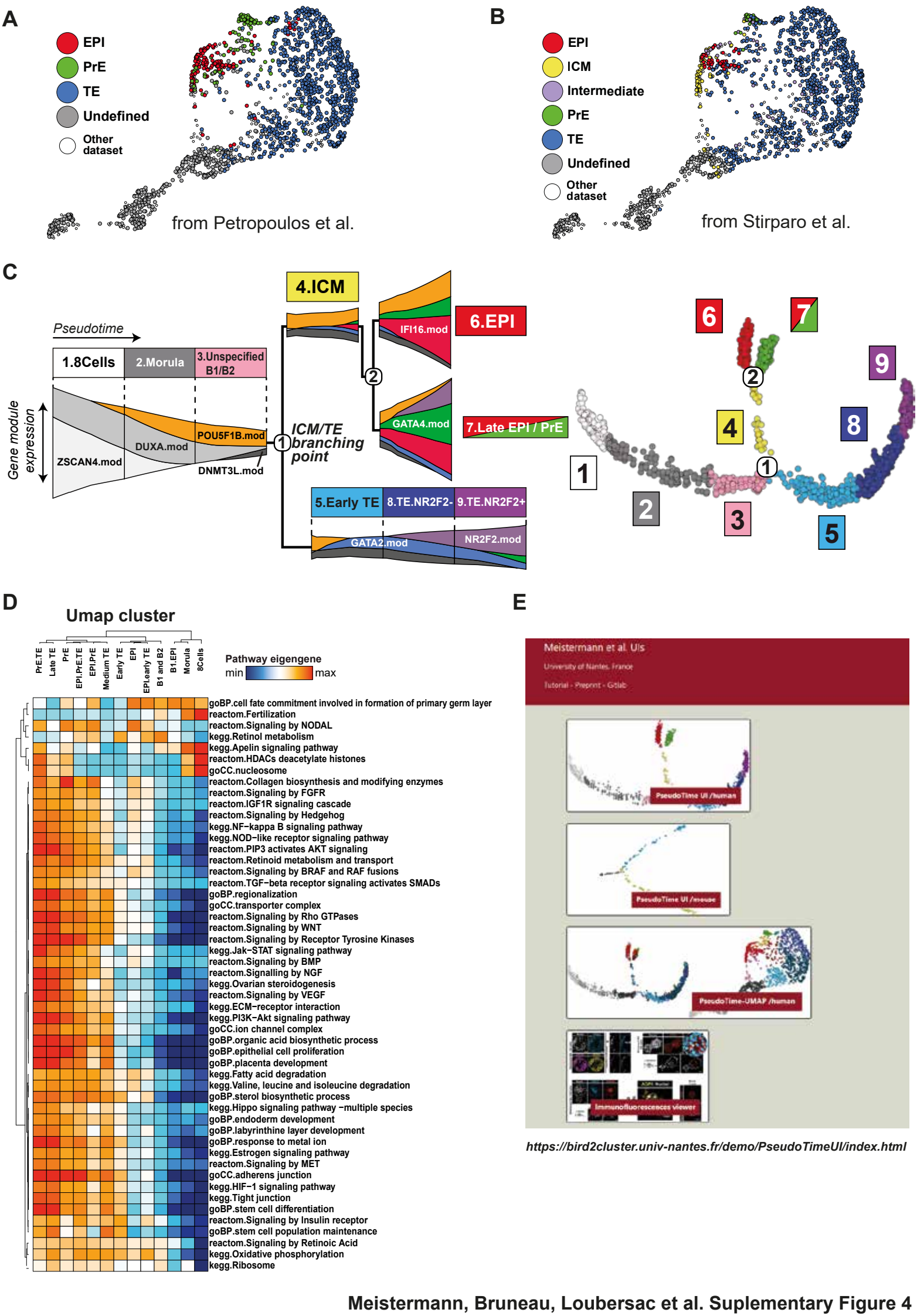
ID	Stage	ZP thickness
B52*	B4	40%
B16*	B4	36%
B24*	B4	37%
B49*	B4	38%
B26*	B4	35%
B18*	B4	36%
B48*	B4	35%

ID	Stage	ZP thickness
B53*	B5	22%
B20*	B5	0%
B56*	B4	19%

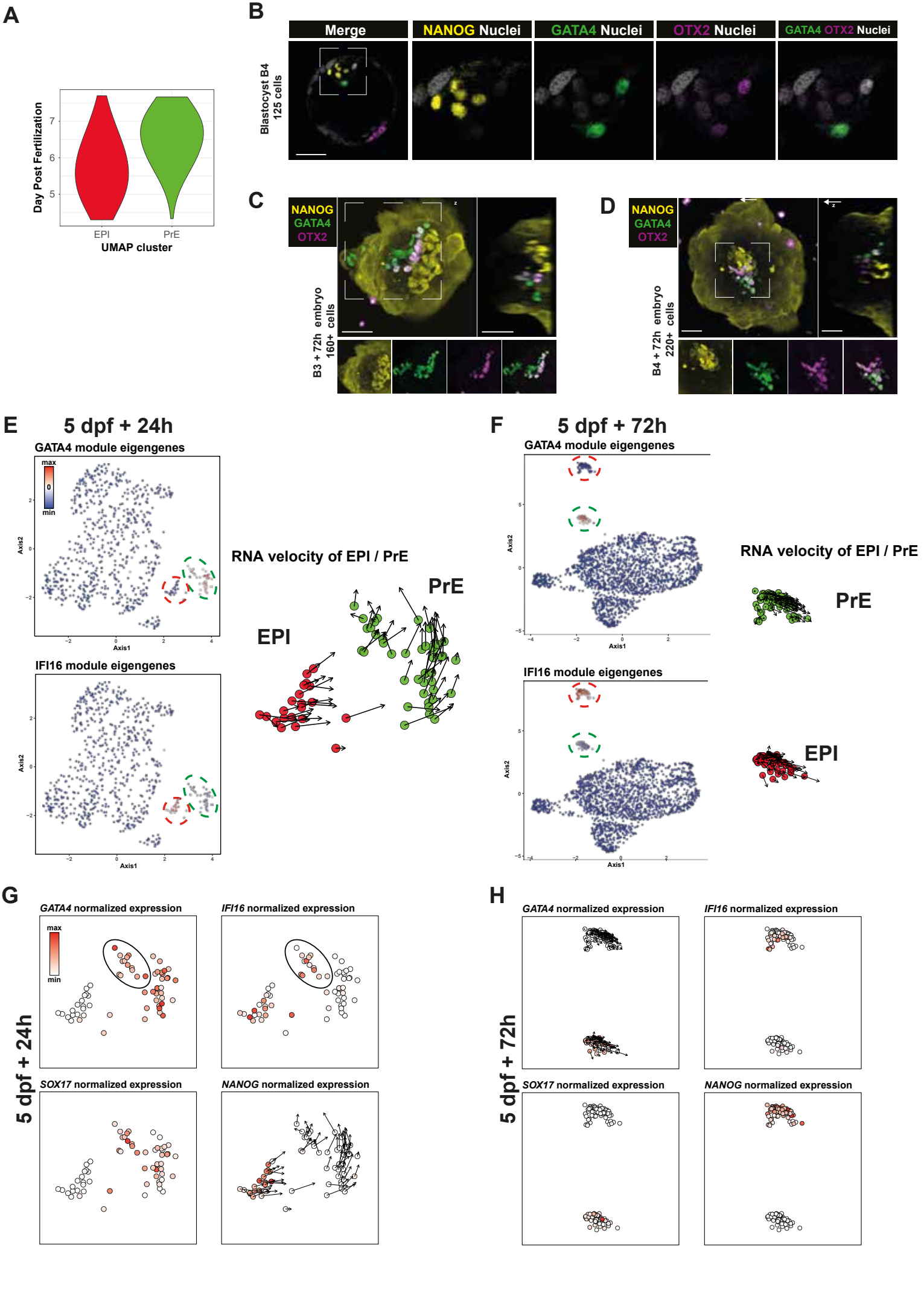
Meistermann, Bruneau, Loubersac et al. Suplementary Figure 2



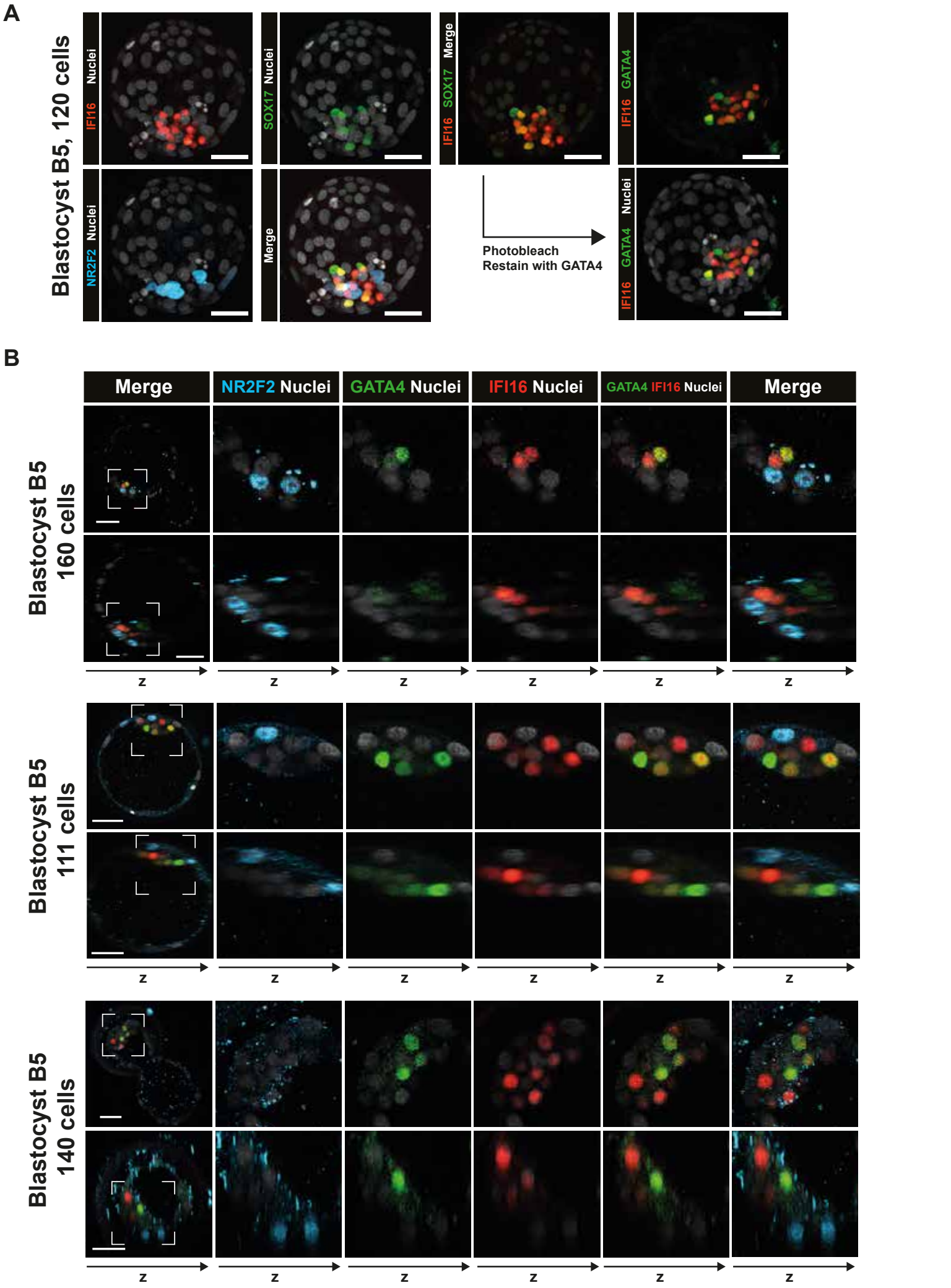
Meistermann, Bruneau, Loubersac et al. Suplementary Figure 3



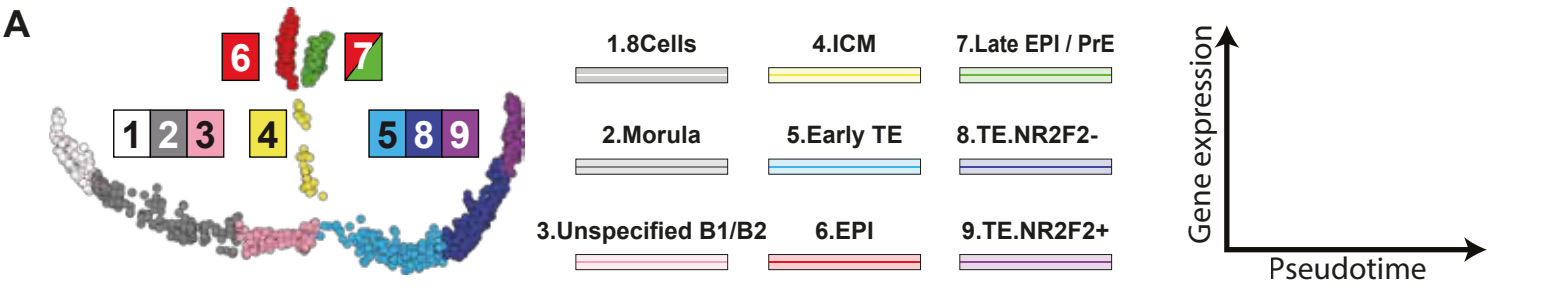
Meistermann, Bruneau, Loubersac et al. Suplementary Figure 4

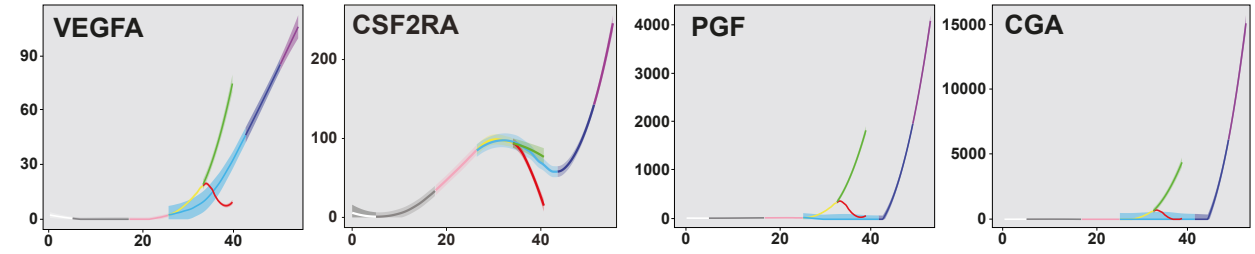


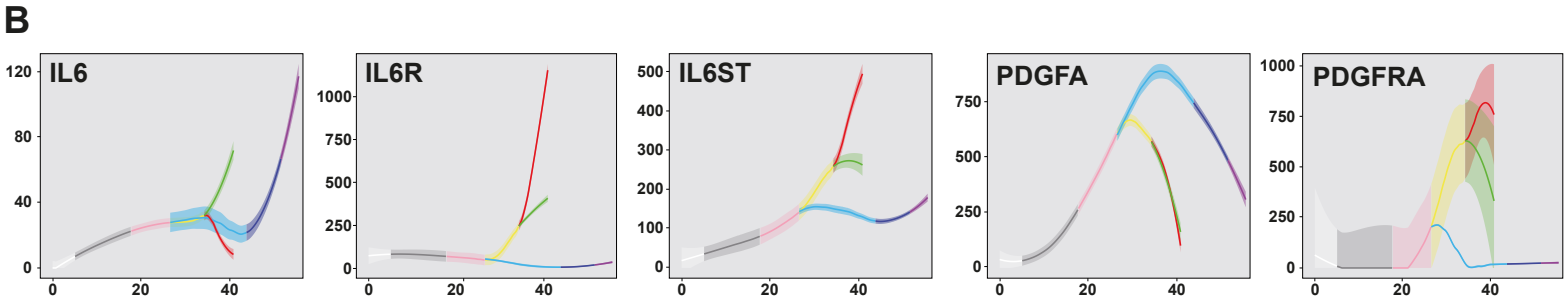
Meistermann, Bruneau, Loubersac et al. Suplementary Figure 5

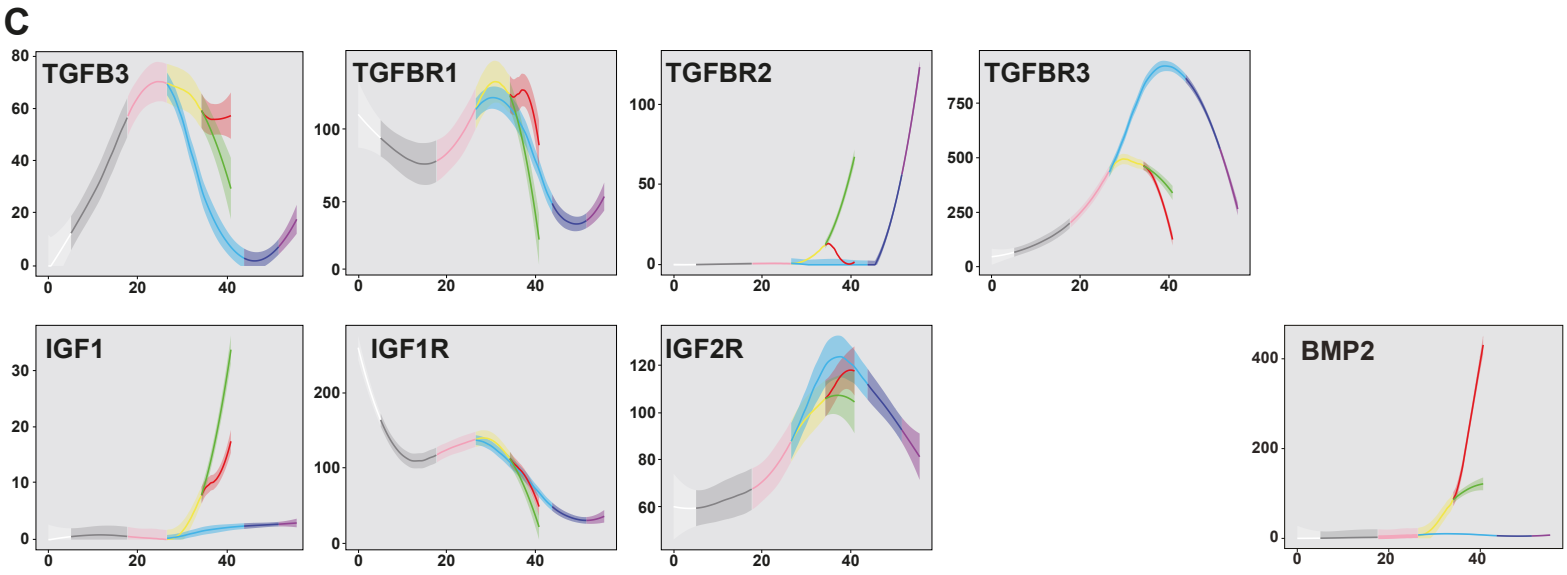


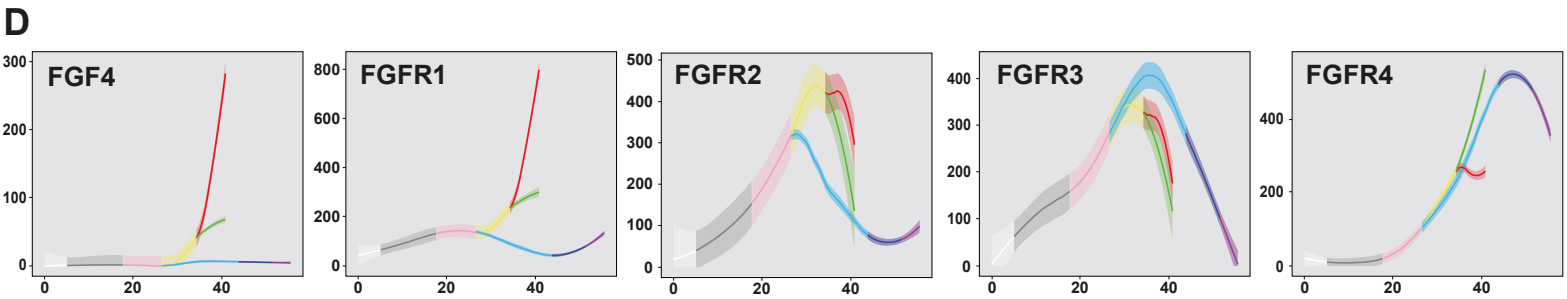
Meistermann, Bruneau, Loubersac et al. Suplementary Figure 6











Meistermann, Bruneau, Loubersac et al. Suplementary Figure 7