Étude de la spécification des lignées cellulaires lors du développement péri-implantatoire humain

Alexandre Bruneau

DOCTORAT BIOLOGIE BRETAGNE SANTE LOIRE

Nantes Université

THESE DE DOCTORAT DE

NANTES UNIVERSITE

ECOLE DOCTORALE N° 605
Biologie Santé
Spécialité : Biologie Cellulaire, Biologie du Développement

Par
Alexandre Bruneau

Etude de la spécification des lignées cellulaires lors du développement péri-implantatoire humain

Thèse présentée et soutenue à Nantes, le 13 décembre 2022
Unité de recherche : CR2Ti UMR 1064

Rapporteurs avant soutenance :

Michel COHEN-TANNOUDJI	Directeur de recherche, Institut Pasteur Paris
Catherine PATRAT	Professeure des Universités - Praticien Hospitalier, APHP Paris

Composition du Jury :

Président :	Claire CHAZAUD	Directrice de recherche, Faculté de médecine, Clermont Ferrand
Examinateurs :	Michel COHEN-TANNOUDJI	Directeur de recherche, Institut Pasteur Paris
	Pablo NAVARRO	Directeur de recherche, Institut Pasteur Paris
	Catherine PATRAT	Professeure des Universités - Praticien Hospitalier, APHP Paris
	Hilde VAN DE VELDE	Professeure des Universités - Praticien Hospitalier, Vrije Université de Bruxelles
Dir. de thèse :	Laurent DAVID	Maître de Conférence des Universités - Praticien Hospitalier, Nantes Université

Remerciements

Je remercie chaleureusement Michel Cohen-Tannoudji et Catherine Patrat d'avoir accepté de rapporter ma thèse, ainsi que Claire Chazaud et Pablo Navarro d'avoir accepté de l'examiner. Je mesure la chance que j'ai de pouvoir vous présenter mes travaux.

Je remercie mon directeur de thèse Laurent David de m'avoir accompagné durant cette aventure. Grâce à toi j'ai pu prétendre à une thèse riche, engageante et pluridisciplinaire. Merci d'avoir cru en moi dans les hauts et les bas. Avoir pu évoluer sous ta direction fut un immense privilège.

Je remercie ma co-superviseuse de thèse Hilde Van de Velde. Merci pour ta bienveillance et tes conseils au long de ma thèse. Merci de m'avoir formé et permis de travailler avec toi.

Je remercie mon laboratoire d'accueil, le CR2Ti. Je remercie Régis Josien son directeur ainsi que Carole Guillonneau, Ignacio Anegon, Jérôme Jullien, et Matthieu Giraud pour leur direction de l'Equipe 2. Merci pour votre écoute et vos questions lors des réunions.

Je remercie l'équipe du service d'Assitance Médicale à la Reproduction du CHU de Nantes. Thomas Freour, Sophie Loubersac, Jenna Lammers, Arnaud Reigier, un immense merci de m'avoir accueilli et formé. Merci d'avoir été patients et pédagogues. Merci de m'avoir permis de faire des choses extraordinaires à vos côtés.

Je remercie Djemil Bencharif et Lamia Amirat-Briand. Merci d'avoir suivi mes travaux, et merci pour ces journées travaux pratiques un peu hors du commun.

Je remercie ma famille. Maman, Papa, mon frère Victor. Merci de m'avoir donné ce cadre de vie si parfait, et de m'avoir soutenu depuis toujours dans mes choix. Merci d'avoir été là et de continuer à l'être. Sans vous, je n'aurais pas pu y arriver.

Je remercie la plateforme iPSCDTC de m'avoir formé et accueilli. Merci à Anne, Caroline, Quentin, Lisa, Isabelle, Delphine, Elsa, Célia, Coline pour votre gentillesse et votre écoute.

Je remercie la plateforme d'imagerie MicroPICell de m'avoir formé à la microscopie confocale. Merci à Perrine Paul-Guilloteaux, Steven Nedellec, Philippe Hulin, Magalie Feyeux pour vos conseils et vos astuces. Merci de rendre visible l'invisible, merci pour ces clichés époustouflants qui resteront, merci de m'avoir ouvert à d'autres horizons.

Je remercie Gaël et Dimitri, premiers membres du bureau des doctorants. Vous avez été une immense source d'inspiration pour moi. Merci pour vos discussions, vos suggestions. Merci pour les soirées, et les cigares. Je vous souhaite le meilleur.

Je remercie Constance, Eva, Océane, Nathan, Simon, dignes occupants du bureau des doctorants. Avec vous la relève est sereinement assurée. Merci pour cet environnement de travail riche et divertissant. Merci à Céline et Florian, loin des yeux près du cœur.

Je remercie Sarah et Arthur. Vous êtes pour moi comme des frères et sœurs. Je vous ai toujours admirés, je suis impatient de voir ce que nous deviendrons.

Je remercie mes oncles et tantes d'avoir été là et de m'avoir encouragé. Fabienne, Christelle, Chantale, Martine, je pense très fort à vous !

Je remercie évidemment les copains et les copines : Aurore, Bastien O., Eric, Erwan, Laurent, Maxime, Morgane, Robinson, Thomas D., Thomas R., Valentin. Mes puimseurs sûrs, mes posteurs sans vergogne(nt). Vous n'imaginez pas à quel point vous avoir à mes côtés a pu m'aider tout au long de ces années. Je vous aime, ne changez pas.

Je remercie également les membres de K&C : Alice, Audrey, Bastien C, Chloé, Léa, Eleonore, Jean-Nicolas, Katell, Maïwenn, Méganne, Noëlie, Tiphanie. Merci de supporter le moldu que je suis.

Je remercie ceux qui sont là et ceux qui ne sont plus. Je dédie cette thèse à ma grand-mère, qui était très fière de dire à toutes ses amies que son petit fils allait devenir Docteur.

Merci à celles et ceux qui se reconnaitront. Vous avez tous compté.

“So we beat on, boats against the current, borne back ceaselessly into the past.”

F. Scott Fitzgerald

The Great Gatsby

Liste des figures

Table des matières

Préambule

En 1956, un médecin de Harvard nommé Arthur Hertig publia un article qui fera date dans l'histoire de l'embryologie humaine. Fruit d'une démarche inédite, il propose la synthèse de plus de dix années de recherche portant sur l'étude des embryons humains suivant leur fécondation. Le public découvre alors 34 planches décrivant autant d'embryons, allant du 2ème au 17ème jour de développement. Pour la première fois, nous pouvions voir et étudier les tous premiers stades de la vie humaine. 70 ans après leur publication, ces images sont toujours des références incontournables, et sont encore présentes dans de nombreux ouvrages d'embryologie.

Ces travaux ont ouvert plus d'un demi-siècle de recherche en biologie du développement humain, dont un aboutissement notable fut la fécondation *in vitro* (FIV). Il était désormais possible de cultiver et d'observer de près des embryons humains, en laboratoire.

Pendant ses deux premières semaines de développement, l'embryon va subir de profonds changements morphologiques et protéiques. Il passe en quelques jours d'une cellule unique à une structure complexe asymétrique, dotée de plusieurs types cellulaires dont l'émergence, la régulation, et le potentiel sont encore mal connus aujourd'hui.

Pour 10% des couples, ces processus n'aboutissent pas à une grossesse. On parle alors d'infertilité. Pour traiter cette infertilité, le législateur a rendu possible l'aide médicale à la procréation (AMP) et la recherche sur l'embryon humain.

L'introduction qui suit récapitule l'état de nos connaissances du développement péri-implantatoire humain dans le contexte médical de l'infertilité.

Introduction

I- INFERTILITE : L'ASSISTANCE MEDICALE A LA PROCREATION

1) Les chiffres et les causes de l'infertilité

L'infertilité est définie comme l'absence de grossesse après un an de rapports sexuels réguliers non protégés. On estime qu'environ 8 à 15% des couples dans le monde sont infertiles (Boivin et al., 2007; Datta et al., 2016; Vander Borght and Wyns, 2018). Cette infertilité peut être de causes diverses. Dans un tiers des cas, elle est d'origine maternelle (trompes de Fallope obstruées, faible réceptivité endométriale, faible qualité ovocytaire). Dans un autre tiers, elle est d'origine paternelle (faible quantité de spermatozoïdes dans l'éjaculat, spermatozoïdes pas ou peu mobiles). Dans 20% des cas environ, la cause est combinée : les deux membres du couple sont infertiles. Enfin, environ 15% des infertilités sont de cause inconnue (Carson and Kallen, 2021; Nardelli et al., 2014). Entre 1990 et 2017, la prévalence de l'infertilité a augmenté chaque année de 0.37% chez les femmes et de 0.29% chez les hommes (Sun et al., 2019). Les raisons invoquées de cette infertilité sont notamment le recul de l'âge de la première grossesse chez la femme (passé en France de 26,5 ans en 1977 à 30 ans en 2016, ENP 2021), et la réduction de plus de 50% de la concentration spermatique chez l'homme entre 1973 et 2011 (ESHRE Capri Workshop Group, 2005; Levine et al., 2017).

Parfois considérée comme tabou ou honteuse, l'infertilité peut être la cause de détresse psychologique, sociale (Bak et al., 2012; Slade et al., 2007), et représente un coût individuel et sociétal important (Wu et al., 2013).

2) La prise en charge médicale

En France, les couples hétérosexuels, les couples de femmes, et les femmes seules peuvent se voir proposer une assistance médicale à la procréation (AMP). L'AMP désigne l'ensemble des techniques médicales permettant une grossesse autrement que par les processus naturels.

L'AMP recourt à plusieurs méthodes comme la stimulation hormonale, l'insémination artificielle (dépôt de sperme directement dans l'utérus), la fécondation *in vitro* (FIV), ou l'injection intra cytoplasmique (FIV-ICSI) qui seront employées au cas par cas selon le type d'infertilité. En 2020 en France, 3,4% des enfants sont nés à la suite d'une AMP (Figure 1). Cette proportion est en constante augmentation depuis la légalisation de cette pratique médicale en 1982 en France (de La Rochebrochard, 2018).

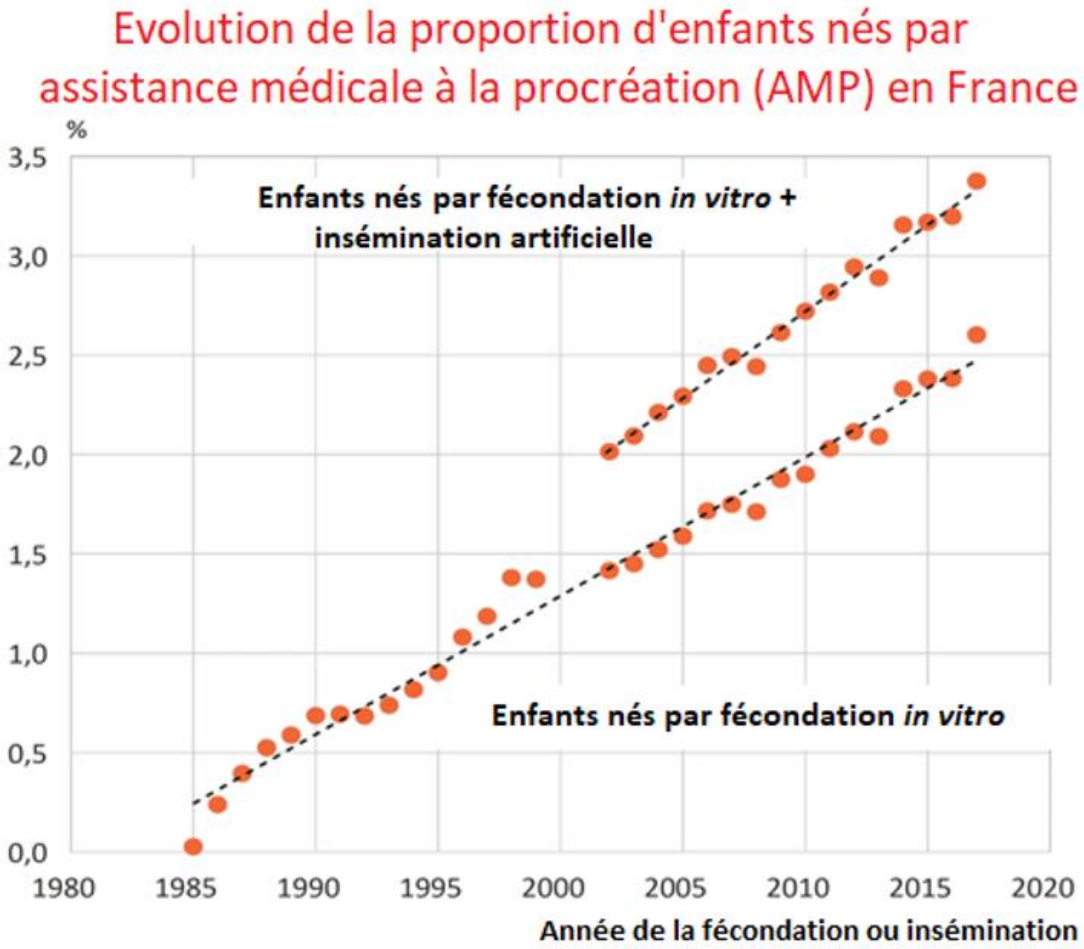


Figure 1 : évolution de la proportion d'enfants nés par AMP en France depuis 1985 La fécondation *in vitro* compte pour environ 2.5% des naissances par AMP, contre 0.9% pour l'insémination artificielle. Adapté de de La Rochebrochard, 2018, mis à jour en juillet 2020

Pour certains couples, la rencontre naturelle (par rapport sexuel) ou artificielle (par insémination) des gamètes dans l'utérus ne peut pas se produire. La stratégie de soin consiste alors en la rencontre artificielle des gamètes en laboratoire. On parle de fécondation *in vitro* (FIV). La prise en charge d'une FIV suit le schéma suivant (Figure 2) : entre 5 et 20 ovocytes sont ponctionnés chez la patiente préalablement stimulée

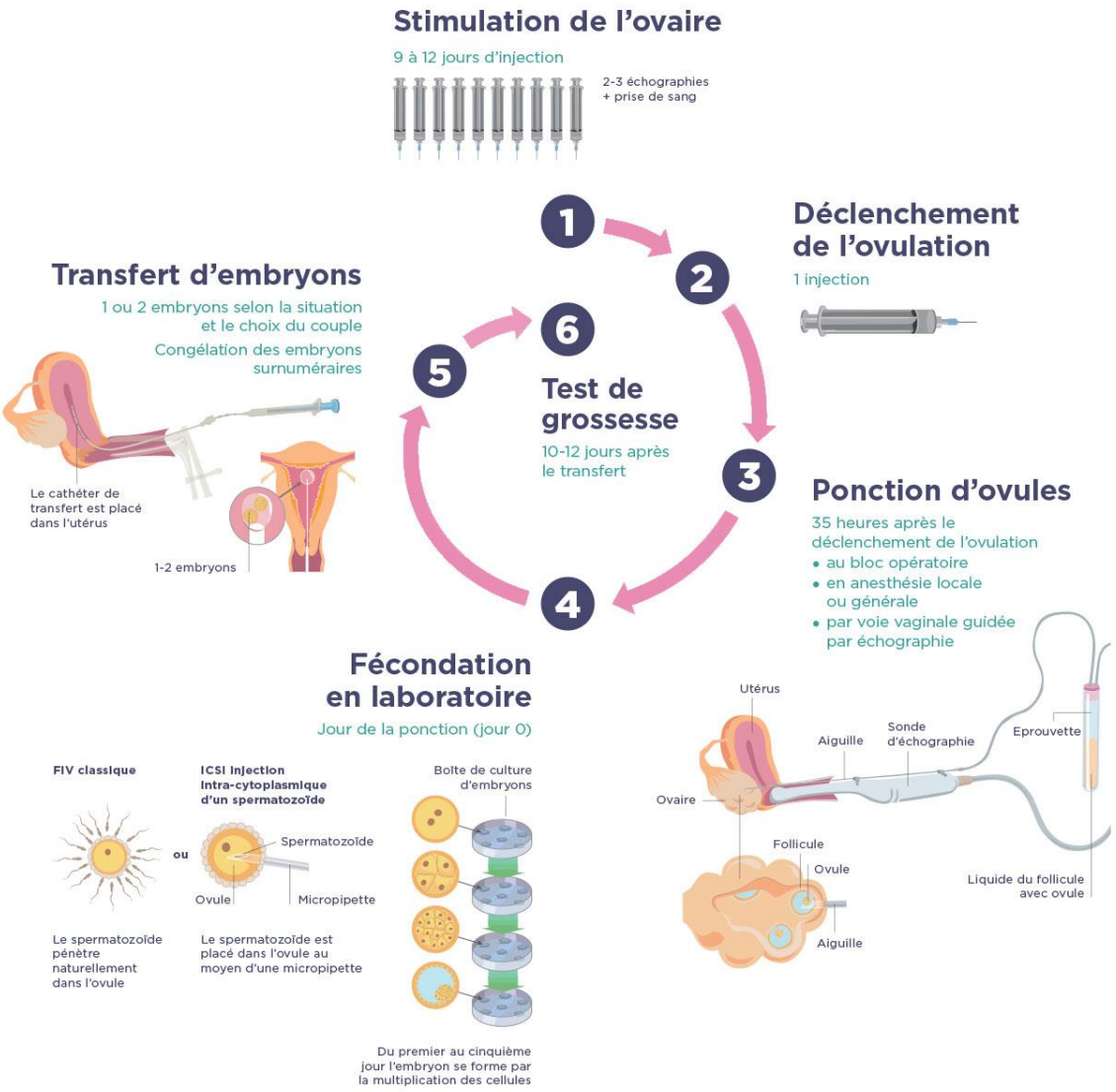


Figure 2 : schéma du cycle de FIV depuis la stimulation ovarienne au transfert utérin

Le traitement de FIV commence par une stimulation ovarienne quotidienne qui peut être effectuée par la patiente elle-même, et qui a pour but de faire produire plusieurs ovocytes en même temps. Ces ovocytes sont ponctionnés par voie vaginale et mis en contact avec le spermatozoïde *in vitro*. Les embryons obtenus sont cultivés entre 2 et 5 jours et transférés par un ou deux dans l'utérus. Figure issue de www.fertigeneve.ch. FertiGeneve n'a pas participé à la rédaction de cette thèse, et n'est pas impliqué dans les travaux qui y sont présentés.

par des hormones ovariennes telles que l'hormone folliculo stimulante (FSH) et des antagonistes de l'hormone de libération des gonadotrophines hypophysaires (GnRH). Les ovocytes sont ensuite isolés et mis en contact avec les spermatozoïdes selon deux stratégies : la FIV classique ou la FIV-ICSI. La FIV classique est principalement employée dans le cadre d'une infertilité féminine (ex : obstruction ou absence de trompes). Au laboratoire, l'ovocyte est mis en contact avec un grand nombre de spermatozoïdes jusqu'à la fécondation et l'obtention du zygote. Quand les spermatozoïdes sont peu mobiles, peu nombreux, ou malformés, on déduit que l'infertilité est masculine. Dans ce cas, il est possible de sélectionner un unique spermatozoïde et de l'injecter directement dans le cytoplasme de l'ovocyte (ICSI). L'ICSI était initialement réservée au traitement de l'infertilité masculine, mais sa sécurité et ses taux élevés de fécondation font qu'elle est utilisée dans plus 70% des FIV en Europe. (De Geyter et al., 2018).

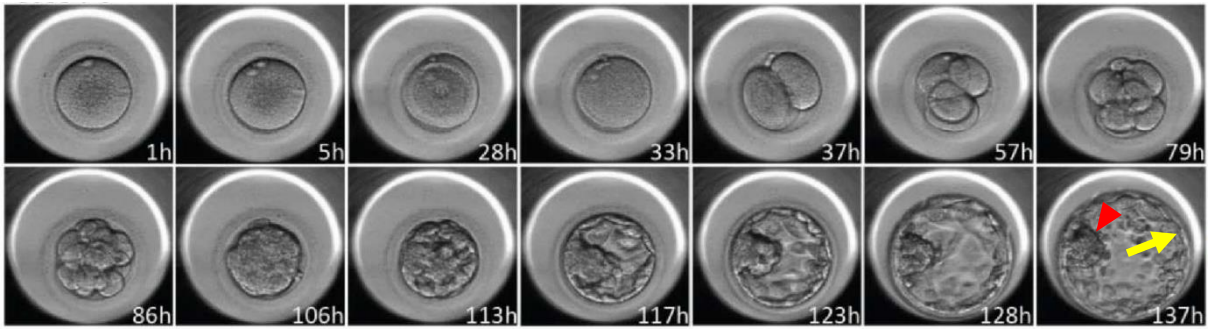


Figure 3 : développement préimplantatoire de l'embryon humain entre la fécondation et J6

Pendant les trois premiers jours de développement, l'embryon humain est composé de cellules identiques appelées blastomères (visibles jusqu'à 86h sur l'image). Ces blastomères se compactent les uns aux autres à J4 pour former la morula (106h). Entre le 4ème et 5ème jour, une cavité croit dans la morula (113h) pour former le blastocyste (117h), composé d'une couche de cellules externes appelée trophectoderme (TE, flèche jaune) et d'une masse cellulaire interne (ICM, pointe de flèche rouge).

Suite à la fécondation, l'embryon est cultivé entre 2 (J2) et 6 (J6) jours en laboratoire (Figure 3) avant d'être transféré dans l'utérus de la patiente. L'implantation et le début de la grossesse sont confirmés par un test de grossesse ou une prise de sang détectant la présence de l'hormone chorionique gonadotrope humaine (β-HCG).

Jusqu'à la fin des années 1990, la stratégie des centres AMP consistait à transférer de multiples embryons J2-3, et blastocystes si disponibles, en une seule fois dans l'utérus, avec l'espoir qu'au moins l'un d'entre eux engendre une grossesse. Cette méthode permet jusqu'à 80% de grossesse à partir de neuf embryons transférés (Gardner et Schoolgraft., 1999), mais augmente le risque de de grossesses multiples à risque pour la patiente et les fœtus (De Geyter et al., 2020). La recherche en AMP a finalement permis la mise au point de milieux adaptés à la culture des embryons jusqu'au stade blastocyste. Cultiver les embryons plus longtemps permet de sélectionner ceux qui ont le meilleur potentiel de survie, d'implantation, et de grossesse. L'ensemble du parcours de soin depuis la ponction ovocytaire jusqu'au transfert de tous les embryons générés suite à cette seule ponction est appelé « cycle de FIV ». Plusieurs cycles peuvent être nécessaires à l'obtention d'une grossesse.

3) Culture *in vitro* et sélection des embryons humains

Un embryon qui se développe *in vitro* n'est pas assuré de s'implanter et de donner une grossesse. Aussi, les embryologistes ont recours à des méthodes d'évaluation de la qualité embryonnaire afin de ne transférer que ceux qui ont le plus de chance, a priori, de s'implanter. Ces évaluations peuvent se faire dès J2 où le nombre de blastomères, leur homogénéité, et le pourcentage de fragmentation est évalué (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). Le stade morula peut également être évalué en mesurant le nombre de blastomères impliqués dans la compaction, la morphologie de la masse compacte, et la présence de fragments (Tao et al., 2002, Reignier 2020). La méthode la plus répandue aujourd'hui est la classification de Gardner et Schoolgraft (Gardner et al., 2000). Elle repose sur l'évaluation de trois critères morphologiques de l'embryon au stade

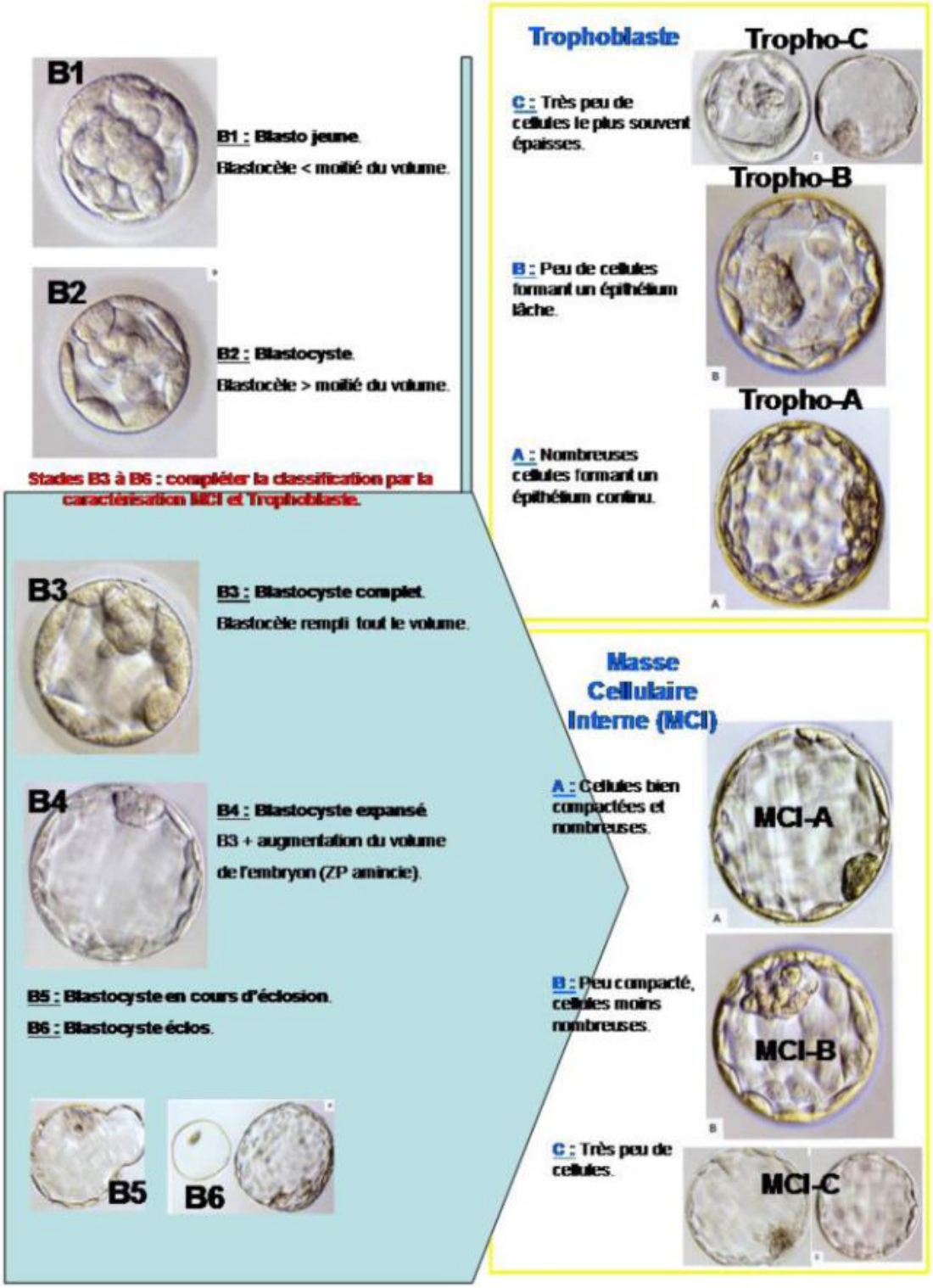


Figure 4 : Classification de Gardner et Schoolgaft de la morphologie embryonnaire

L'apparition de la cavité marque le début du stade blastocyste. Le chiffre 1 (B1) est attribué quand la cavité fait moins de 50% du volume de l'embryon, et 6 (B6) quand l'embryon est sorti de sa zone pellucide. A partir du stade B3, quand la cavité remplit tout le volume, l'ICM et le TE peuvent être évalués de A (bonne qualité) à C (mauvaise qualité). Exemple : un embryon classé « B5 AA » est en cours d'éclosion, et son ICM ainsi que son TE sont de bonne qualité. Figure issue de Alter et al., 2014

blastocyste que sont l'expansion (taille de la cavité, amincissement de la zone pellucide, éclosion), qualité apparente de l'ICM (nombre de cellules, cohésion), qualité apparente du TE (nombre de cellules, cohésion, épaisseur) (Figure 4).

En transférant des blastocystes de bonne qualité « AA », le taux de grossesse est de 65% alors qu'il est de 50% pour les embryons de qualité moyenne (4BB, 4AC, 4CA, 2AB, 2BA), et de seulement 33% pour les embryons de mauvaise qualité (3BC, 4CB, 4CC, 2BB) (Zhao et al., 2018). Plus que l'ICM, la qualité du TE semble être le critère qui corrèle le mieux avec une implantation de l'embryon dans l'utérus (Ahlström et al., 2011; Ebner et al., 2016; Thompson et al., 2013). Nous pouvons lier cette observation avec le fait que le TE est le premier tissu de l'embryon à interagir avec l'endomètre, et que cette interaction peut être moins efficace si le TE est moins bien formé.
L'application de cette classification couplée à la stratégie de transfert unique d'embryon a drastiquement diminué le nombre de grossesses gémellaires dans le cas de procédures d'AMP. Ainsi, entre 2000 et 2017 aux Etats-Unis, le taux de transfert unique d'embryon est passé de 5.7% à 64.2%. Dans le même temps, les taux de grossesses multiples sont passés de 53.1% à 26,4% (Ma et al., 2022).

Chez les mammifères comme le lapin, le macaque rhésus, ou le hamster, le taux d'oxygène (O2) dans les oviductes et l'utérus varie entre 2% et 8%, contre 21% dans l'air ambiant. (Fischer and Bavister, 1993; Mastroianni and Jones, 1965). Chez l'humain, le peu d'informations disponibles suggèrent que le taux d'O2 y est également plus faible que dans l'air ambiant (Kigawa, 1981; Yedwab et al., 1976). Ce taux peut varier considérablement d'une patiente à une autre, de manière cyclique, et d'une minute à l'autre par l'influence de la contractilité utérine et des hormones (Ng et al., 2018; Ottosen et al., 2006). La concentration d'O2 optimale pour la culture *in vitro*

d'embryons humains a longuement été débattue (Gardner, 2016). Certaines études font état d'une meilleure qualité embryonnaire en culture hypoxique sans amélioration du taux de grossesse (Dumoulin et al., 1999; Kea et al., 2007), quand d'autres attestent de meilleurs taux de naissances vivantes (Meintjes et al., 2009; Waldenström et al., 2009). Une méta-analyse a soutenu que la culture en hypoxie améliore le succès global des prises en charges en AMP (Bontekoe et al., 2012). Plus précisément, il semblerait que l'embryon humain ait au moins besoin d'une culture hypoxique les deux-trois premiers jours de développement jusqu'au stade 8 cellules, et que la culture au stade blastocyste peut se faire en hypoxie ou normoxie (De los Santos et al., 2013; Guo et al., 2014; Herbemont et al., 2021; Kovacic et al., 2010). La Société Européenne de Biologie de la Reproduction en Embryologie (ESHRE) recommandera finalement la culture en hypoxie pour la culture des embryons humains quel que soit leur stade (ESHRE Guideline Group on Good Practice in IVF Labs et al., 2016).

L'évaluation morphologique des embryons nécessite une observation au jour le jour, donc de sortir les embryons de leur incubateur quotidiennement. Afin de préserver au maximum les conditions de culture, des incubateurs « timelapse » ont été développés. Ces incubateurs maintiennent les embryons en hypoxie et les prennent en photo à intervalles réguliers. Ils renseignent en temps réel les données morphocinétiques de l'embryon, permettant ainsi une classification morphologique plus fine et plus facile pour l'embryologiste. L'intérêt clinique de cette méthode de culture a lui aussi été débattu. Des méta-analyses et études rétrospectives récentes tendent à souligner leur supériorité en terme de naissances vivantes comparé à la culture conventionnelle (Bartolacci et al., 2021; Pribenszky et al., 2017; Reignier et al., 2021).

Dans le cas où les patients sont porteurs de maladies génétiques mortelles ou fortement invalidantes pour le futur enfant (myopathie de Duchenne, mucoviscidose…), un diagnostic pré-implantatoire (DPI) peut être effectué avant le transfert de l'embryon. L'embryon cultivé à J2-J3 se voit prélevé d'une à deux cellules à l'aide d'un laser à dissection. La ploïdie et l'intégrité du génome des cellules peuvent être analysées par hybridation *in situ* (FISH), réaction en chaine par polymérase quantitative (qPCR), ou séquençage du génome (Sermon et al., 2016; Treff and Scott, 2013, Reignier 2020). Les résultats sont obtenus en moins de 24h-48h afin de sélectionner les embryons sains avant le transfert.

4) Les limites de l'AMP, la recherche sur l'embryon

Bien que la prise en charge de l'infertilité ait évolué en 40 ans, les taux de grossesse suite à une AMP restent d'environ 25-30% par cycle (Bodri et al., 2014; De Geyter et al., 2018; Reignier et al., 2021). Ce taux varie à la baisse avec l'âge de la patiente

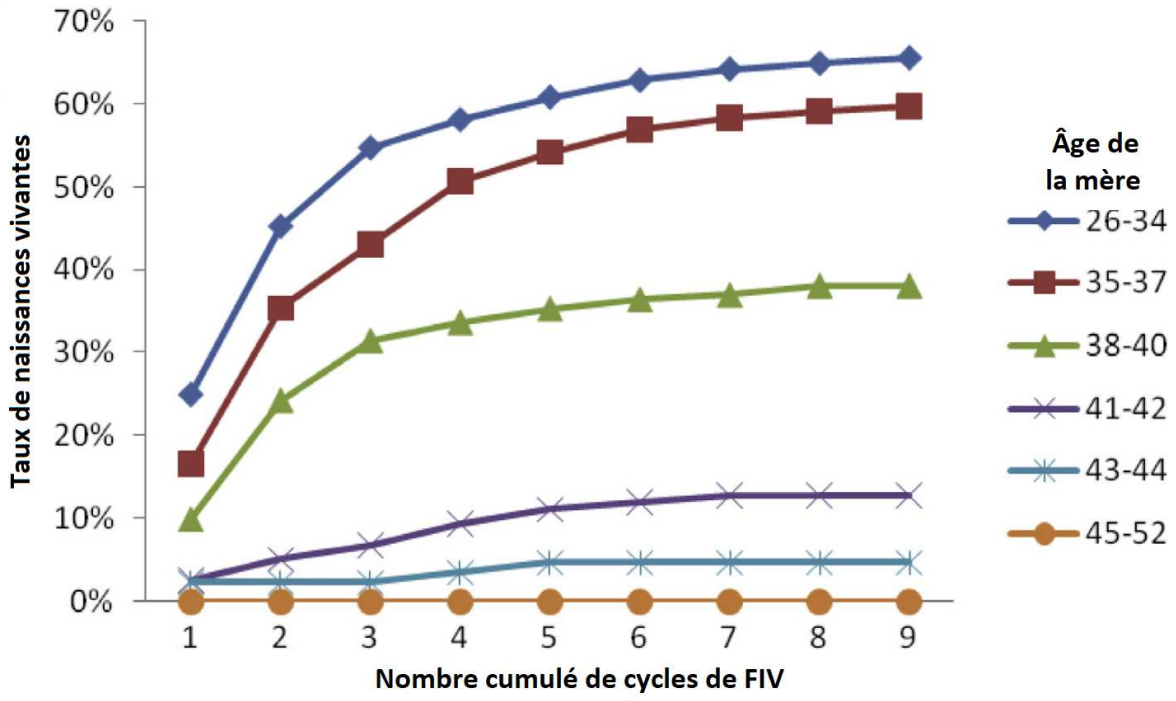


Figure 5 : taux cumulés de naissances vivantes selon l'âge de la mère en FIV

On note qu'après 41-42 ans, les taux de naissances sont inférieurs à 10%. Figure adaptée de Bodri et al., 2014.

quand il s'agit de ses propres ovocytes (Figure 5), mais varie peu lorsque que l'ovocyte est issu d'un don (2015 ART National Summary Report; CDC, 2017). Les embryons issus de dons, potentiellement de femmes plus jeunes, ont en effet une probabilité plus faible d'être aneuploïdes (Makhijani and Grow, 2020).

Une explication possible à ces taux bas de réussite en PMA est notre méconnaissance des mécanismes développementaux de l'embryon, et leur importance dans la formation et la sélection d'un embryon viable pour un transfert. En effet, le développement embryonnaire humain est notoirement inefficace. Dans une cohorte d'ovocytes fécondés, seuls 40% atteignent le stade blastocyste avec une qualité suffisante pour un transfert (Chadid et al., 2015; De Geyter et al., 2018). En dehors des critères morphologiques de l'embryon, nous ne disposons pas d'autres critères prédictifs du succès d'une FIV.

Afin d'améliorer la prise en charge de l'infertilité, la loi française a rendu légale sous conditions la recherche sur l'embryon humain. Aucun embryon ne peut être créé dans un but de recherche uniquement. Tous les embryons utilisés en recherche proviennent exclusivement de parcours d'AMP, et ont été donnés par les couples informés sans compensation financière. Depuis la révision de la Loi n° 2021-1017 du 2 août 2021 relative à la bioéthique, les embryons humains peuvent être modifiés génétiquement. Aucun embryon donné à la recherche ne peut être transféré dans un utérus.

Ces autorisations de recherche nous ont permis d'en apprendre davantage sur l'embryon humain, notamment la dynamique d'expression de ses gènes et de ses protéines, à certains moments, dans certaines cellules au cours du développement. Nous avons surtout appris à quel point l'embryon humain était différent de l'embryon de souris, modèle d'étude incontournable du développement des mammifères. Dans le prochain chapitre, nous aborderons les aspects mécanistiques, transcriptomiques, et protéiques du développement embryonnaire murin et humain.

II- LES 14 PREMIERS JOURS DE DEVELOPPEMENT DE L'EMBRYON HUMAIN

1) La première semaine de développement

La première semaine de développement de l'embryon humain est appelée « développement pré-implantatoire ». Elle comprend les évènements survenant depuis la fécondation de l'ovocyte par le spermatozoïde, jusqu'à l'attachement de l'embryon dans l'utérus au 7ème jour. Pendant cette période, l'embryon humain va multiplier son nombre de cellules, former le blastocyste, et spécifier plusieurs lignées cellulaires que sont :

- L'épiblaste (EPI), pluripotent, et formant le futur fœtus
- Le trophectoderme (TE), extra-embryonnaire, formant le placenta.
- L'endoderme primitif (PrE), extra-embryonnaire, formant le sac vitellin.

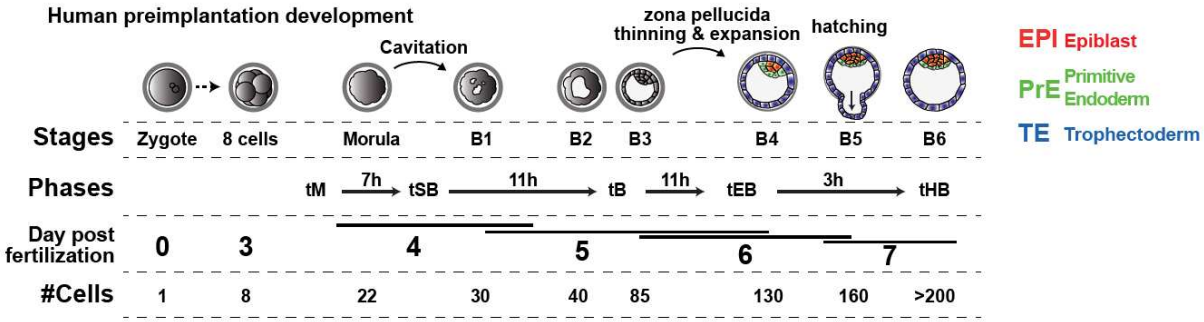


Figure 6 : schéma du développement pré-implantatoire humain

Différents critères sont utilisés pour décrire le stade développemental des embryons : les stades morphologiques (Gardner et al., 2000), les phases de morphocinétiques, les jours post fécondation (dpf) ou encore le nombre de cellules totales dans l'embryon. L'embryon blastocyste est composé du trophectoderme externe et d'une masse cellulaire interne (ICM) comportant l'épiblaste (EPI) et l'endoderme primitif (PrE). tM = début du stade morula, tSB = début du stade blastocyste, tB = blastocyste plein, tEB = début de l'amincissement de la zone pellucide, tHB = embryon éclos sorti de sa zone pellucide.

A – De la fécondation à la première spécification

Suite à la fusion des gamètes mâles et femelles, l'embryon humain va subir des divisions cellulaires successives au rythme d'une toutes les 24h environ. Entre la fécondation et jusqu'au 3ème jour de développement, les cellules de l'embryon sont appelées blastomères. Elles sont morphologiquement identiques et sont considérées comme totipotentes. Une seule d'entre elle isolée est capable de reformer un embryon entier et de permettre une grossesse (Van de Velde et al., 2008; Veiga et al., 1987). Pendant les trois premiers jours de développement, le génome de l'embryon n'est pas ou peu actif. Son métabolisme est assuré par les protéines et ARN messagers (mRNA) maternels présents initialement dans l'ovocyte. L'activation du génome embryonnaire (EGA) est progressive, elle est initiée dès le stade zygote (Asami et al., 2022) et s'achève au 3ème jour de développement (Asami et al., 2022; Braude et al., 1988; Niakan et al., 2012; Vassena et al., 2011). Peu après l'EGA, les 6-12 blastomères sphériques augmentent leur surface de contact et se compactent les uns aux autres. Il devient alors impossible de les discerner au microscope en contraste de phase : l'embryon est à son 4ème jour de développement, c'est le stade morula. C'est à partir de ce moment qu'émergent les premières différences entre les cellules de l'embryon.

Dans la morula, les cellules peuvent être localisées soit à l'extérieur soit à l'intérieur de l'embryon. En fonction de leur localisation, les cellules auront un destin différent. Les cellules internes de la morula formeront la masse cellulaire interne « ICM », d'où émergerons l'EPI et le PrE plus tard dans le développement. Les cellules externes de la morula elles formeront le TE, extra-embryonnaire, futur placenta. Ce phénomène de ségrégation de l'ICM et du TE à partir de la morula est communément appelé « première spécification » (Shahbazi, 2020).

Pour des raisons pratiques et réglementaires, le lien entre la position des cellules au stade morula et leur destin a surtout été étudié chez la souris (Figure 7). Au stade

morula, les cellules externes compactées acquièrent une polarité (Johnson and Ziomek, 1981). Les cellules internes elles restent apolaires. Cette différence de polarité est visible par la localisation spécifique de certaines protéines membranaires. Le pôle apical des cellules externes est enrichi en microvillosités, la F-actine, l'Ezrine, et le complexe protéique apical contenant PAR3, PARD6, et la protéine kinase atypique C (aPKC). Leur pole basolatéral exprime entre autres la E-cadhérine et la β-caténine (Chazaud and Yamanaka, 2016; Yamanaka et al., 2006). Dans les cellules internes apolaires, la E-cadhérine et la β-caténine sont exprimées sur toute la surface de la membrane (de Vries et al., 2004).

Polarité et destin cellulaire sont vraisemblablement liés par la voie HIPPO. Il s'agit d'une voie de signalisation largement conservée chez les mammifères modulant la prolifération cellulaire et l'apoptose en réponse à des signaux de l'environnement, dont le contact cellule-cellule et la polarité (Chazaud and Yamanaka, 2016; Yu et al., 2015). Dans une cellule polarisée, la voie HIPPO est inactive. Son effecteur YAP1 transloque librement dans le noyau et fixe son co-activateur TEAD4. Le complexe YAP1/TEAD4 active la transcription de gènes régulateurs du destin TE tels que *Cdx2* et *Gata3* (Nishioka et al., 2009; Yagi et al., 2007). Ainsi, les cellules externes polarisées de la morula s'orientent vers un destin TE. Dans les cellules internes apolaires, où la voie HIPPO est active, AMOT et NF2 forment un complexe qui permet à la kinase LATS1/2 de phosphoryler YAP1, qui ne peut pas entrer dans le noyau et rejoindre TEAD4. Les cellules internes de la morula exprimeront les facteurs SOX2, OCT4 et s'orienteront vers un destin ICM.

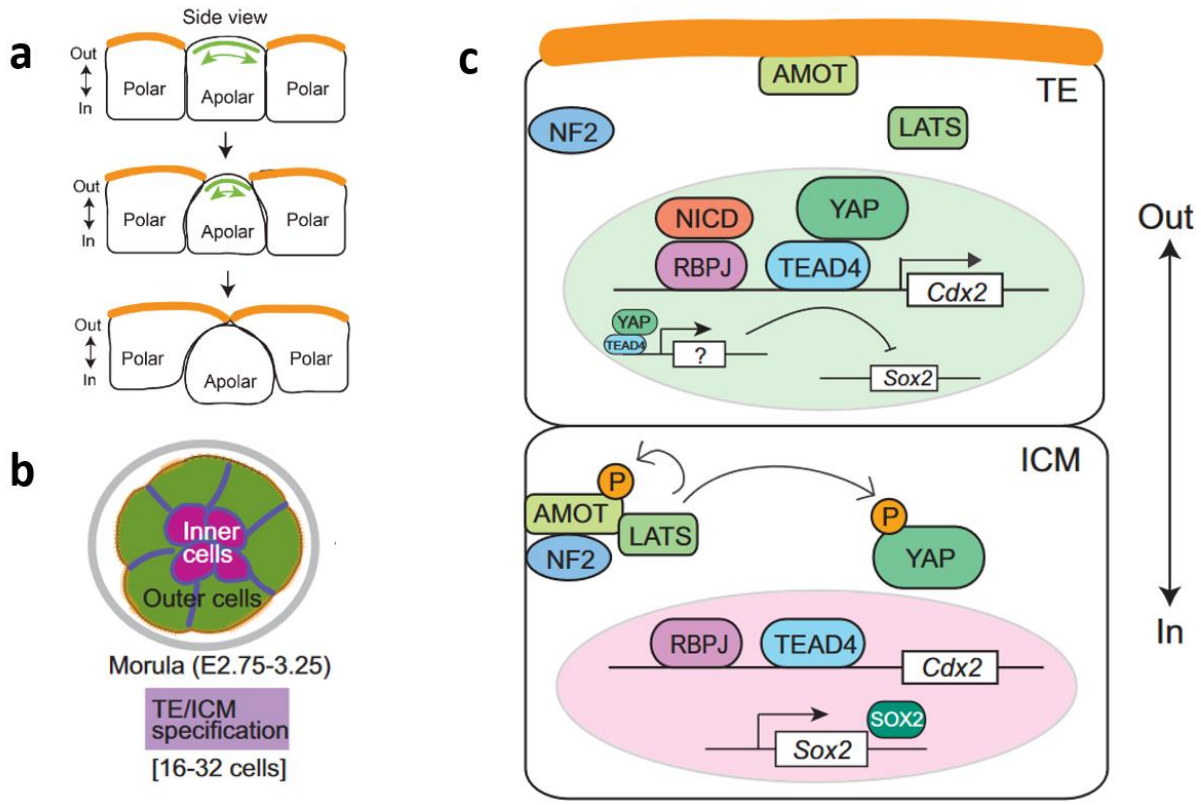


Figure 7 : la première spécification dans l'embryon de souris

Au moment de la compaction, certaines cellules apolaires peuvent être localisées au niveau des cellules externes de la morula. Ces cellules initient une constriction apicale là où il n'y a pas de contact cellule-cellule, et finissent internalisées (a). La morula de souris est alors formée d'environ 16 à 32 cellules polaires ou apolaires (b). Les cellules internes apolaires s'orienteront vers un destin ICM, et les cellules externes polaires vers un destin TE (c)

La polarisation est nécessaire à la spécification des cellules en TE. En bloquant le processus de polarité, par inhibition de l'activité de la aPKC ou PARD6, les cellules externes expriment SOX2 et le développement de l'embryon se stoppe au stade morula (Alarcon, 2010; Frum et al., 2013; Gerri et al., 2020; Korotkevich et al., 2017; Plusa, 2005). Dans les embryons murins *Nf2-/-*, on observe une localisation ectopique de YAP1 qui transloque librement dans le noyau des cellules internes de la morula. Les embryons *Nf2-/-* blastulent normalement, mais l'ICM des blastocystes exprime CDX2, et non les facteurs de pluripotence OCT4 et NANOG.

Contrairement à son coactivateur YAP1 dont la localisation dépend de la polarité, TEAD4 est exprimée dans le noyau de toutes les cellules de la morula. La présence de TEAD4 est également nécessaire à la spécification du TE. Les cellules externes

des embryons de souris *Tead4*-/- polarisent normalement, mais n'expriment pas ou peu ni CDX2, ni GATA3. L'embryon ne cavite pas, et son développement se stoppe au stade morula (Nishioka et al., 2008; Yagi et al., 2007). La formation en bonne et due forme du complexe de polarité au stade morula semble être une étape indispensable à l'initiation du programme TE, sans quoi l'embryon ne forme pas le blastocyste.

Les embryons murins *Cdx2*-/-, bien qu'incapables de s'implanter, peuvent caviter et expriment toujours GATA3 dans le TE (Ralston et al., 2010; Strumpf et al., 2005). Ceci suggère que *Cdx2* seul n'est pas strictement nécessaire à l'initiation du programme TE, et que l'expression de *Gata3* n'est pas directement dépendante de *Cdx2*. En effet, bien qu'ils s'auto régulent positivement, *Cdx2* et *Gata3* sont exprimés en parallèle l'un de l'autre, en aval de *Tead4* (Home et al., 2009; Ralston et al., 2010). Cette apparente redondance de fonction entre *Cdx2* et *Gata3* n'est que relative : le phénotype *Cdx2*-/- est lui aussi léthal *in vivo.*

La première spécification est une cascade d'évènements précisément orchestrés dans le temps et dans l'espace de l'embryon. Si nous en comprenons les grandes étapes chez la souris, nous en ignorons la plupart des rouages chez l'Homme. Une différence notable est l'expression de CDX2. Chez l'Homme, CDX2 n'est pas détectable avant le stade blastocyste au 5ème jour de développement (Niakan and Eggan, 2013). C'est également le cas chez le bovin et le porc (Bou et al., 2017; Goissis and Cibelli, 2014; Liu et al., 2015; Madeja et al., 2013). Aussi, chez l'Homme, OCT4 semble avoir une fonction différente de chez la souris. Les embryons *OCT4-/-* humains blastulent moins efficacement, avec une plus petite cavité, et expriment plus faiblement des gènes associés au TE comme *CDX2, HAND1, DLX3, PLAC8* et *GATA2* (Fogarty et al., 2017). Les embryons *OCT4-/-* n'expriment pas NANOG, et le développement du PrE semble perturbé avec une baisse d'expression de *GATA4* et SOX17. *OCT4* chez l'Homme

semble donc nécessaire à la spécification des trois lignées embryonnaires EPI, TE, et PrE. Ces observations sont différentes chez la souris, où l'embryon *Oct4/-* maintient l'expression de NANOG, et où l'expression de gènes du TE comme *Cdx2*, *Hand1* et *Gata3* est augmentée (Frum et al., 2013; Nichols et al., 1998).

B – La seconde spécification

Suite à la cavitation, l'embryon est appelé le blastocyste. Il est composé d'un TE périphérique et d'une masse cellulaire interne (ICM) accolée au TE. Au sein de cette ICM, deux types cellulaires vont progressivement émerger : l'épiblaste « EPI », pluripotent, formant le futur fœtus, et l'endoderme primitif « PrE », extra-embryonnaire, formant le futur sac vitellin du fœtus. L'apparition de l'EPI et du PrE à partir de l'ICM est communément appelée « seconde spécification ». Comme la première spécification, elle est un processus dynamique qui s'inscrit dans le temps et l'espace de l'embryon.

A E4.5 chez la souris, l'EPI est caractérisé par l'expression de NANOG. Le PrE exprime lui GATA6. Ces deux facteurs de transcription sont co-exprimés dès le stade 8-cellules dans tous les blastomères de l'embryon de souris, soit avant la cavitation (Chazaud et al., 2006). Entre E3.0 et E3.75, certaines cellules de l'ICM vont progressivement perdre NANOG et former les progéniteurs PrE, et d'autres vont perdre GATA6 et former les progéniteurs EPI. Cette ségrégation est asynchrome, l'EPI se formant avant le PrE (Bessonnard et al., 2014; Plusa et al., 2008; Saiz et al., 2016). A E3.75 au stade jeune blastocyste 32-cellules, la plupart des cellules expriment soit NANOG, soit GATA6, dans une disposition dite « poivre et sel » (Artus et al., 2011; Chazaud et al., 2006; Schrode et al., 2013). Les mécanismes de spécification de l'EPI et du PrE sont débattus.

Dans des embryons murins *Nanog*-/-, toutes les cellules de l'ICM expriment GATA6. Inversement, dans des embryons *Gata6*-/-, toutes les cellules de l'ICM expriment NANOG. Ceci suggère que NANOG et GATA6 se régulent négativement l'un l'autre, et directement comme montré par des études de co-immunoprécipitation (Singh et al., 2007; Wamaitha et al., 2015).

En plus de l'action de NANOG et GATA6 l'un sur l'autre, la spécification de l'EPI et du PrE est dépendante de la voie fibroblast growth factor (FGF)(Nichols et al., 2009; Yamanaka et al., 2010). La voie FGF est une voie de signalisation de type tyrosine kinase, dont l'activité consiste en la fixation d'un agoniste sécrété sur un récepteur membranaire. Dans l'ICM de souris, les cellules NANOG sécrètent du FGF4 dont le récepteur FGFR2 est exprimé à la membrane des cellules GATA6. En bloquant la voie FGF, toutes les cellules de l'ICM adoptent une identité EPI en exprimant NANOG (Chazaud et al., 2006; Kang et al., 2013). Et inversement, un embryon cultivé avec un excès de FGF4 n'aura dans son ICM que des cellules GATA6 (Yamanaka et al., 2010).

L'action dirigée de la voie FGF explique en grande partie la restriction de NANOG et GATA6 à certaines cellules de l'ICM. Pour autant, les raisons d'une hétérogénéité initiale -l'expression plus importante de NANOG et FGF4 dans certaines cellules et pas dans d'autres- sont longtemps restées un mystère.

Récemment, il a été mis en évidence que l'initiation de l'EPI depuis le stade 16-32 cellules (morula-blastocyste) est purement aléatoire. A ce stade, toutes les cellules de l'ICM co-expriment à des niveaux variables des gènes de l'EPI (dont *Nanog*) et du PrE (notamment *Gata6*). Aléatoirement, certaines de ces cellules vont exprimer davantage de gènes de l'EPI et vont activer, via NANOG, un programme de spécification de l'EPI et d'expression de FGF4. Les cellules adjacentes, qui expriment moins de gènes de

l'EPI vont s'orienter vers un destin PrE suite à l'activation de la voie FGF (Allègre et al., 2022).

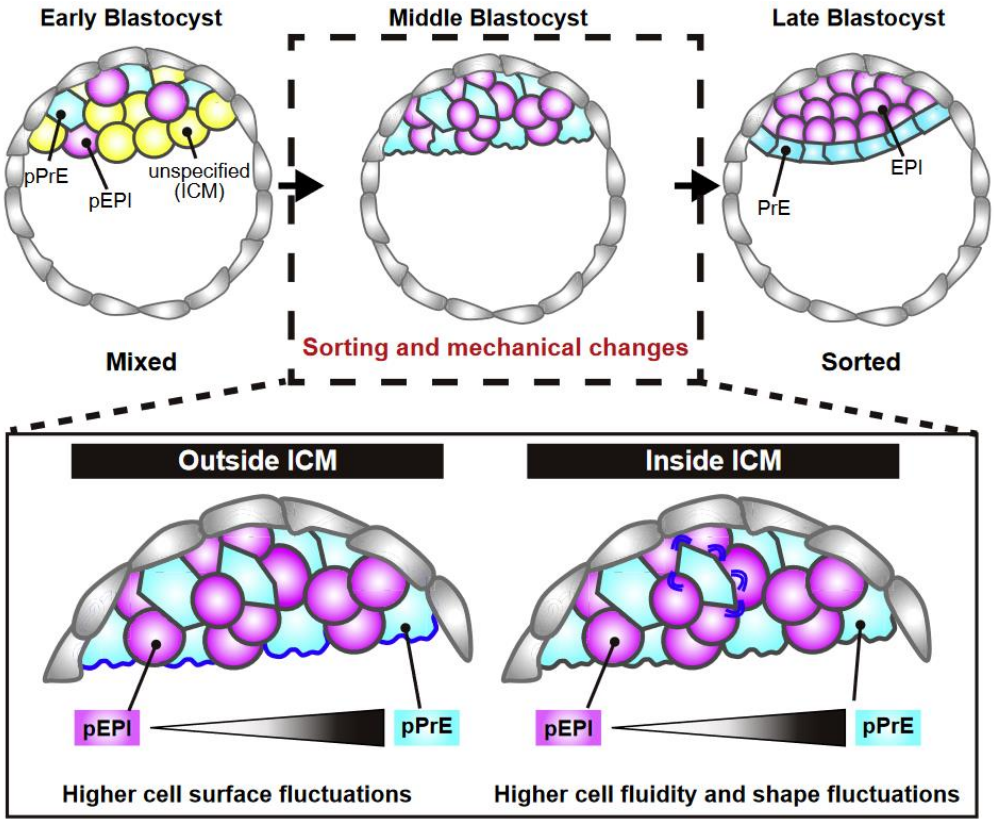


Figure 8 : ségrégation de l'EPI et du PrE par fluctuations membranaires

La membrane des cellules du PrE est plus fluide, et peut plus facilement changer de forme sous la contrainte des cellules de l'EPI qui ont une membrane moins mobile. Les contraintes physiques entrainent une relocalisation des cellules du PrE face à la cavité du blastocyste. Figure issue de Yanagida et al., 2022.

Bien que l'axe FGF-NANOG-GATA6 soit capital dans la spécification EPI-PrE, il n'explique pas à lui seul la ségrégation spatiale des cellules du PrE (exprimant GATA6) en une monocouche adjacente à la cavité du blastocyste. Pendant longtemps, les hypothèses se sont succédées sans vraiment conclure de manière formelle sur le mécanisme sous-jacent. Récemment, il a été montré que la fluidité membranaire des cellules du PrE était plus variable et plus élévée que celles de l'EPI (Figure 8). Cette fluidité membranaire est caractérisée par la forme active phosphorylée du complexe membranaire Ezrin-Radixin-Moesin (pERM). Des cellules du PrE isolées ont davantage de « blebbing » (proéminence membranaire) que les cellules de l'EPI. Des

cellules souches embryonnaires (ES) de souris, modèle de l'EPI, transfectées avec des siRNA ERM et ACT4 (Actinine-4, une protéine liée à l'actine) vont davantage « blebber » que les ES contrôles. Si ces cellules sont intégrées à un blastocyste à E3.25-3.75, elles vont se positionner dans l'ICM en face du blastocèle, soit une position de cellules du PrE (Yanagida et al., 2022). Ainsi, chez la souris, l'action de NANOG, GATA6, de la voie FGF et les fluctuations membranaires forment une boucle de régulation finement contrôlée permettant l'émergence et la ségrégation de l'EPI et du PrE.

Chez l'Homme, certains de ces mécanismes ne semblent pas conservés. Dans l'embryon humain, NANOG n'est détectable qu'à partir du stade blastocyste expansé B3-B4 à J5 (Cauffman et al., 2009). GATA6 est détectable dès le stade morula comme chez la souris, et est exprimé dans toutes les cellules de l'embryon puis est progressivement restreint à l'ICM. La plupart des études font état de cellules NANOG et GATA6 mutuellement exclusives dans l'ICM à J6-J7, mais montrent également des cellules exprimant à la fois NANOG et GATA6 dans l'ICM à J6, suggérant une spécification inachevée avant l'implantation (Deglincerti et al., 2016; Roode et al., 2012). A J6, la majorité les cellules de l'ICM GATA6+/NANOG- expriment GATA4, et certaines cellules NANOG expriment GATA6. L'ICM est composée de cellules double positives NANOG+/OCT4+, et de cellules SOX17 et/ou GATA6 dont la plupart expriment toujours OCT4. Ces observations suggèrent que l'EPI serait formé plus tôt que le PrE, et que la spécification de ce dernier n'est pas complète avant l'implantation. Contrairement à la souris, les embryons humains cultivés avec des inhibiteurs de la voie FGF conservent un nombre normal de cellules NANOG ou GATA6 (Kuijk et al., 2012; Roode et al., 2012). Aussi, les embryons humains traités avec du FGF2 entre J2 et J6 n'expriment pas NANOG, quand le même traitement effectué entre J5 et J6 n'affecte pas l'embryon (Wamaitha et al., 2020). Le nombre de cellules GATA6 n'est

pas modifié quel que soit le jour du traitement. Ceci suggère que le FGF2 a peut-être un rôle dans l'initiation de l'EPI, mais pas son maintien.

Il est important de souligner que ces données sur l'embryon humain n'ont été obtenues que sur faible nombre de réplicas par expérience (<6), expériences qui n'ont pour la plupart jamais été refaites dans d'autres laboratoires. Bien qu'en nombre limité, les informations dont nous disposons suggèrent que les mécanismes de la seconde spécification chez l'Homme sont différents de ceux chez la souris.

C – La révolution du séquençage ARNm en cellule unique

Ces dix dernières années, une avancée technologique majeure a bouleversé notre approche dans l'étude intégrée des systèmes biologiques : le séquençage ARN à haut débit (RNAseq). Avant d'aborder son utilisation -et sa pertinence- dans l'étude du développement humain, nous commencerons par une brève introduction à cette méthode d'analyse.

Toutes les cellules de notre organisme possèdent le même ADN, qui est le support de l'information génétique. Bien qu'ayant les mêmes gènes, les cellules peuvent être de tailles, de formes, et de fonctions différentes. Cette diversité de phénotypes est due à l'expression différentielle du génome, c'est-à-dire l'activation au cas par cas d'un ensemble de gènes nécessaires à la réalisation d'une ou plusieurs fonctions biologiques. Quand un gène est actif (« transcrit »), il produit une molécule d'ARNm complémentaire à ce gène qui sera ensuite traduite en protéine. L'ensemble des ARN présents dans une cellule à un instant donné est appelé le transcriptome. Le transcriptome reflète l'état d'activation global du génome d'un échantillon biologique.

Le séquençage ARN en masse « RNA-seq » (Mortazavi et al., 2008) est un ensemble de techniques d'extraction, de préparation, et d'analyse de tous les ARN (ARNm mature, pré-ARNm, miRNA, lncRNA) présents dans un échantillon. Elles commencent généralement par une étape d'extraction des ARN de l'échantillon, qui subiront un processus de transcription inverse et d'amplification (RT-PCR) à partir de leur queue poly-A. Pendant la RT-PCR, et selon la stratégie de séquençage, divers éléments artificiels peuvent être ajoutés aux amplicons pour faciliter leur identification (des codes-barres), leur séquençage (des adapteurs), ou leur comptage (UMI, identifiant moléculaire unique). Les adapteurs sont nécessaires aux premières étapes du séquençage une fois la libraire insérée dans le séquenceur. Les codes-barres permettent d'identifier un échantillon, ce qui est utile si l'on décide de mélanger des échantillons entre eux afin de réduire les coûts de séquençage. Enfin, les UMI sont des séquences aléatoires qui, séquencées avec l'ADNc auxquelles elles sont fixées, permettent de compter le nombre absolu d'ARNm présents initialement dans l'échantillon (Islam et al., 2014). Le produit d'amplification est généralement fragmenté en plusieurs petits bouts de 300-1000 paires de bases avant d'être chargé dans un séquenceur haut débit qui en sortira un fichier (généralement de type FASTQ), comportant l'ensemble des fragments séquencés appelés « *read*s». Chaque *read* comporte un identifiant, la séquence, et le score de qualité de chaque nucléotide. Ce fichier brut sera ensuite aligné sur un génome de référence afin d'attribuer chaque *read* à un gène. Suite à l'alignement, une table de comptage est obtenue avec le nombre de *reads* par gène pour chaque échantillon. Dans le cas de séquençages dits *full lenght* ou *whole transcript,* les multiples fragments obtenus lors de la préparation de la librairie sont séquencés. Cette technique est utile pour mesurer l'expression de transcrits alternatifs, mais peut comptabiliser plusieurs fragments d'un seul transcrit initial.

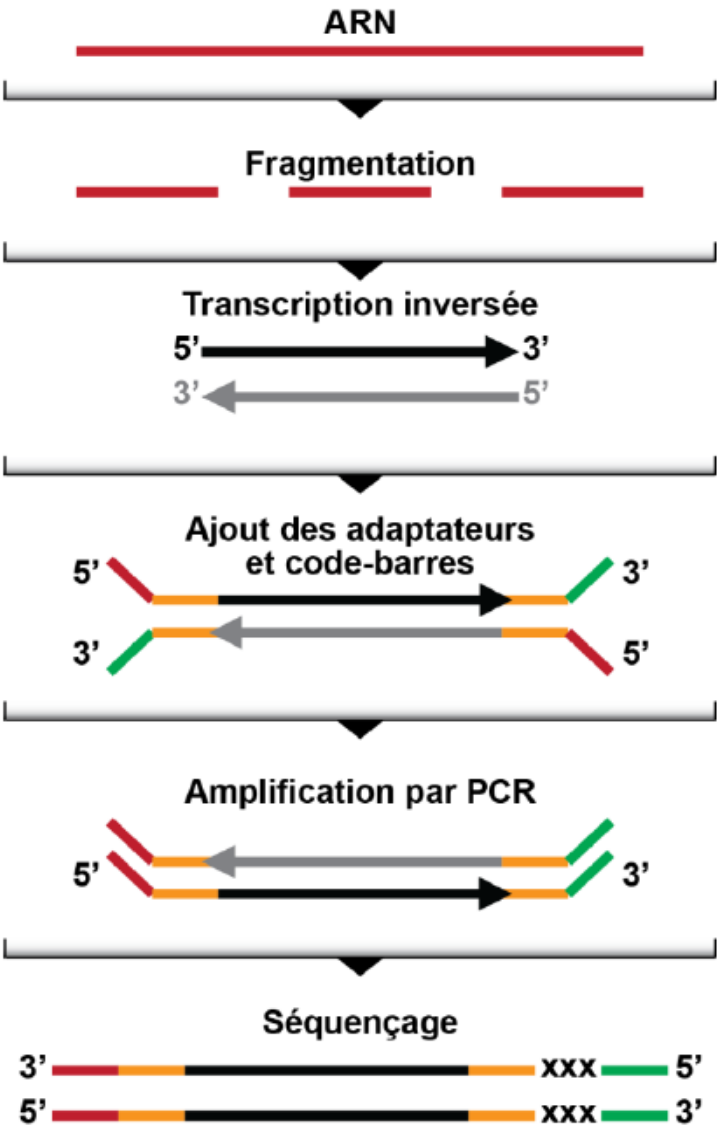


Figure 9 : principales étapes d'une préparation de libraire RNAseq

Figure issue de Hrdlickova et al., 2017, modifiée par Cimarosti 2022

De ce fait, plus un gène est grand, plus il génère de *reads* (Oshlack and Wakefield, 2009), ce qui induit un biais dans la mesure d'expression des gènes. Pour corriger ce biais inhérent à la technique, on peut diviser le nombre de transcrits normalisés par million (TPM) par la longueur totale en kilobases des exons du gène (Mortazavi et al., 2008). La valeur obtenue est donnée en *reads per kilobase of transcript per million reads mapped* « RPKM ». Dans le cas où la librairie utilise des UMI fixés à l'extrémité 3' de l'ADNc, un r*ead* peut être considéré comme un ARNm. On parle alors de

séquençage *3'.* Pour comparer l'expression de gènes entre différents échantillons, on peut diviser le nombre de UMI par le nombre total de *reads* de la librairie et multiplier par un million. On obtient ainsi des UPM (UMI par million). Le séquençage 3' a en partie été développé pour surmonter le biais dû à la taille des transcrits en séquençage *full lenght*, et proposer une méthode de séquençage quantitative. En revanche, en ne séquençant que quelques dizaines de bases en 3', on perd l'information sur les éventuels épissages alternatifs des gènes transcrits. Les premières techniques de RNAseq « classiques », aussi appelées « *bulk* » reposent sur l'analyse moyennée de tous les ARNm présents dans un échantillon composé en général de plusieurs milliers de cellules (ex : lysats tissulaires, cellules en culture …). C'est une technique d'analyse particulièrement adaptée à l'étude de changements transcriptomiques globaux entre deux conditions, ou pour obtenir rapidement et à moindre coût le profil d'expression d'un tissu mal caractérisé. Cependant elle n'est pas adaptée à l'étude d'échantillons contenant peu d'ARNm, peu de cellules, ou pour étudier l'hétérogénéité de tissus non-homogènes, c'est-à-dire composés de plusieurs types cellulaires différents (Li and Wang, 2021).

C'est pour répondre à ces situations spécifiques que le séquençage ARN en cellule unique (scRNAseq) a été développé. Le premier article démontrant la faisabilité de cette technique a porté sur l'analyse de blastomères uniques de souris (Tang et al., 2009). L'utilisation du scRNAseq est en effet particulièrement adaptée à l'étude des embryons de mammifères qui sont composés de peu de cellules, parfois différentes les unes des autres, et au profil d'expression changeant au cours du développement.

Sa mise au point a requis de surmonter certains défis techniques inhérents à la méthode notamment :

- l'isolation des cellules uniques. Avant une expérience de scRNAseq, chaque cellule doit en effet être individualisée de son tissu d'origine.

- la faible quantité d'ARNm dans une cellule, et son expression pulsatile. La méthode de séquençage doit être suffisamment sensible pour capturer la faible quantité d'ARN présent dans une seule cellule. De plus, l'ARNm d'un gène est exprimé de manière pulsatile et non en continu. (Jiang et al., 2022; Raj et al., 2006; Suter et al., 2011). Un gène non exprimé à un instant t dans une cellule est peut-être actif et nécessaire dans cette même cellule à un autre moment.

De part ces singularités techniques, les analyses de scRNAseq sont confrontées à une « inflation des zéros », c'est-à-dire ne pas détecter un grand nombre de gènes. Si les données de RNAseq *bulk* ont une proportion de zéros de l'ordre de 10-40%, elle peut monter à 90% dans les analyses scRNAseq (Jiang et al., 2022).

L'apport de nouvelles connaissances et d'outils bio-informatiques plus modernes au fil des années a été déterminant dans l'étude de ce type de données. Ces nouveaux outils peuvent notamment être utilisés pour réanalyser d'anciens jeux de données scRNAseq afin d'en extraire davantage d'informations. Cette stratégie est à la base des études d'analyse par scRNAseq du développement embryonnaire humain que nous allons aborder dans les prochaines lignes.

Le scRNAseq au service de la biologie du développement pré-implantatoire

La première analyse de scRNAseq d'embryons humains fut publiée en 2013 (Yan et al., 2013). Le jeu de données contient 124 cellules provenant d'ovocytes (n=3 cellules), d'embryons entre le stade zygote et blastocyste éclos à jour 6 (n=87 cellules), et des hESC (n=34 cellules). Les auteurs décrivent les trois lignées EPI, TE, PrE de l'embryon humain à partir d'une liste de gènes provenant de données murines et humaines (Assou et al., 2012; Guo et al., 2010). Leur analyse non supervisée discrimine le TE mural et polaire, suggérant que ces derniers ont des transcriptomes différents. Ils

comparent également l'EPI humain et les hESC et montrent des différences significatives entre les deux. Le hESC expriment plus que l'EPI certains gènes associés à la pluripotence comme *SALL2, TCF3, ZIC2* et *ZIC3* et les gènes de la famille *ID* (*ID1, ID2, ID3*). Cette analyse soutient l'existence chez l'homme de deux états pluripotents distincts comme chez la souris, l'un préimplantatoire dit « naïf » ou « *naive* », et l'autre post-implantatoire dit « amorcé » ou « *primed* » (Brons et al., 2007; Nichols and Smith, 2009; Tesar et al., 2007).

Bien que cet article jette les bases des futures analyses en scRNAseq humaines, il souffre de certaines limites. Par exemple, les blastocystes analysés sont uniquement des blastocystes éclos à J6. Or, le stade blastocyste couvre un ensemble de morphologies, et donc potentiellement de transcriptomes, allant de J5 à J7. En l'absence de blastocystes J5, les auteurs ne peuvent d'ailleurs pas reconstituer de continuum transcriptomique entre les cellules de morulas séquencées à J4 et celles des blastocystes à J6. Le problème le plus notable de cet article porte sur la stratégie d'analyse des cellules. Les auteurs ont choisi d'inclure dans leur analyse des cellules contenant peu d'ARNm, et des gènes avec un niveau d'expression fixé à RPKM > 0.1. Ce seuil d'expression bas est susceptible de fausser l'analyse en considérant des gènes n'ayant probablement aucune fonction biologique (Hebenstreit et al., 2011).

Ce jeu de données sera réanalysé avec un seuil RPKM >5 et complété par 30 cellules de blastocystes afin de mieux caractériser l'EPI et le PrE (Blakeley et al., 2015). Cette nouvelle étude met en lumière certains gènes exclusifs à l'EPI humain comme le facteur de transcription KLF17, initialement exprimé au stade morula puis restreint à l'EPI du blastocyste. Les auteurs montrent également l'expression spécifique dans l'EPI d'éléments de la voie TGF-β comme *NODAL, GDF3, TGFBR1/ALK5, LEFTY1, SMAD2, SMAD4* et *TDGF1*, faisant écho aux hESC dont la culture requiert l'activation de la voie TGF-β (Bertero et al., 2015; James et al., 2005; Vallier et al., 2005). En

inhibant la voie TGFB-β entre J3 et J5, les embryons humains blastulent normalement mais n'expriment plus les facteurs de transcription NANOG, marqueur de l'EPI, ainsi que SOX17, marqueur du PrE. Cela suggère que la voie TGF-β est un régulateur important de la seconde spécification chez l'Homme, contrairement à chez la souris où l'inhibition de TGF-β n'affecte ni l'expression ni la localisation de NANOG, de OCT4, ou de SOX17 (Blakeley et al., 2015).

Finalement, en 2016, Petropoulos et collègues publient un jeu de données de scRNAseq d'embryons humains préimplantatoires incluant plus de 1500 cellules uniques provenant d'embryons du stade 8-cellules (J3) jusqu'au blastocyste éclos (J7), stade précédant l'implantation (Petropoulos et al., 2016). Cette grande quantité de cellules permet non seulement de caractériser les stades développementaux manquants (jeune blastocyste J5 et blastocyste J7), mais surtout de mieux définir les stades préalablement étudiés. En effet, en augmentant drastiquement le nombre de cellules séquencées par stade, on limite les variabilités techniques et biologiques inhérentes à l'analyse en scRNAseq. Cette stratégie permet de reconstituer un continuum transcriptomique de l'embryon humain en développement, et notamment de couvrir la transition entre les stades morula et blastocyste. Pour capturer l'émergence des lignées EPI, TE, PrE, les auteurs ont recours à une analyse de pseudotemps. Le pseudotemps permet de visualiser l'avancement transcriptomique au sein d'un échantillon en développement à partir du transcriptome des cellules uniques qui le composent (Saelens et al., 2019; Van den Berge et al., 2020). Ils en concluent que les trois lignées émergent simultanément au stade blastocyste J5. Dans les lignées TE tardives à J6-J7, ils observent quelques centaines de gènes différentiellement exprimés entre deux sous-populations TE murales et polaires, confirmant ainsi les observations de Yan et al. en 2013.

2) La deuxième semaine de développement

Aux alentours du 7ème jour de développement, l'embryon humain s'implante dans l'utérus et initie sa deuxième semaine de développement (Niakan et al., 2012). Bien que la grossesse ait débuté, il n'est pas encore possible de la suivre par échographie : l'embryon mesurant moins d'1mm est trop petit. Cette période complexe à étudier est appelée « la boite noire du développement humain » (Goedel and Lanner, 2021). Cette boite noire a été partiellement ouverte au cours de la seconde moitié du XXème siècle, notamment par le prélèvement et l'étude de tissus utérins en cours de grossesse.

A – L'étude d'embryons humains *in vivo*

Les premières images d'embryons humains en développement *in vivo* de la deuxième semaine de développement proviennent des travaux de Hertig et Rock (Hertig et al., 1956) (Figure 10). Entre 1938 et 1954, ces deux embryologistes ont effectué 211 hystérectomies et/ou ovariectomies de patientes selon les critères suivants :

- Les patientes sont incapacitées pour cause de « déplacement utérin » ou tumeur bénigne, symptômes pouvant être guéris par hystérectomie
- Leur cycle menstruel est régulier
- Elles « vivent avec leur mari, sont consentantes, et suffisamment intelligentes pour coopérer ».

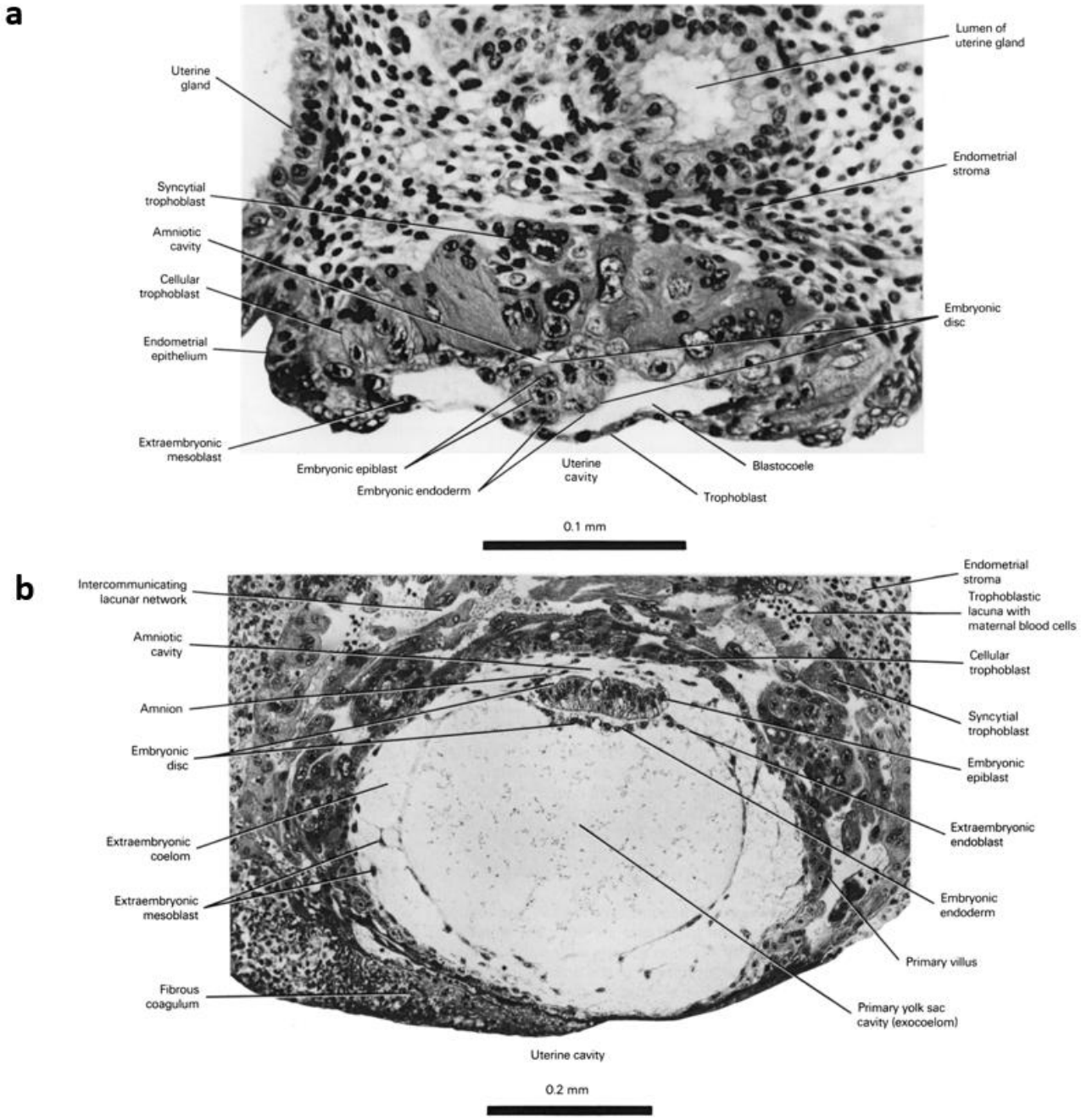


Figure 10 : embryons humains *in vivo* prélevés par hystérectomie

(a) Embryon humain en cours d'implantation à J7. On note que l'embryon s'implante par le TE polaire, et que l'endomètre recouvre progressivement le trophoblaste. La présence de syncytia indique la différenciation du trophoblaste en syncytiotrophoblaste. La cavité amniotique émerge dans l'épiblaste.

(b) Embryon humain à J12. L'embryon possède une structure en trois dimensions. La présence de lacunes à l'interface embryon-endomètre suggère une différenciation du trophoblaste en trophoblaste extra-villeux (EVT). L'EPI s'est épaissi et forme avec le PrE le disque bilaminaire. La grande cavité visible au centre de l'embryon est la cavité vitelline formée à l'intérieur du PrE. Images issues de Hertig and Rock, 1949, Hertig et al., 1956, et modifiées par Gasser RF, 1975.

Il était demandé à ces patientes d'avoir un rapport sexuel avant l'opération et d'en noter la date, permettant ainsi de récupérer les embryons à différents stades de développement. Hertig et Rock ont obtenu 34 embryons entre J2 et J17 (8 flottants dans les trompes ou dans l'utérus, 26 implantés). Les embryons ont été colorés à l'hématoxyline et éosine pour marquer les noyaux et le cytoplasme respectivement, permettant une étude morphologique des échantillons.

A cette époque, le recueil de ces embryons s'est fait dans un cadre éthique incompatible avec nos standards actuels. Pour continuer à étudier les mécanismes d'implantation de l'embryon, des techniques de culture *in vitro* ont dû être développées.

B – Implantation *in vitro* de l'embryon humain

La Loi des 14 jours

Dans les années 1980, en plein essor de la FIV humaine et de la recherche sur l'embryon humain, la Couronne Britannique commanda un rapport pour examiner « les implications sociales, éthiques, et légales des développements potentiels en assistance médicale à la procréation » (Warnock, M. et al., 1984).

Ce rapport avait pour objectif de renseigner le législateur sur ce domaine d'expertise récent, et de lui dresser une liste de recommandations à intégrer aux futures lois. Il y est suggéré d'accorder un statut légal à l'embryon humain qui doit être « protégé ». Il est aussi noté que toute recherche sur l'embryon humain devrait être soumise à autorisation, et que les membres du couple devraient donner leur accord pour ce type de recherche. Une de ces recommandations anticipe la probabilité future de

l'ectogenèse, soit la gestation dans un utérus artificiel. Elle suggère alors d'interdire toute culture *in vitro* d'embryons humains au-delà de 14 jours suivant la fécondation. Ce délai de 14 jours correspond au jour approximatif de la formation de la ligne primitive, un épaississement de l'épiblaste qui a lieu pendant la gastrulation. La gastrulation est la mise en place depuis l'épiblaste des trois feuillets embryonnaires ectoderme, mesoderme, endoderme, à l'origine de tous les types cellulaires dont les premières structures nerveuses (Hurlbut et al., 2017). Cette limite de 14 jours est arbitraire, elle fut choisie comme un compromis éthique, social, et religieux entre les sociétés savantes consultées pour la réalisation du rapport Warnock. Elle est communément désignée comme la *14 days rule*.

Cette règle des 14 jours a depuis intégré la Loi de plusieurs pays comme le Royaume-Uni, l'Espagne, l'Australie ou la France (Hyun et al., 2016).

L'auto-organisation de l'embryon humain postimplantatoire

Les premières méthodes de culture postimplantatoire utilisaient des cellules endométriales primaires ou cancéreuses sur lesquelles était déposé l'embryon. Dans ces conditions, l'embryon est capable d'initier certaines étapes de l'implantation comme l'apposition (contact avec l'endomètre), l'attachement (établissement d'une liaison protéique avec l'endomètre), et l'invasion (pénétration de l'embryon dans l'endomètre) (Ojosnegros et al., 2021). Ces techniques de co-culture embryonnaire sont particulièrement adaptées pour l'étude de l'interaction embryon-endomètre, mais ne permettent pas de cultiver les embryons humains plus de 48h-72h après leur attachement, soit J9 maximum (Teklenburg and Macklon, 2009).

Etonnement, il semblerait que l'embryon humain soit capable de survivre et de s'auto-organiser après J7 indépendamment de tout support cellulaire. Avec un milieu de culture adapté (dit « IVC »), l'embryon humain est capable de s'attacher à une plaque de culture *in vitro* sans cellules ni matrice, survivre, croître, et former des structures morphologiques spécifiques de la deuxième semaine de grossesse comme la cavité amniotique et la cavité vitelline (Deglincerti et al., 2016; Shahbazi et al., 2016)(Figure11).

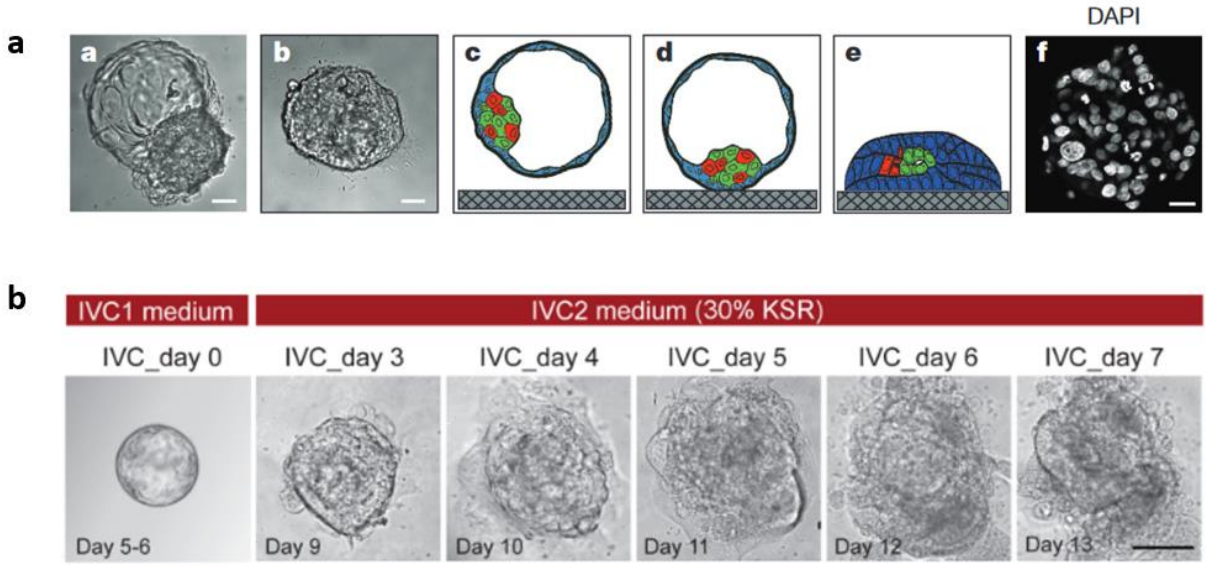


Figure 11 : culture postimplantatoire d'embryons humains en auto-organisation

(a) Panel supérieur : image et schéma de l'attachement de l'embryon entre J7 et J8. L'embryon s'attache à la plaque de culture par le TE polaire, du côté de l'ICM (Deglincerti et al., 2016). Une fois l'embryon attaché, la cavité du blastocyste s'affaisse complètement et l'embryon forme un amas cellulaire dense. Figure issue de Deglincerti et al., 2016.

(b) Panel inférieur : développement de l'embryon entre J5 et J13. Entre J5 et J7, l'embryon humain est cultivé avec du milieu IVC1 contenant 20% de sérum de veau fœtal (FBS). A partir de J7 quand l'embryon est attaché, le FBS est remplacé par 30% de Knock-out Serum Replacement (KSR). Figure issue de Shahbazi et al., 2016.

Comme le milieu IVC est défini et que la culture ne requiert pas de cellules nourricières, cette technique permet une meilleure reproductibilité et un contrôle plus fin des conditions de culture que les méthodes de co-culture. Elle permet la survie et la culture d'embryons humains jusqu'à J13.

Il est à noter que l'intégralité de la culture IVC, dès le stade blastocyste J5 et pendant la postimplantation, s'effectue à 37°C avec un taux d'oxygène (O2) de 21%, soit en normoxie. Le fait que l'embryon humain survive en normoxie entre J5 et J7 suggère que l'hypoxie n'est pas strictement nécessaire au bon développement embryonnaire à partir du stade blastocyste. Les auteurs montrent également qu'aucun embryon cultivé en hypoxie dès J5 ne maintenait son EPI à J9-10 en postimplantation, contrairement aux embryons cultivés en normoxie. Cette observation suggère que le développement embryonnaire requiert une concentration d'O2 supérieure en postimplantation, probablement à cause de l'augmentation de la taille de l'embryon et la densité cellulaire qui limite la pénétration d'O2 dans les cellules au centre de l'embryon (Shahbazi et al., 2016).

Les informations dont nous disposons chez l'Homme suggèrent que la spécification de l'EPI et du PrE n'est pas achevée dans toutes les cellules avant l'implantation. Deglincerti et al. observent en effet des colocalisations OCT4+/GATA6+ et OCT4+/NANOG+ dans l'ICM de blastocyste J6. Les cellules OCT4+/NANOG+ correspondent probablement à de l'EPI spécifié là où les cellules OCT4+/GATA6+ correspondent à du PrE non spécifié. En utilisant ces marqueurs, ils montrent que la ségrégation spatiale de l'EPI et des futures cellules du PrE n'est pas en place dans l'ICM des blastocystes à J6, mais est visible à J8 dans l'embryon attaché. A J8, les cellules de l'EPI OCT4+ sont regroupées les unes à côté des autres sous la forme d'une sphère. Les cellules GATA6+/OCT4- du PrE sont de part et d'autre des cellules de l'EPI.

Entre J8 et J9, l'épiblaste se polarise en son centre et forme la cavité amniotique. Cette cavité est caractérisée par l'expression apicale d'aPKC, et par un fort signal membranaire en marquage phalloïdine. Chez la souris, l'ouverture de la cavité

amniotique est promue par l'action combinée dans l'EPI d'OCT4 et OTX2 qui activent l'expression de la podocalyxine (PODXL) (Shahbazi, 2020). PODXL est également présent dans la cavité amniotique chez l'Homme (Molè et al., 2021; Xiang et al., 2020), mais son mécanisme d'expression est probablement différent de chez la souris. Chez la souris, OTX2 est un régulateur de l'intégrité de l'EPI murin qui se fixe sur les promoteurs d'*Oct4*, *Sox2*, et *Nanog* (Acampora et al., 2016). Chez l'humain, OTX2 n'est pas détecté dans l'EPI mais dans les cellules GATA6+ du PrE postimplantatoire (Zhou et al., 2019). Sa dynamique d'expression n'est pas documentée dans le blastocyste humain.

Il a été montré que la modulation de la voie FGF n'avait pas d'incidence sur le nombre de cellules de l'EPI et du PrE pré-implantatoire (Kuijk et al., 2012; Roode et al., 2012). En revanche, l'embryon humain postimplantatoire exprime dans son EPI divers ligands de la voie FGF tels que FGF2, FGF4, et plus modestement FGF8, FGF17, FGF18 et FGF19. Les lignées EPI, TE, et PrE en expriment les récepteurs associés (FGFR1/2/3/4). En inhibant ces voies FGF entre J6 et J8, une réduction du nombre de cellules peut être observée dans l'EPI et le PrE, et dans une moindre mesure le TE (Molè et al., 2021). Ces observations témoignent d'un rôle régulateur de la voie FGF dans le maintien des lignées postimplantatoires humaines.

En utilisant des plaques de culture à faible attachement et du matrigel, il est possible de cultiver l'embryon humain postimplantatoire en trois dimensions, et ainsi conserver sa structure en volume comme *in vivo* (Xiang et al., 2020). Les auteurs y documentent notamment la présence des trois sous-populations trophoblastiques à l'origine du placenta : le cytotrophoblaste (CT), le syncytiotrophoblaste (ST), et le trophoblaste extra-villeux (EVT) (James et al., 2012). Le CT est détecté entre J8-J9 et est mononucléé, proche de l'EPI, et exprime les marqueurs CK7/TEAD4/CDH1. Entre J9 et J12, le CT se différencie en ST, multinucléé et exprimant la β-HCG, puis en EVT

exprimant la HLA-G. Les auteurs montrent que les voies de signalisation actives dans les cellules EVT sont liées à l'immunité et à l'angiogenèse. Ces informations corroborent une étude récente montrant que l'EVT de placenta de premier trimestre joue un rôle crucial dans la modulation de la réponse immunitaire et le remodelage des artères spiralées à l'interface fœto-maternelle (Vento-Tormo et al., 2018).

Les protocoles de culture *in vitro* d'embryons humains sans cellules ont permis des avancées notables dans notre compréhension du développement. Ils permettent la survie des embryons jusqu'à J13 dans des conditions reproductibles, l'imagerie aisée des marqueurs des lignées embryonnaires, et modélisent certaines grandes étapes du développement *in vivo*. La plupart des immunofluorescences et jeux de données de scRNAseq aujourd'hui disponibles sont issus d'embryons cultivés dans ces conditions. Bien qu'ils soient moins adaptés à l'étude de l'interaction embryon-endomètre, ils constituent une méthode pertinente pour l'étude de la deuxième semaine de développement.

III- MODELISATION DE L'EMBRYON HUMAIN PAR LES CELLULES SOUCHES

L'utilisation d'embryons humains dans le cadre de projets de recherche demeure un challenge tant d'un point de vue réglementaire que logistique. Leur nombre limité, leur qualité variable, et la variabilité entre les échantillons ont amené les chercheurs à développer des modèles d'étude alternatifs de l'embryon.

Les cellules souches sont cultivables indéfiniment en laboratoire, et reflètent à des degrés variables les lignées embryonnaires présentes aux différents stades du développement. Elles peuvent être modifiées génétiquement, et permettre une étude mécanistique de la différenciation des lignées embryonnaires.

1) Les cellules souches embryonnaires (ES)

Les cellules souches embryonnaires humaines (hES) sont des cellules dérivées de l'épiblaste de l'ICM de l'embryon. Ces cellules sont pluripotentes (hPSC), c'est-à-dire capables de se différencier dans les trois feuillets embryonnaires (ectoderme, endoderme, mésoderme), et par extension en tout type cellulaire adulte, mais sans pouvoir former les annexes embryonnaires comme le placenta ou le sac vitellin.

Des lignées de hES ont été dérivées chez l'Homme à partir de blastocyste en 1998 (Thomson et al., 1998), 17 ans après la souris (Evans and Kaufman, 1981; Martin, 1981). Il est également possible d'obtenir des cellules pluripotentes par reprogrammation de cellules somatiques par surexpression des facteurs de pluripotence *OCT4, SOX2, KLF4,et c-MYC* (OSKM) (Takahashi et al., 2007). Ces

dernières sont appelées cellules souches pluripotentes induites (hiPSC), et sont considérées comme des équivalents des hES (Liang and Zhang, 2013; Narsinh et al., 2011).

La culture des hPSC (hiPSC et hES) nécessite l'activité des voies FGF2 et Activine A/transforming growth factor β-1 (TGFβ1) (James et al., 2005; Weinberger et al., 2016). En revanche, la culture des hPSC ne dépend pas du Leukemia inhibitory factor (LIF) qui stabilise la culture des ES murines (mES)(Smith, 2001). Bien que toutes deux dérivées du blastocyste, les cellules souches embryonnaires humaines diffèrent de leurs analogues murins. Les hPSC ont un ADN plus méthylé, expriment plus faiblement des gènes associés à la préimplantation comme *KLF17*, et les lignées femelles ont un chromosome X inactif (Mekhoubad et al., 2012; Smith et al., 2014). Ces caractéristiques sont finalement proches des cellules EpiSC, des cellules souches pluripotentes murines dérivées de l'EPI murin postimplantatoire (Brons et al., 2007; Tesar et al., 2007). Pour ces raisons, les hPSC sont considérées comme un modèle de l'EPI postimplantatoire dit amorcé ou *primed* (Nichols and Smith, 2009).

Les hPSC pré-implantatoires dites *naives* (hNPSC) peuvent être obtenues par conversion de hPSC (Hanna et al., 2010; Takashima et al., 2014; Theunissen et al., 2014) ou reprogrammation de cellules somatiques (Kilens al., 2018). Leur culture nécessite la présence de LIF et de 2i, un mélange d'inhibiteurs des voies MAPK et GSK3β qui empêche la différenciation des ES murines ou de rats (Buehr et al., 2008; Ying et al., 2008). Les hNPSC se distinguent des hPSC par un métabolisme plus actif, une hypométhylation globale de leur génome, et l'activation des deux chromosomes X dans les lignées femelles (Figure 12)(Kilens et al. 2018).

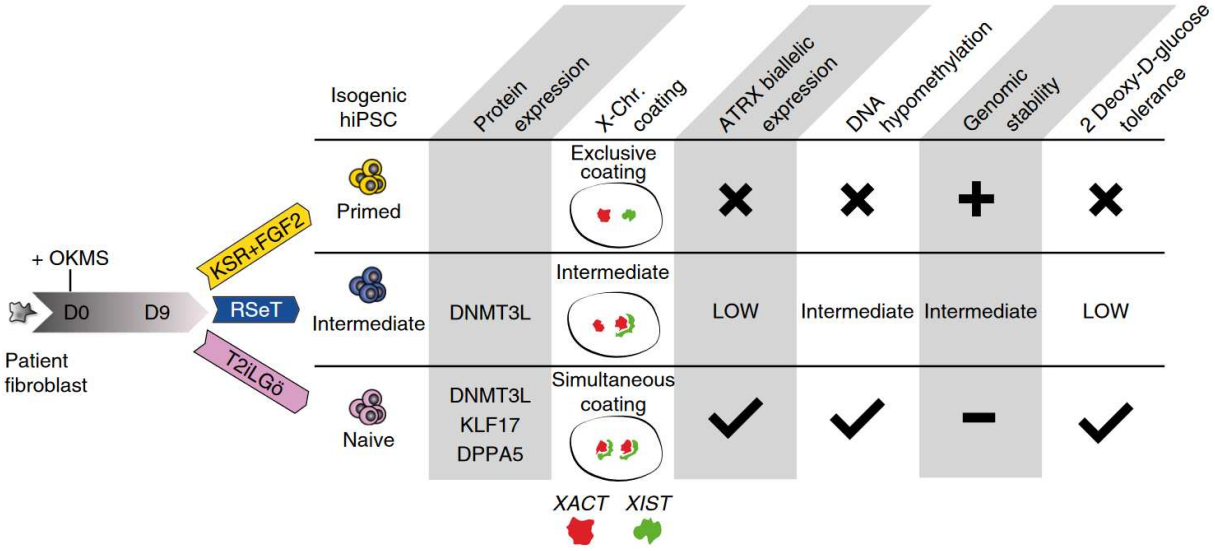


Figure 12 : caractéristiques des cellules PSC primed, intermédiaires et naives

Les fibroblastes sont reprogrammés par l'expression des facteurs de pluripotence OSKM. A J9 de la culture, les cellules sont transférées en milieu KSR+FGF2, RSeT, ou T2iLGö pour obtenir respectivement des cellules pluripotentes primed (postimplantatoires), des cellules pluripotentes intermédiaires, et des cellules pluripotentes naïves. Figure issue de Kilens et al. 2018.

2) Les cellules souches trophoblastiques et la lignée placentaire

Les cellules trophoblastiques sont à l'origine de la lignée placentaire qui permet l'apport de nutriments et d'oxygène au foetus, la sécrétion d'hormones, et le remodelage du tissu utérin maternel (Turco and Moffett, 2019; Viukov et al., 2022). *In vivo*, le compartiment trophoblastique est composé de trois types cellulaires. Les cellules du cytotrophoblaste (CT) sont des cellules indifférenciées prolifératives. Elles constituent le réservoir auto-réplicatif pouvant se différencier en syncytiotrophoblaste (ST) et trophoblaste extra-villeux (EVT).

Les cellules souches trophoblastiques murines (mTSC) modélisant le CT ont été dérivées en 1998 à partir d'embryons postimplantatoires. Elles sont maintenues en culture *in vitro* par l'action combinée de FGF4 et TGFb1/Activin, et peuvent se différencier en ST et EVT (Tanaka et al., 1998). Il a fallu attendre 20 ans pour dériver des TSC humaines (hTSC) à partir de blastocyste et de placenta de premier trimestre (Okae et al., 2018). Les hTSC se maintiennent en culture grâce à l'activation de la voie

Wnt, EGF, l'inhibition des histones acétylases HDAC, de la voie ROCK et de la voie TGF-β. Les hTSC expriment des protéines caractéristiques du TE comme GATA3, TP63, ou TEAD4. Elles peuvent se différencier *in vitro* en ST et EVT. Les hTSC sont un modèle d'étude prometteur des pathologies placentaires comme la prééclampsie, les fausses couches, ou les cancers placentaires. Cependant, l'accès aux embryons humains et la dérivation de leur trophoblaste, ou l'obtention de placenta de premier trimestre, ne sont pas aisés.

Récemment, il a été montré que l'obtention de hTSC est possible par conversion de hNPSC et hPSC ou par reprogrammation de cellules somatiques avec les facteurs OSKM (Castel et al., 2020; Cinkornpumin et al., 2020; Dong et al., 2020; Guo et al., 2021; Io et al., 2021; Liu et al., 2020; Viukov et al., 2022). Ces hTSC ont un transcriptome proche du trophoblaste au 8ème jour de développement (Castel et al., 2020; Zhou et al., 2019). L'obtention de hTSC depuis les h(N)PSC, et notamment par reprogrammation OSKM, suggère une plasticité entre les compartiments pluripotents et trophoblastiques. En revanche, il est difficile de déterminer si cette capacité est un artefact de culture *in vitro* ou si les cellules pluripotentes de l'épiblaste *in vivo* sont aussi capables de se différencier en trophoblaste.

3) Le blastoïde : un modèle d'embryon créé sans gamètes

Cultivés isolés les uns des autres, les modèles cellulaires des lignées embryonnaires ne permettent pas d'étudier le développement de l'embryon dans son ensemble, et notamment les mécanismes de communication entre les lignées, ou l'implantation de l'embryon dans l'utérus.

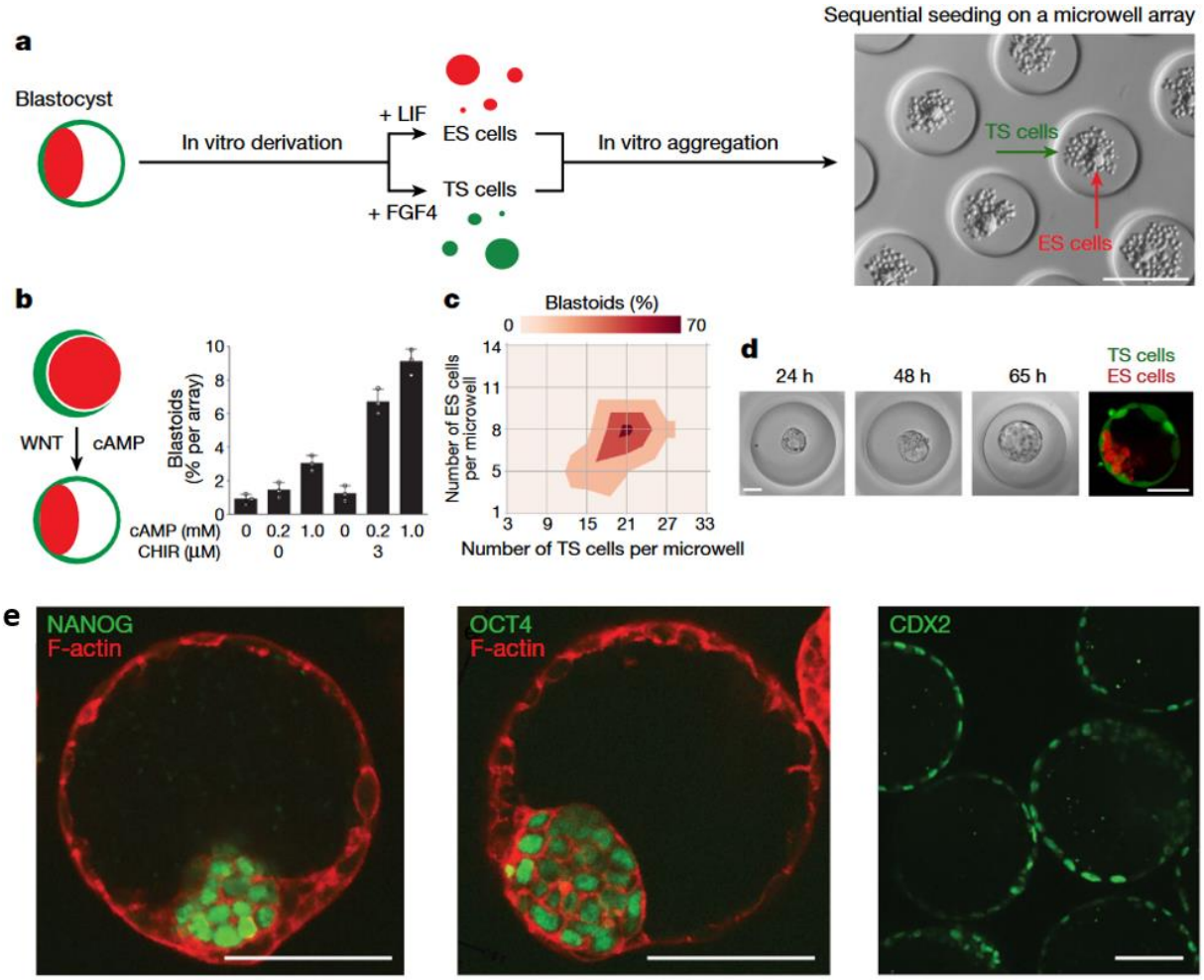


Figure 13 : génération d'un blastoïde murin

Les blastoïdes sont issus d'agrégats de cellules ES et TSC dérivées de blastocystes (a). Ces agrégats sont cultivés en présence d'activateur de la voie WNT et d'AMP cyclique (b). Avec un ratio de 21 cellules TSC pour 8 cellules ES, la formation du blastoïde survient dans 70% des puits (c). Les blastoïdes cavitent en 24h (d) et expriment le facteur de pluripotence NANOG dans leur ICM et CDX2 dans les cellules externes délimitant la cavité. Figure adaptée de (Rivron et al., 2018)

En 2018, une équipe est parvenue à générer un modèle de blastocyste murin en agrégant des cellules pluripotentes et trophoblastiques murines (Rivron et al., 2018). Ce modèle appelé « blastoïde » cavite en 24h et forme une structure sphérique creuse comportant un amas cellulaire interne ressemblant à un blastocyste (Figure 13). Les cellules externes du blastoïde expriment CDX2 comme le TE, et les cellules de l'amas cellulaire interne expriment les facteurs de pluripotence OCT4 et NANOG comme l'EPI. Quand ces blastoïdes de souris sont transférés dans un utérus de souris pseudo-gestante, ils sont capables d'initier une grossesse. On y observe une decidualisation

de l’endomètre, des anastomoses blastoïde-endomètre, et l’expression locale de ALDH3A1 qui est un marqueur d’implantation chez la souris (McConaha et al., 2011). En revanche, le blastoïde transféré n’atteint pas le stade fœtal, il n’y a pas d’organisation tissulaire, et aucune naissance n’est possible.

Résultats

La partie résultats qui suit comporte trois publications.

Meistermann, Bruneau, Loubersac et al, représente le cœur de mon travail de thèse (Meistermann et al., 2021). Cet article a permis d'établir la hiérarchie des évènements moléculaires associés au développement péri-implantatoire humain. Ces connaissances nous ont permis d'engager deux collaborations :

- Avec Claudia Gerri et Kathy Niakan, où nous avons exploré les mécanismes de la première spécification chez l'Homme (Gerri et al., 2020).
- Avec Harunobu Kagawa, Alok Javali, Heidar Heidari Khoei et Nicolas Rivron, où nous avons travaillé à l'élaboration du modèle de blastoïde humain (Kagawa et al., 2021).

Ces travaux se lisent dans un ordre logique, les deux premières publications ayant servi de référence à la troisième.

1.1 Manuscrit principal : “Integrated pseudotime analysis of human preimplantation embryo single-cell transcriptomes reveals the dynamics of lineage specification“

1.1.1 Le contexte de l’étude

J’ai rejoint ce projet pendant mon Master 2. J’ai eu l’opportunité d’être formé au travail sur l’embryon humain par les biologistes de la reproduction du service d’AMP du CHU de Nantes, et à l’imagerie confocale. J’ai réalisé la plupart des immunofluorescences de l’article, la culture postimplantatoire d’embryons humains, ai participé à la rédaction de l’article, la mise en forme des figures, et le processus de révision et de soumission aux éditeurs. L’article a été soumis une première fois dans Nature en décembre 2018 pour être finalement publié dans Cell Stem Cell en 2021, soit deux ans et demi de révision. Les révisions nous ont amené à implémenter au fil de l’eau de nouvelles technologies d’analyses bioinformatiques et de culture embryonnaire postimplantatoire.

Notre papier s’inscrit dans un contexte de recherche effervescent sur le développement humain et des modèles cellulaires qui y sont associés. Ce projet a débuté en 2013 par une demande d’autorisation à l’Agence de la biomédecine (ABM) et l’envoi des premiers échantillons pour séquençage début 2014. Dimitri Meistermann, co-premier auteur de ce projet, a débuté l’analyse des données en 2014, à une époque où les outils d’analyses spécifiques au scRNAseq n’étaient pas encore développés.

Nous avons intégré nos données de séquençage à celles de jeux de données préalablement publiés (Blakeley et al., 2015; Petropoulos et al., 2016; Yan et al., 2013). Au cours de nos analyses, nous avons observé que :

- Nos données de morula étaient bien moins hétérogènes que les données de Petropoulos et al.
- Les embryons à J5 de Petropoulos et al. regroupaient des embryons avant et après la spécification des lignées.

Grâce à notre évaluation morphologique clinique, nous avons pu transformer ces observations en hypothèses biologiques. Nous étions sûrs du stade développemental des embryons que nous avions analysés par scRNAseq, notamment pour différencier les morulas des jeunes blastocystes B1-B2 entre J4 et J5.

A J5, une grande variété de morphologies embryonnaires peut être observée (Figure 6). L'annotation embryonnaire clinique nous a permis d'affiner nos analyses en associant précisément un stade morphologique à un transcriptome, et de réannoter les jeux de données publiés précédemment

Nous proposons ici un patron transcriptomique et protéique du développement humain entre J3 et J7. Notre travail a notamment permis de caractériser l'émergence des lignées embryonnaires, ainsi que leur maturation péri-implantatoire.

1.1.2 Manuscrit Meistermann, Bruneau, Loubersac et al., 2021

Meistermann D*, **Bruneau A***, Loubersac S*, Reignier A, Firmin J, François-Campion V, Kilens S, Lelièvre Y, Lammers J, Feyeux M, Hulin P, Nedellec S, Bretin B, Castel G, Allègre N, Covin S, Bihouée A, Soumillon M, Mikkelsen T, Barrière P, Chazaud C, Chappell J, Pasque V, Bourdon J, Fréour T, David L. 2021. Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. Cell Stem Cell S1934590921001855. doi:10.1016/j.stem.2021.04.027

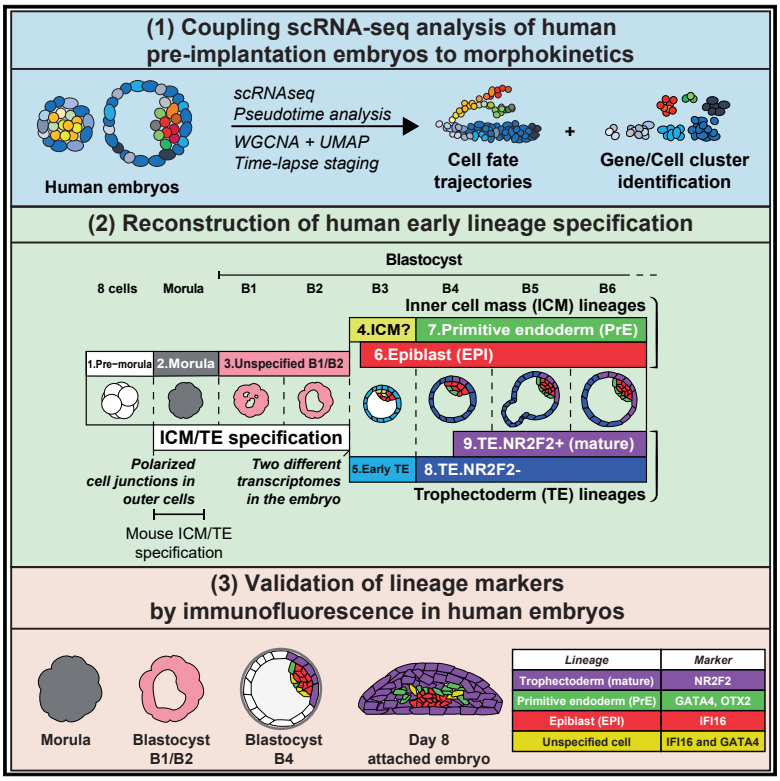
*Contribution égale

Cell Stem Cell

Resource

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

Graphical abstract



Highlights

- Distinct trophectoderm/epiblast signatures arise at the B2–B3 blastocyst stages
- IFI16 is broadly expressed in the ICM and then restricted to epiblast after implantation
- NR2F2 arises from the polar TE in late blastocysts and then spreads to all TE cells
- Transcriptomics analysis suggests that PrE emerges from EPI in late blastocysts

Authors

Dimitri Meistermann, Alexandre Bruneau, Sophie Loubersac, ..., Jérémie Bourdon, Thomas Fréour, Laurent David

Correspondence

thomas.freour@chu-nantes.fr (T.F.), laurent.david@univ-nantes.fr (L.D.)

In brief

Meistermann et al. use scRNA-seq, pseudotime analysis, and immunofluorescence validation to reconstruct the sequence of molecular events occurring in time-lapse-staged human preimplantation embryos. Their findings clarify specification, timing, and cell transcriptomes across early stages of human development.

Meistermann et al., 2021, Cell Stem Cell *28*, 1625–1640
September 2, 2021
https://doi.org/10.1016/j.stem.2021.04.027

CellPress

Cell Stem Cell

CellPress

Resource

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

Dimitri Meistermann,[1,3,13] Alexandre Bruneau,[1,13] Sophie Loubersac,[1,2,13] 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,[6,7,8,11] Tarjei Mikkelsen,[6,7,8,12] Paul Barrière,[1,2] Claire Chazaud,[9] Joel Chappell,[10] Vincent Pasque,[10] Jérémie Bourdon,[3] Thomas Fréour,[1,2,*] and Laurent David[1,4,14,*]

[1]Université de Nantes, CHU Nantes, INSERM, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ITUN, 44000 Nantes, France
[2]CHU Nantes, Université de Nantes, Service de Biologie de la Reproduction, 44000 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
[7]Broad Institute, Cambridge, MA 02142, USA
[8]Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA
[9]GReD Laboratory, Université Clermont Auvergne, CNRS, INSERM, Faculté de Médecine, CRBC, 63000 Clermont-Ferrand, France
[10]KU Leuven – University of Leuven, Department of Development and Regeneration, Institute for Single Cell Omics, Leuven Stem Cell Institute, Herestraat 49, 3000 Leuven, Belgium
[11]Present address: Flexomics LLC, 38 Wareham Street, Fl 3, Boston, MA 02118, USA
[12]Present address: 10x Genomics, 7068 Koll Center Pkwy #401, Pleasanton, CA 94566, USA
[13]These authors contributed equally
[14]Lead contact
*Correspondence: thomas.freour@chu-nantes.fr (T.F.), laurent.david@univ-nantes.fr (L.D.)
https://doi.org/10.1016/j.stem.2021.04.027

SUMMARY

Understanding lineage specification during human pre-implantation development is a gateway to improving assisted reproductive technologies and stem cell research. Here we employ pseudotime analysis of single-cell RNA sequencing (scRNA-seq) data to reconstruct early mouse and human embryo development. Using time-lapse imaging of annotated embryos, we provide an integrated, ordered, and continuous analysis of transcriptomics changes throughout human development. We reveal that human trophectoderm/inner cell mass transcriptomes diverge at the transition from the B2 to the B3 blastocyst stage, just before blastocyst expansion. We explore the dynamics of the 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. Our study pinpoints the precise timing of lineage specification events in the human embryo and identifies transcriptomics hallmarks and cell fate markers.

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. Concomitant with morphological changes and growth of the embryo, pluripotent epiblast and trophectoderm cells are established. Because human pluripotency is established during human preimplantation development, a better understanding of early human embryogenesis has a strong effect on research in the field of pluripotent stem cells. Human preimplantation development is also important for fertility treatment and *in vitro* fertilization (IVF) in particular.

The main morphological events 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 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

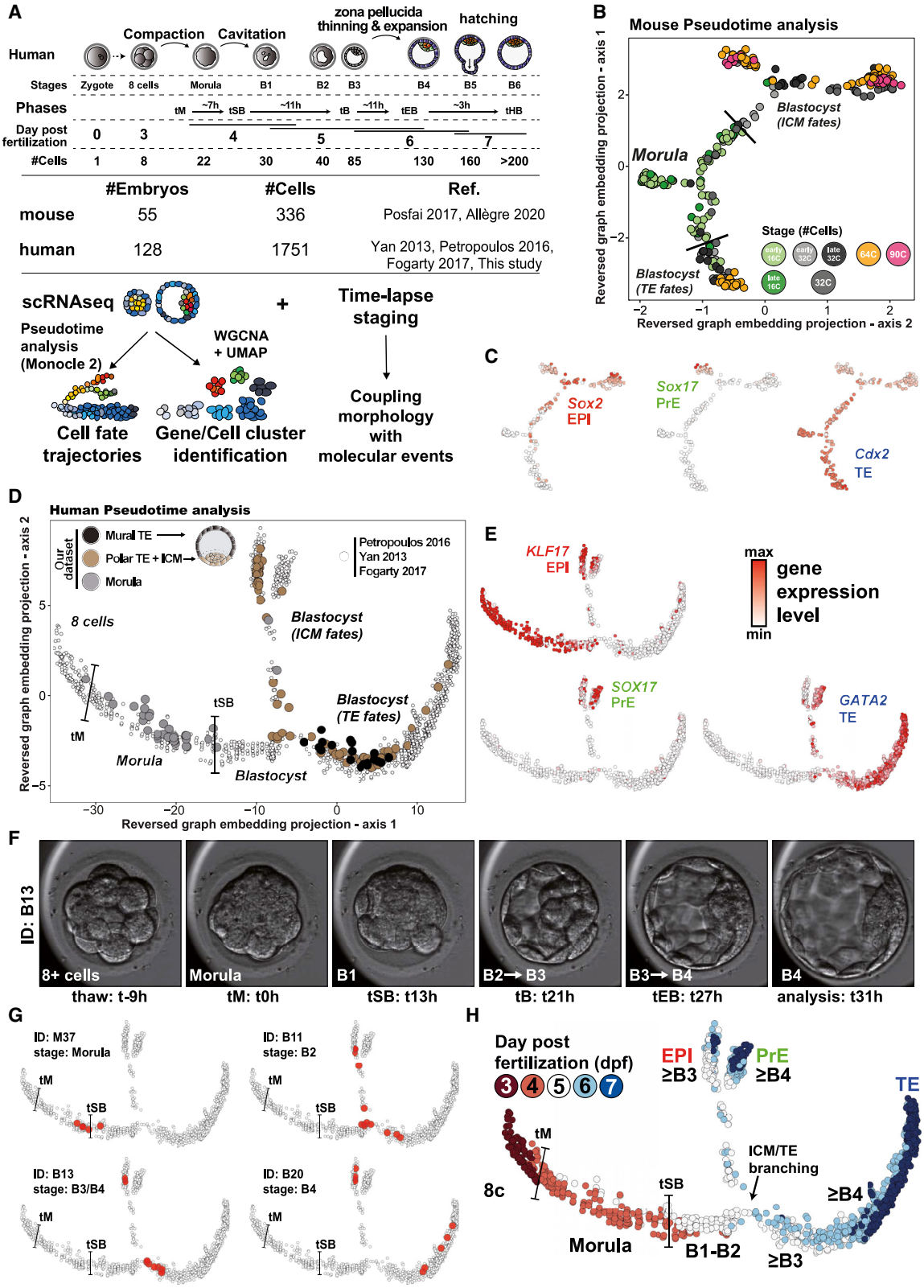


Figure 1. Pseudotime analysis of human and mouse pre-implantation embryo development

(A) Top: Schematic of human pre-implantation development, highlighting the link between developmental stages, morphological phases, dpf, and average number of cells observed. Center: summary of datasets used. Bottom: description of our experimental design. We used 2 mouse single-cell datasets and 4

(legend continued on next page)

(Hirate et al., 2013), and cavitation has been shown to be necessary for proper trophectoderm (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 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), precursors 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 transcriptomics analysis of human pre-implantation embryos suggested 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 analyses 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)

In the 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, whereas non-polarized inner cells engage in an ICM fate.

A recent study showed that polarization/the Hippo pathway was involved in initiation of the TE fate in humans (Gerri et al., 2020). However, little is known regarding the progression of TE; 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 the mouse. Only a 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 is 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 that specific molecular events occur on the polar TE side of the blastocyst; however, the molecular events and underlying mechanisms remain to be identified.

The second cell fate decision, separating ICM cells into EPI and PrE, occurs at the blastocyst stage in the 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 induces 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, whereas it occurs at the 90-cell blastocyst stage in the mouse (Deglincerti et al., 2016). Additionally, fibroblast growth factor (FGF) signaling does not seem to regulate EPI/PrE specification in humans (Kuijk et al., 2012; Roode et al., 2012). These results demonstrate differences from the mouse model. Single-cell RNA sequencing (scRNA-seq) studies have provided atlases of gene expression during human pre-implantation development, but it is still not clear whether and when cell fate decisions take place (Blakeley et al., 2015; Petropoulos et al., 2016; Singh et al., 2019; Yan et al., 2013). How distinct molecular regulators relate to specification of TE, EPI, and PrE lineages in humans remains to be 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, highlighting differences and similarities of mouse and human embryos. Last, we identify proteins that mark EPI, TE, and PrE fate progression. This study is 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 in each species. In the mouse, embryonic days are accurately linked to morphology and molecular events (Figure S1A). In humans, however, *in vitro* culture of embryos shows discrepancies between the timing of culture and morphology. Therefore, for humans, reproductive biologists rely on morphology-based annotation with precise criteria (Figure 1A; Alpha Scientists in Reproductive

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 the activation score of the WGCNA module (module eigengenes) to map cell identities during pre-implantation development. Human embryos used for this study were imaged using a time-lapse microscope.

(B) Projection from the mouse scRNA-seq and single-cell qPCR (scqPCR) samples (Allègre et al., 2019; Posfai et al., 2017) 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 pseudotime: Sox2 (EPI), Cdx2 (TE), and Sox17 (PrE).

(D) Projection of stage and positional information for the samples we analyzed for this study, on the pseudotime. Morula cells (gray) and blastocyst dissection origin of cells are indicated (mural TE, black; polar TE or ICM, brown). tM, morula compaction; tSB, blastulation.

(E) Projection of lineage marker expression levels on the human pseudotime: KLF17 (EPI), SOX17 (PrE), and GATA2 (TE).

(F) Frames from time-lapse microscopy for embryo B13. For each embryo sequenced in this study, its morphokinetics were acquired using time-lapse microscopy. Developmental events include tM, tSB leading to the B1 stage, full blastocyst (tB) at the B3 stage, and expanded blastocyst when the *zona pellucida* thickness is halved (tEB) at the 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 (3–7 dpf) for all samples combined for this study and the result of our refined staging.

Medicine and ESHRE Special Interest Group of Embryology, 2011). The numerous morphological events 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 to better understand human pre-implantation development.

Pseudotime analysis of scRNA-seq 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 most likely path that temporally orders transcriptomics signatures of single cells (Qiu et al., 2017; Figure 1A). To identify lineage signatures, an essential benchmark for stem cell models, we clustered cells based on their transcriptomes. Finally, because we used embryos donated for research, staged carefully using time-lapse monitoring and *zona pellucida* thickness measurements (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 into 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 where cells make their first fate decision. Future cells, from early 32- and late 32-cell morula, occupy one of the two branches toward the ICM and TE fates. Cells from 64- and 90-cell blastocysts are located at the tips of specified branches of the pseudotime (Figure 1B). Enrichment of Sox2, Cdx2, and Sox17 expression allows identification of EPI, TE, and PrE branches, respectively (Frum et al., 2018; Figure 1C). Pseudotime analysis allowed us to reconstruct the developmental paths taken during early mouse embryo development, including the expected TE/ICM branchpoint as well as EPI/PrE bifurcation.

Having established the developmental pathways of early embryo cells in the mouse, we turned our focus to human embryos. We first generated scRNA-seq data of 25 human embryos from morula to B5 blastocysts (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 1,751 cells from 128 embryos. Pseudotime analysis revealed 5 branches, with each branch containing samples from at least 2 datasets (Figure 1D; Figure S2A). Because our embryos were followed by time-lapse microscopy prior to scRNA-seq, we could univocally associate developmental stages for each embryo. Moreover, because 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; Figures S1B–S1E). This revealed 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 associated predominantly 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 analyzed together. We then regrouped embryos based on their position on the pseudotime (Figure S2C). 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 humans, TE and ICM cells become transcriptionally distinct between the B2 and B3 stages, after the beginning of blastocyst cavitation (Figure 1H; Figure S2C). These results were in stark contrast to mouse embryos, where branching between ICM and TE fate occurred earlier, between the 16- and 32-cell morula stage (Figure 1B). Our results suggest that, in human embryos, ICM cells are transcriptionally distinct from TE cells shortly after formation of the blastocyst, unlike in the mouse, where this event takes place earlier.

Defining the hierarchy of gene expression and lineage signature during human pre-implantation development

The next step was to define lineages associated with each cell. Pseudotime analysis is particularly suited to studying 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 lineages. Gene module activation or repression can be summarized by a linear combination of gene expression values called module eigengenes. To produce precise mapping of transcriptome identity, we used module eigengenes from gene modules reduced in two dimensions by a UMAP (uniform manifold approximation and projection) approach (McInnes et al., 2018). Finally, we clustered cells based on their relative position on the UMAP (Ester et al., 1996). This featured 8 major clusters indicative of specific lineages and stages; we qualified cells located between clusters as undefined or "intermediates" (Figure 2A; Figure S3).

Module eigengenes classify EPI, PrE, and TE cells in an unsupervised manner. In contrast to Petropoulos et al. (2016), who used successive rounds of principal-component analysis (PCA) to identify different cell types and loadings to identify the associated markers, our method gives a more confined identification of PrE and 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 embryonic day 5 (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 be considered a proper ICM cluster. However, our analysis highlighted 3 clusters of unspecified cells and subdivided the TE cells into 3 subclusters based on their progression on the pseudotime (Figure 2B). Our unsupervised analysis therefore identified developmental molecular states during pre-implantation development and

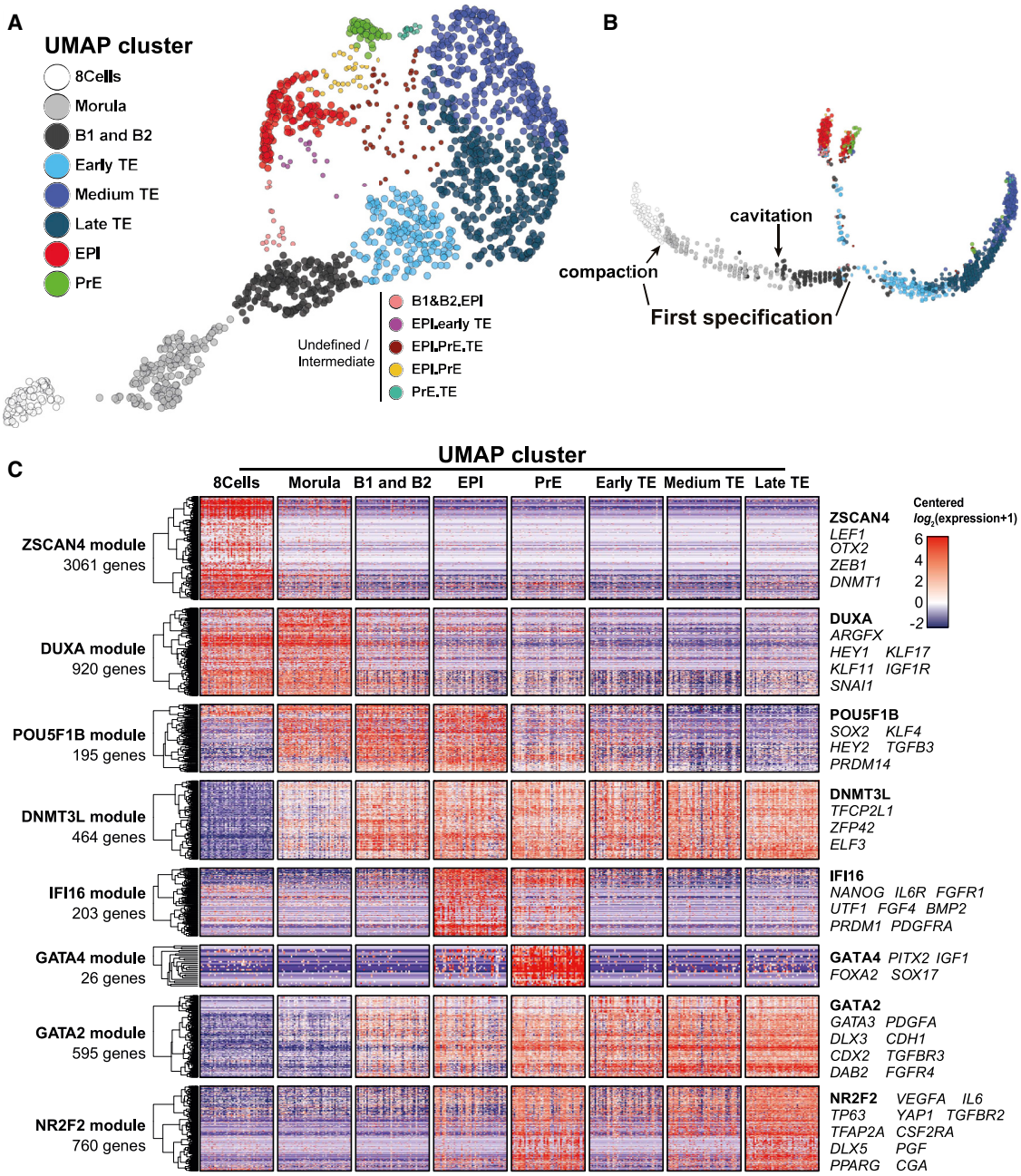


Figure 2. Gene modules define cell fate transcriptional signatures, and pseudotime indicates the hierarchy of events

(A) Dimension reduction (UMAP) of WGCNA module eigengenes, colored by density-based clusters. Small dots indicate a minor cluster composed by the intermediate transcriptome. Medium TE and late TE were segregated by 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 represents a WGCNA module, and each column is a set of 50 cells drawn from a UMAP cluster (A). The height of each row represents the number of genes 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.

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 gene module expression with each cluster of cells obtained with the UMAP (Figure 2C). This revealed unique behavior of 8 modules of genes, named after a representative gene (Figure S3). The 3 clusters of unspecified cells are distinguished based on expression of (1) the ZSCAN4 module, exclusively within 8-cell stage embryos; (2) the DUXA module and onset of the POU5F1B and DNMT3L modules in morula cells; and (3) absence of the ZSCAN4 module, fading of the DUXA modules, and the presence

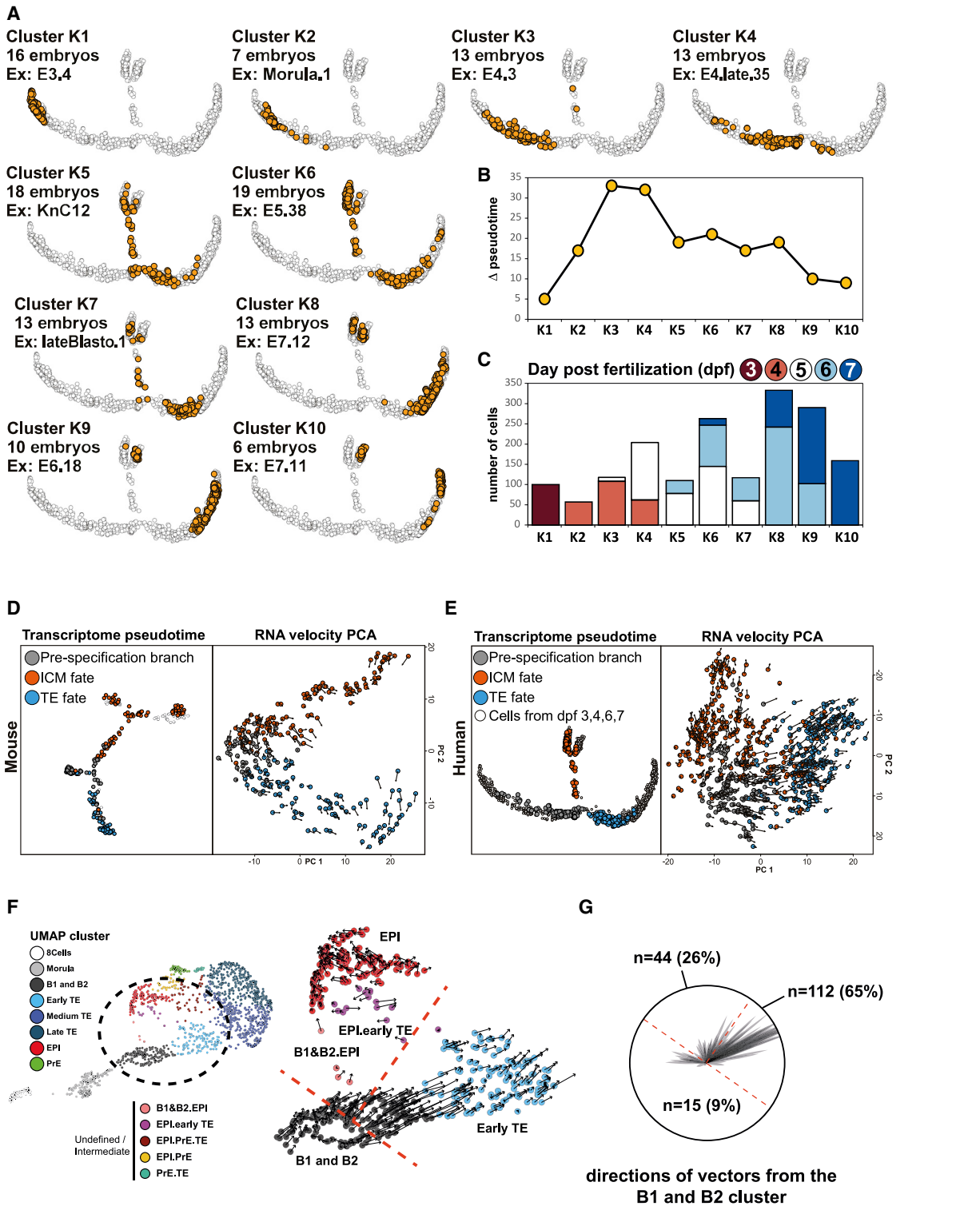


Figure 3. Early blastocysts contain a heterogeneous cell population that is preparing to specify despite having similar transcriptomes

(A) Projection of each embryo cluster (k-means) according to the position of its cells on the pseudotime. For each embryo cluster, the corresponding number of embryos and a representative embryo are 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 dpf.

(D) Transcriptome pseudotime and RNA velocity comparison of cells before and after specification in the mouse.

(E) Transcriptome pseudotime and RNA velocity comparison of cells before and after specification in humans.

(legend continued on next page)

of the POU5F1B/DNMT3L modules in 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 called early TE; (2) a cluster expressing the GATA2 module but not POU5F1B; and (3) a cluster expressing the 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 in each module. Some modules are not associated with fates but 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, whereas 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 and B1 and B2 blastocysts and is then maintained in the EPI but reduced in the PrE and TE. Consequently, the POU5F1B module is the first module that 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, the GATA4 module for PrE, the GATA2 module for early TE/medium TE/late TE, and the NR2F2 module for late TE. Expression of several genes within lineage-specific modules aligns 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 the 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 progression of molecular signatures that lead to specification. All annotations and expression profiles can be found in a user-friendly online browser we generated to facilitate 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, suggests 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 because 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 develop faster (Figure 3C).

To test whether morula cells, located just prior to the first pseudotime bifurcation, have already initiated a transcriptional program biased toward a fate, we used RNA velocity, a measurement of non-spliced mRNA that infers cellular fate (La Manno et al., 2018). In the 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 toward 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 humans, RNA velocity showed that, despite having similar mRNA profiles, unspecified cells localized on the pseudotime prior to branching are already biased toward ICM or TE fates, in agreement with mouse work (Graham and Zernicka-Goetz, 2016; Figure 3E). This can be observed when projecting the velocity vectors on the UMAP cluster B1/B2 (Figures 3F and 3G). Heterogeneity in human embryos seemed to be established at the morula stage, but inner and outer cells are not transcriptionally distinct until the 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 the mouse, compaction is initiated or characterized by sequential establishment of adherent junctions, marked by CDH1, followed by establishment of tight junctions, marked by ZO1 (Barcroft et al., 1998, 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 the theory that apico-basal polarity in human embryos is linked to molecular events, similar to other mammalian species (Gerri et al., 2020; Reijo Pera and Prezzoto, 2016).

After acquisition of apico-basal polarity and compaction, an important event in mammalian pre-implantation development is formation of a blastocoel cavity (White et al., 2018). This is mediated by 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

(F) RNA velocity vectors projected on the B1 and B2, B1&B2.EPI, EPI, EPI.early TE, and early TE clusters from the UMAP analysis. Cells are colored by cell clusters.

(G) Summary 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 of 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 a dashed orange line.

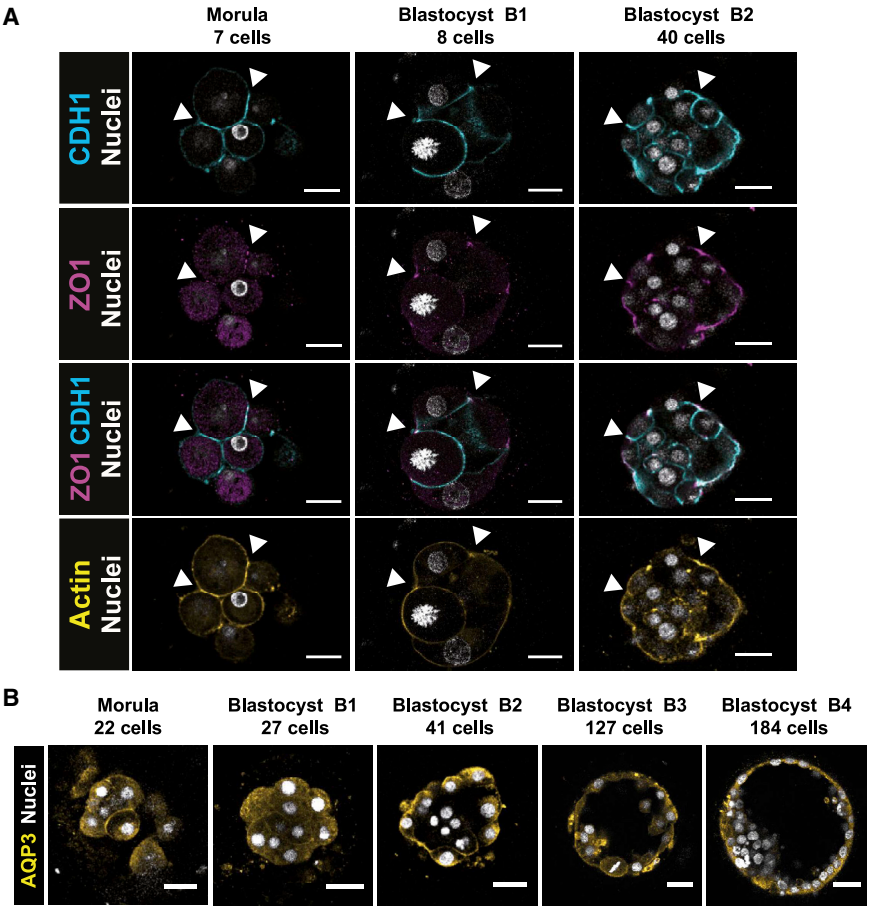


Figure 4. Polarization of outer cells of the morula precedes restriction of the TE marker AQP3

(A) IF of CDH1 (cyan), ZO1 (purple), Actin (yellow), and nuclear counterstaining (white) at the indicated stages. An arrowhead points to the apical part of the cell, with ZO1 staining next to CDH1 staining (n = 3).

(B) IF of AQP3 (yellow) and DAPI (white) as a nuclear counterstain at the indicated developmental stages (n = 6).

Scale bars, 47 μm.

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 onset of a program linked to the ICM, consistent with our pseudotime analysis.

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 allow us to track progression of cells within EPI or PrE fates. Using our integrated transcriptomics 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 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. 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 72 h *in vitro*, which is considered equivalent to "day 8" human embryos. We observed NANOG-positive cells that clearly expressed IFI16, whereas GATA4-positive cells were IFI16 negative or expressed IFI16 weakly (Figure 5D). Quantification of immunofluorescence in pre-implantation embryos and B3 or B4 embryos cultured for 72 h *in vitro* suggests that, before implantation, ICM cells broadly express IFI16, whereas after implantation, cells expressing high levels of IFI16 are negative for GATA4 (Figures 5E and 5F).

We found that IFI16 is restricted to EPI post-implantation in human embryos. Loss of IFI16 might therefore indicate advancement toward the PrE fate. This observation raised an interesting hypothesis: PrE could be specified later than EPI in humans (EPI and PrE are linked differentially to dpf; linear regression, $p = 1.51 \times 10^{-12}$; Figure S5A). The difficulty to identify a distinct ICM molecular signature that would correspond to progenitors of EPI and PrE like in the mouse might indicate that EPI and ICM molecular fates in humans are highly similar or that PrE in humans differentiates from EPI. To better understand PrE fate, we analyzed a recent scRNA-seq dataset of human embryos cultured up to 120 h (day 5) and then switched to IVC medium for up to 216 h (day 14) (Zhou et al., 2019). We focused our analysis on day 5 + 24 h and day 5 + 72 h prolonged culture. UMAP showed that EPI and PrE cells were more related to each other (closer) at day 5 + 24 h than at day 5 + 72 h (Figures S5E and S5G). Interestingly, day 5 + 24 h embryos had cells co-expressing *IFI16* and *GATA4*, whereas this was not the case in day 5 + 72 h embryos (Figures S5F and S5H). We completed this analysis

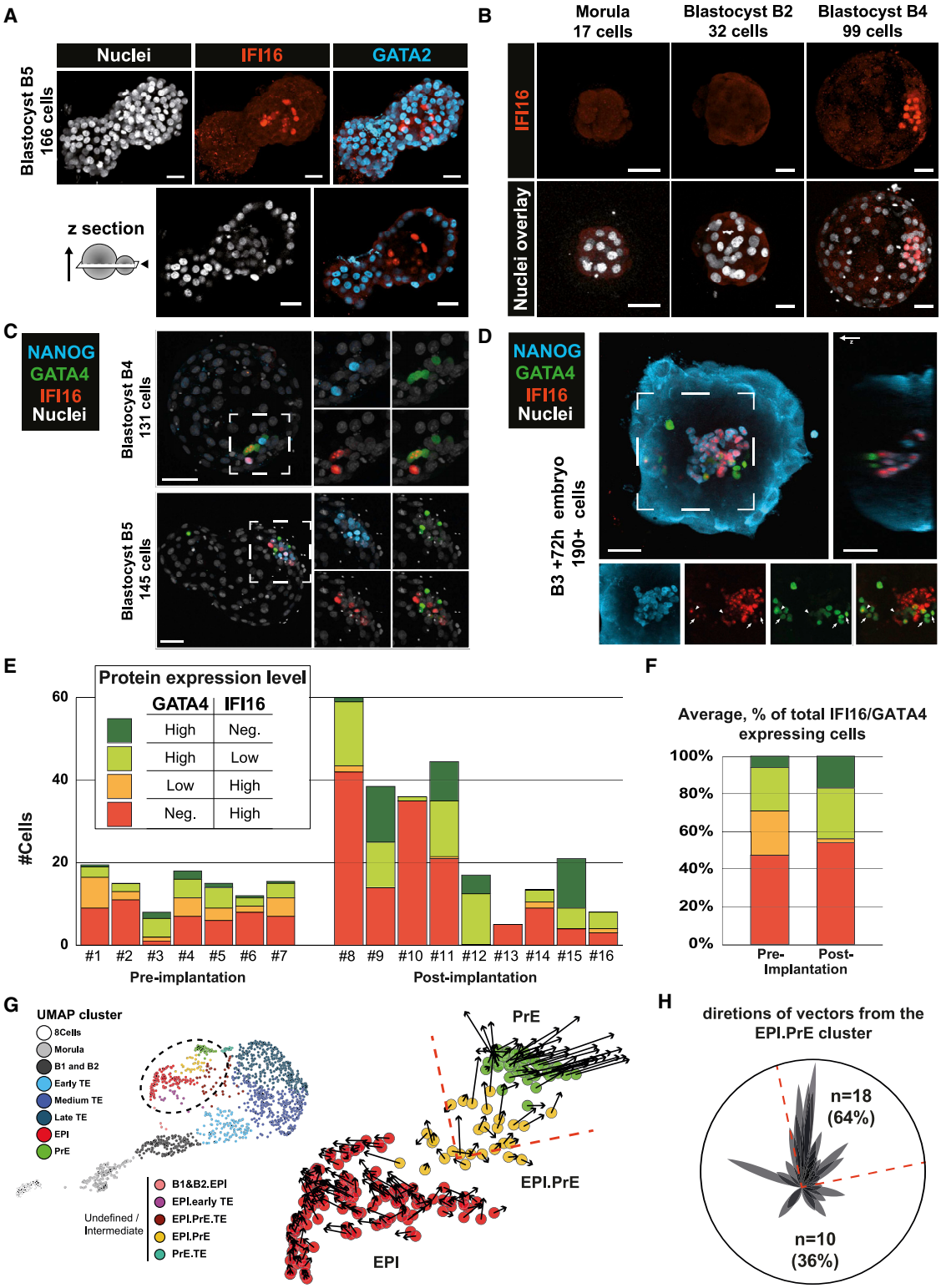


Figure 5. NANOG/IFI16/GATA4 dynamics highlight progression of EPI/PrE specification

(A) IF of IFI16 (red) and GATA2 (cyan) at the B5 blastocyst stage. The z section indicates a z cutting plane (n = 3).

(B) IF of IFI16 at the indicated stages. Nuclear counterstaining is white (n = 13).

(C) IF of NANOG (cyan), GATA4 (green), and IFI16 (red) at the indicated stages. Nuclear counterstaining is white (n = 3).

(D) IF of NANOG (cyan), GATA4 (green), and IFI16 (red) in a B3 + 72 h attached embryo. Arrow, GATA4^{+}/IFI16^{+} cell. Arrowhead, GATA4^{+}/IFI16^{-} cell (n = 4).

(legend continued on next page)

by projected velocity vectors on the UMAP from our dataset or the post-implantation dataset (Figures 5G and 5H; Figures S5E and S5F). The vectors show that, at late blastocyst stage, some EPI cells show a shift toward PrE gene expression. This further supports our hypothesis that PrE specification initiates at the B4 stage and that PrE might differentiate from cells that are transcriptionally close to EPI in humans. Further analysis is required at the B3 stage to identify the molecular states of cells that would eventually become PrE.

To further validate our observations, we investigated 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 humans (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 (Figures S5B–S5D). Interestingly, in day 8 embryos, OTX2 was expressed in GATA4-positive cells closer to NANOG-positive cells (Figures S5B–S5D). It remains to be defined whether OTX2 indicates more advanced PrE development or a different subset of PrE.

Our transcriptomic analyses, validated by immunofluorescence (IF), support that EPI/PrE specification in humans occurs in B4 (late) blastocysts, which is different from the mouse, where EPI and PrE fates are established in 90-cell (early) blastocysts.

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 because of its expression in the later part of the TE branch, which corresponds to developmental stages of expanded/hatched blastocysts (B4 and thereafter; Figures S3A and S3B). In B5 blastocysts, IF showed that nuclear localization of NR2F2 was restricted to polar TE cells juxtaposed to EPI cells, expressing NANOG (Figure 6A). We then investigated NR2F2 expression in B4 blastocysts. Immunostaining revealed GATA3 and GATA2 expression throughout TE, whereas 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 because they are at the end of the trajectory. However, IF 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 suggests that maturation of TE is initiated on the polar side by an EPI-TE dialog.

To further characterize onset of NR2F2^{+} cells, we co-stained NR2F2 with markers of EPI (IFI16) and PrE specification (GATA4). This revealed that NR2F2 expression is concomitant with IFI16 and GATA4 expression (Figure 6C; Figure S6). Staining of NR2F2 in B3/B4 + 72 h *in vitro* cultured embryos confirmed spreading of NR2F2 to all TE cells (Figure 6D). We therefore propose that maturation of TE is initiated by contact with EPI and 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 are 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^{+} and NR2F2^{-} 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 transforming growth factor β (TGF-β), insulin growth factor 1 (IGF1), BMP2, and FGF4, and from TE toward EPI, such as platelet-derived growth factor (PDGF), interleukin-6 (IL-6), and WNT (Figure S7; Figure 6E). This suggests that FGF4 might still have a role in pre-implantation development despite being ruled out from regulating PrE specification (Kuijk et al., 2012; Roode et al., 2012). BMP signaling has been shown recently to inhibit cavitation (De Paepe et al., 2019), but its function after the first lineage specification remains to be studied. The role of IL-6 has been studied recently in pigs; IL-6 knockout (KO) or inhibition of JAK signaling perturbed pre-implantation development, supporting the hypothesis that, in humans and pigs, in the absence of LIF, JAK/STAT signaling is triggered by IL-6 (Ramos-Ibeas et al., 2019). Polar TE-EPI interactions in humans is reminiscent of the ICM-polar TE molecular dialog observed in mouse blastoids (Rivron et al., 2018). Mouse blastoids showed that EPI is necessary for proper TE maturation. However, human-specific events might be involved because some genes critical for mouse TE maturation are not restricted to TE in humans, such as KLF6 (Rivron et al., 2018).

We found that NR2F2 marks mature TE in B4 expanded blastocyst and spreads to all TE cells. Identification of NR2F2^{+} cells will give rise to numerous hypothesis regarding the dual molecular dialog between polar TE and the EPI and 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-cell stage, immediately after differential Cdx2

(E) Manual counting of GATA4-high/IFI16^{-}, GATA4-high/IFI16-low, GATA4-low/IFI16-high, and IFI16-high/GATA4^{-} cells in pre-implantation and post-implantation embryos.

(F) Average percentage of total IFI16/GATA4 expressing cells in pre- and post-implantation embryos.

Scale bars, 47 μm.

(G) RNA velocity vectors projected on the EPI, EPI.PrE, and PrE clusters from the UMAP analysis. Cells are colored by cell clusters.

(H) Summary of the direction and magnitude of RNA velocity vectors 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 of 28 are oriented within an angle of 45° in the direction of the PrE cluster ($p \approx 2.21.10^{-6}$, obtained from the binomial distribution of parameter p = 0.25 and n = 28). The sectors are represented by dashed orange lines.

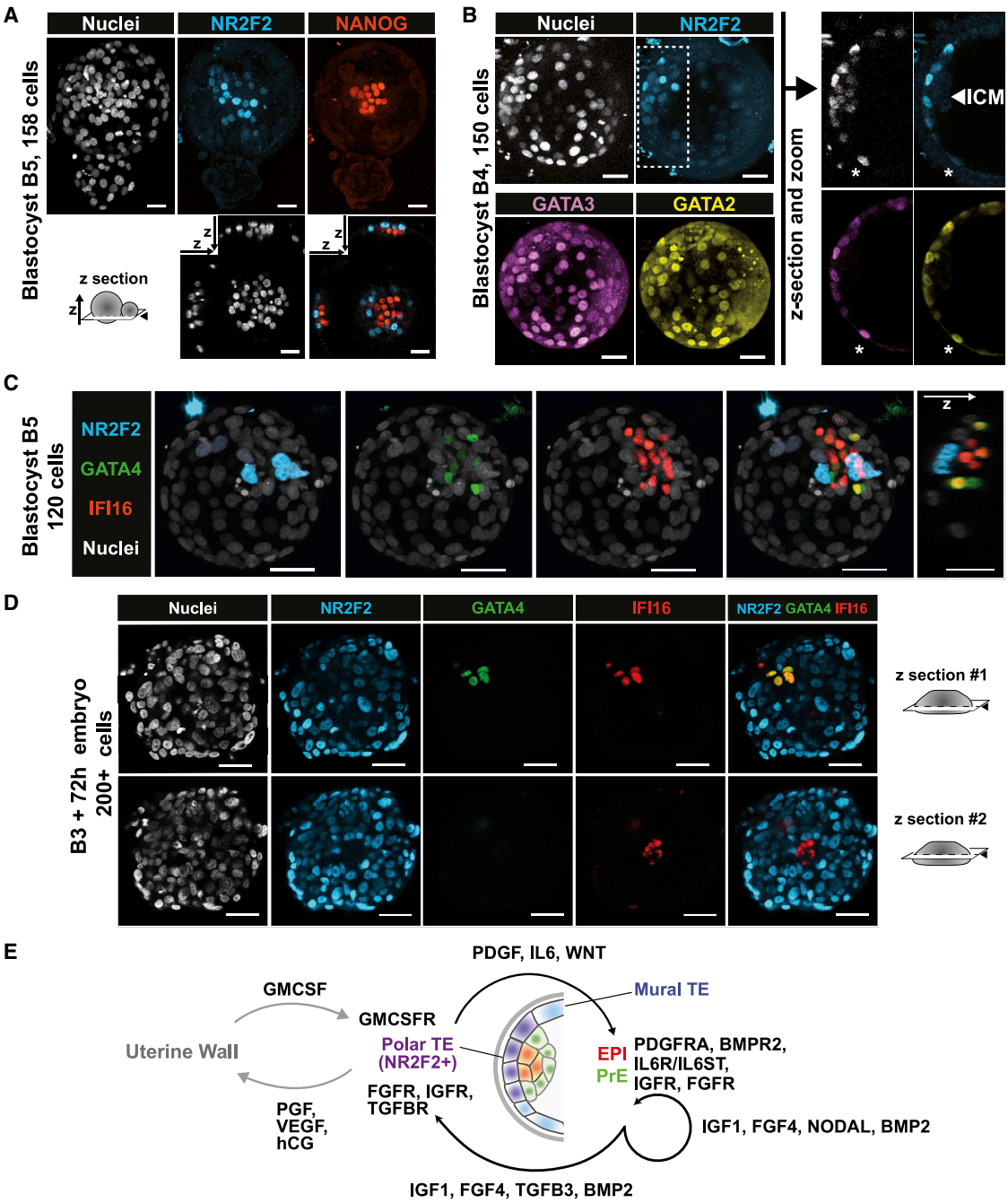


Figure 6. NR2F2 reveals differential maturation of polar and mural TE

(A and B) IF of NR2F2 (cyan) and NANOG (red) at the B5 stage (A) or with GATA2 (yellow) and GATA3 (purple) at the B4 stage (B). Asterisks mark cells that are negative for NR2F2 and positive for GATA3/GATA2 (n = 3).

(C) IF of NR2F2 (cyan), GATA4 (green), and IFI16 (red). Nuclear counterstaining (DAPI) is white (n = 4).

(D) IF of NR2F2 (cyan), GATA4 (green), and IFI16 (red) after 72 h of prolonged culture of a B3 blastocyst (n = 3). Scale bars, 47 μm.

(E) Schematic representing identified cytokine-receptor loops.

protein expression at the late 16-cell morula stage. Experimentally, mouse TE cells at the 32-cell stage are committed, whereas 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 the B2 to the 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

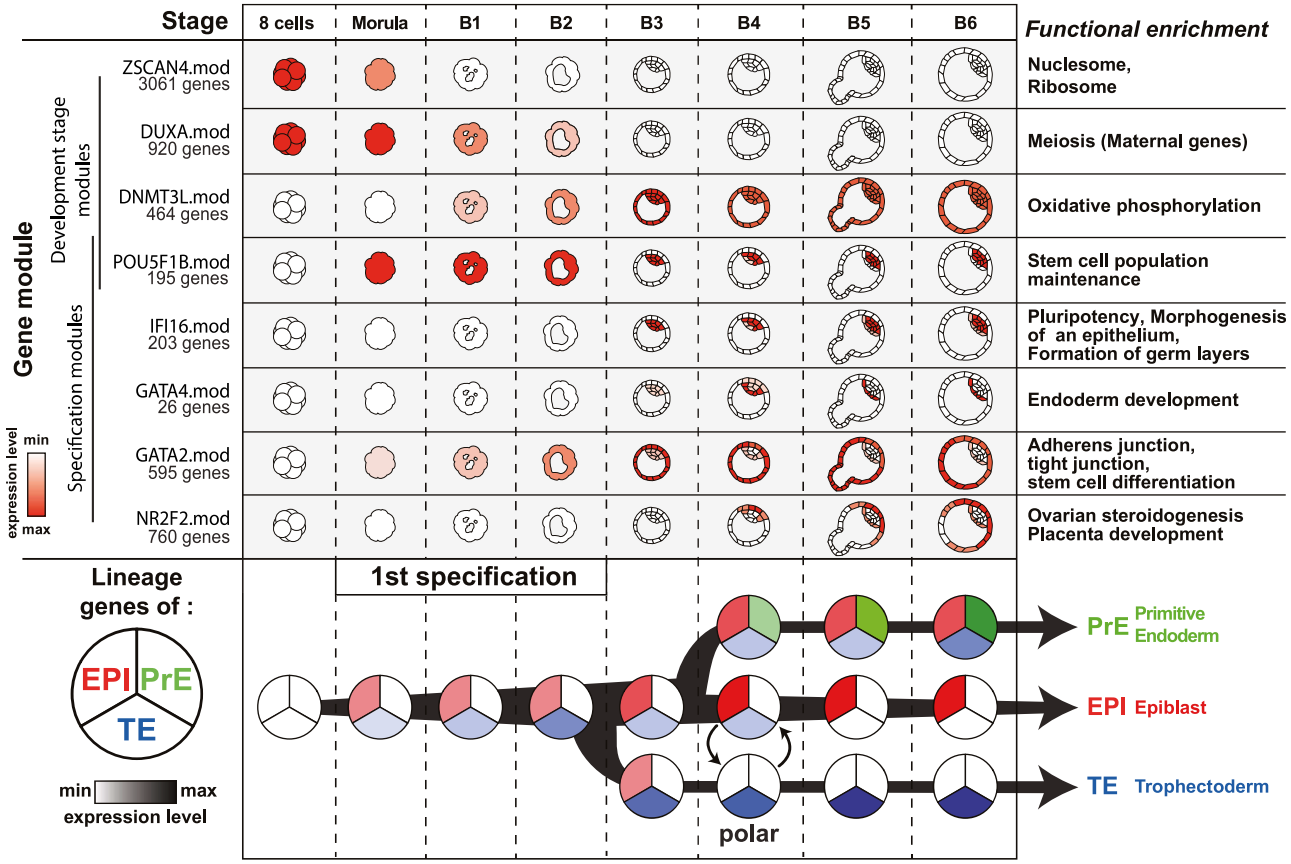


Figure 7. Human pre-implantation development model

Each annotation is represented according to morphology-based stages of human pre-implantation development. Top: schematic representation of gene module global expression. Terms highlighted by functional enrichment are presented on the right. Complete enrichment data are presented in Figure S4D. Bottom: expression of lineage-specific stage per stage and per cell fate.

several gene modules that could be associated with specific states within each branch of the pseudotime. This led us to reveal, by IF, 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. 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 stages to date is to record the development of human embryos with time-lapse devices (Apter et al., 2020). 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 translation of those datasets into a sequence of events, pacing human pre-implantation development. Indeed, binning embryos at 5 dpf results in averaging of transcriptome information from cavitating B1 blastocysts to B4 expanded blastocysts. Because 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, 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 mouse and human (White et al., 2018). Gerri et al. (2020) showed that compaction coupled to the Hippo pathway triggers initiation of TE fate at the morula/B1 stage, which is characterized by expression of GATA3 in external cells. We found that transcriptional variability was increased greatly 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 humans compared with the 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 blastocysts, generally composed of 40 cells. This suggests that two rounds of cellular division, after compaction, are required before cells can initiate/complete distinct transcriptomic programs relating to the ICM/TE. This occurs similarly but slightly earlier

in mouse development, where cell compaction takes place at the 8-cell stage and emergence of distinct cellular profiles is first seen at the 32-cell stage. Regarding cavitation, our analysis also revealed expression of AQP3, which was restricted to TE cells at the B2 stage. However, it remains to be demonstrated whether AQP3 expression depends on TE fate regulators or whether TE fate regulators depend on AQP3.

EPI/PrE specification is difficult to resolve in humans 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 of the mouse raised similar conclusions: EPI cells need to secrete Fgf4 for PrE cells to specify (De Mot et al., 2016). Regarding PrE fate, we found that the PrE signature is defined better when analyses include embryos in prolonged culture; day 6 embryos cultured in IVC medium seem to have more developed PrE cells than day 6 embryos cultured in IVF medium (Petropoulos et al., 2016; Zhou et al., 2019). This could be because IVF media are not currently designed to support extended culture of embryos. 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 medium favors rapid growth/development of human embryos. This observation shows how important it is to understand pre-implantation development in humans to improve embryo culture media.

We consider two main hypotheses for the EPI/PrE specification in humans: (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 these hypotheses, but our study provides a framework; we have pinpointed the morphology associated with specific marker expression and 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 the 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 evidence showing that an EPI/TE molecular dialog could be involved in TE maturation. Interestingly, it is TE cells adjacent to the EPI, on the polar side, that typically initiate implantation prior to maturation of the remaining mural TE cells. However, in mouse embryos, secretion of Fgf4 and nodal by EPI contributes to the progression of polar and mural TE fates, and attachment is mediated by the mural, mature 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 the mouse and humans.

Studies have demonstrated the importance of TE morphological quality for prediction of successful IVF (Ahlström 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 to improve of IVF. We will now be able to focus on specific secreted markers associated with NR2F2^{+} cells that could inform clinicians about the implantation potential of the blastocyst before transfer. Indeed, secreted markers that correlate with expression of NR2F2 could be predictive of the implantation capacity of embryos. This discovery will support the much needed hypothesis regarding human implantation and complement knowledge resulting from study of murine reproduction (Aplin and Ruane, 2017).

Our study also provides key knowledge regarding modeling pre-implantation development with cellular models. Embryonic signature comparison with pluripotent stem cells residing in various states of pluripotency showed that the established hierarchy of pluripotent stem cells mimics EPI: naive pluripotent cells are closest to pre-implantation EPI (5–8 dpf), whereas 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 naive pluripotent stem cells can convert to human trophoblast stem cells (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., 2021). Recently, we proposed that human trophoblast stem cells correspond to human day 8 post-implantation cytotrophoblast (Castel et al., 2020). Our study highlights the importance of knowledge about human peri-implantation development to understand pluripotent and trophoblast stem cell biology.

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 medium formulation for embryo culture during IVF procedures. Improvement of IVF procedures is necessary because the procedural average efficiency is below 27% (European Society of Human Reproduction and Embryology, 2015). Ultimately, our study contributes to understanding human pre-implantation, a gateway to improve IVF success rates.

Limitations of study

A current limitation of human embryo work is the variability of human developmental time *in vitro*, which impairs the reproducibility of experiments. Our study provides a molecular blueprint of human development coupled to time-lapse staging. We unveiled numerous hypotheses and interactions that remain to be validated. Positional information could be improved to discriminate between polar TE and ICM. Fate specification needs to be addressed with lineage tracing approaches, and the importance of specific transcription factors will need to be assessed with CRISPR-Cas9 strategies, which remain restricted in most countries.

STAR★METHODS

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

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stem.2021.04.027.

ACKNOWLEDGMENTS

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. D.M. is supported by FINOX, and MSD contributed to the project. We thank the core facilities BIRD, PFiPSC, and MicroPicell. This work was supported by "Paris Scientifique region Pays de la Loire: HUMPLURI" and IHU CESTI.

AUTHOR CONTRIBUTIONS

T.F. and L.D. designed the study. D.M., A. Bruneau, S.L., and L.D. wrote the manuscript with input from all authors. S.L., A.R., J.F., J.L., and A. Bruneau performed embryo manipulation. P.H., S.N., S.L., A. Bruneau, A.R., and L.D. performed IF analyses. A.R. and S.K. prepared scRNA-seq samples. M.S. and T.M. performed scRNA-seq. D.M., V.F.-C., and Y.L. performed bioinformatics analyses under the supervision of A. Bihouée and J.B. and with input from B.B. and G.C. T.F. and P.B. supervised human embryo donation. All authors approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We worked to ensure gender balance in the recruitment of human subjects. We worked to ensure that the study questionnaires were prepared in an inclusive way. One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.

Received: August 6, 2019
Revised: July 16, 2020
Accepted: April 22, 2021
Published: May 17, 2021

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. Mol. Hum. Reprod. *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 Rep. *15*, 2651–2664.

Ahlström, A., Westin, C., Reismer, E., Wikland, M., and Hardarson, T. (2011). Trophectoderm morphology: an important parameter for predicting live birth after single blastocyst transfer. Hum. Reprod. *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. https://doi.org/10.1101/707679.

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology (2011). The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. Hum. Reprod. *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. J. Cell Sci. *130*, 15–22.

Apter, S., Ebner, T., Freour, T., Guns, Y., Kovacic, B., Le Clef, N., Marques, M., Meseguer, M., Montjean, D., Sfontouris, I., et al. (2020). Good practice recommendations for the use of time-lapse technology(dagger). Hum. Reprod. Open. *2020*, hoaa008.

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. J. Reprod. Fertil. *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. Dev. Biol. *256*, 342–354.

Bentin-Ley, U., Horn, T., Sjögren, A., Sorensen, S., Falck Larsen, J., and Hamberger, L. (2000). Ultrastructure of human blastocyst-endometrial interactions in vitro. J. Reprod. Fertil. *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*, dev167833.

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 Rep. *33*, 108419.

Chan, C.J., Costanzo, M., Ruiz-Herrero, T., Mönke, G., Petrie, R.J., Bergert, M., Diz-Muñoz, 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. Dev. Cell *10*, 615–624.

Chen, X., Zhang, J., Wu, X., Cao, S., Zhou, L., Wang, Y., Chen, X., Lu, J., Zhao, C., Chen, M., and Ling, X. (2014). Trophectoderm morphology predicts outcomes of pregnancy in vitrified-warmed single-blastocyst transfer cycle in a Chinese population. J. Assist. Reprod. Genet. *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 *15*, 198–213.

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

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., 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. Hum. Reprod. *28*, 740–749.

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. Mol. Reprod. Dev. *86*, 53–62.

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*, e52504.

Dumortier, J.G., Le Verge-Serandour, M., Tortorelli, A.F., Mielke, A., de Plater, L., Turlier, H., and Maître, J.L. (2019). Hydraulic fracturing and active coarsening position the lumen of the mouse blastocyst. Science *365*, 465–468.

Ester, M., Kriegel, H.P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 226–231.

European Society of Human Reproduction and Embryology (2015). Data collection and research. https://www.eshre.eu/.

Feyeux, M., Reignier, A., Mocaer, M., Lammers, J., Meistermann, D., Barrière, P., Paul-Gilloteaux, P., David, L., and Fréour, T. (2020). Development of automated annotation software for human embryo morphokinetics. Hum. Reprod. *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*, e42298.

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. Fertil. Steril. *73*, 1155–1158.

Gerri, C., McCarthy, A., Alanis-Lobato, G., Demtschenko, A., Bruneau, A., Loubersac, S., Fogarty, N.M.E., Hampshire, D., Elder, K., Snell, P., et al. (2020). Initiation of a conserved trophectoderm program in human, cow and mouse embryos. Nature *587*, 443–447.

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. Hum. Reprod. *10*, 3259–3263.

Giocanti-Auregan, A., Vacca, O., Bénard, R., Cao, S., Siqueiros, L., Montañez, 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., Zyzyńska, 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? Curr. Top. Dev. Biol. *117*, 671–695.

Greenberg, M.V.C., and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. Nat. Rev. Mol. Cell Biol. *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. Proc. Natl. Acad. Sci. USA *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.E., Spindlow, D., Yang, J., Clarke, J., Dattani, A., Yanagida, A., Li, M.A., Myers, S., et al. (2021). Human naive epiblast cells possess unrestricted lineage potential. Cell Stem Cell.

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. Proc. Natl. Acad. Sci. USA *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? Biol. Reprod. *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. Fertil. Steril. *99*, 1283–1289.e1.

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. Curr. Biol. *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 Dev. *22*, 2692–2706.

Kilens, S., Meistermann, D., Moreno, D., Chariau, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S., et al.; Milieu Intérieur Consortium (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. https://doi.org/10.1101/060012.

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., Lönnerberg, 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.

Lindenberg, S. (1991). Experimental studies on the initial trophoblast endometrial interaction. Dan. Med. Bull. *38*, 371–380.

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 *2015*, bav095.

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. Nat. Methods *14*, 1055–1062.

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.

McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv, 1802.03426, arXiv. https://arxiv.org/abs/1802.03426.

Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. Dev. Biol. *375*, 54–64.

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R., and Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Cell *165*, 1012–1026.

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*, e22906.

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. Nat. 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. Curr. Top. Dev. Biol. *120*, 401–420.

Rivron, N.C., Frias-Aldeguer, J., Vrij, E.J., Boisset, J.C., Korving, J., Vivié, J., Truckenmüller, R.K., van Oudenaarden, A., van Blitterswijk, C.A., and Geijsen, N. (2018). Blastocyst-like structures generated solely from stem cells. Nature *557*, 106–111.

Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A., and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. Dev. Biol. *361*, 358–363.

Rossant, J., and Lis, W.T. (1979). Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. Dev. Biol. *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*, e49500.

Sergushichev, A.A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv. https://doi.org/10.1101/060012.

Shahbazi, M.N., Jedrusik, A., Vuoristo, S., Recher, G., Hupalowska, A., Bolton, V., Fogarty, N.N.M., Campbell, A., Devito, L., Ilic, D., et al. (2016). Self-organization of the human embryo in the absence of maternal tissues. Nat. Cell Biol. *18*, 700–708.

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. Hum. Reprod. *13*, 676–681.

Singh, M., Widmann, T.J., Bansal, V., Cortes, J.L., Schumann, G.G., Wunderlich, S., Martin, U., Garcia-Canadas, M., Garcia-Perez, J.L., Hurst, L.D., and Izsvak, Z. (2019). The selection arena in early human blastocysts resolves the pluripotent inner cell mass. BioRxiv *318329*.

Stirparo, G.G., Boroviak, T., Guo, G., Nichols, J., Smith, A., and Bertone, P. (2018). Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. Development *145*, dev158501.

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.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell *15*, 471–487.

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.

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. J. Assist. Reprod. Genet. *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. Hum. Reprod. *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*, 4.22.1-17.

White, M.D., Zenker, J., Bissiere, S., and Plachta, N. (2018). Instructions for Assembling the Early Mammalian Embryo. Dev. 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. Biochem. Biophys. Res. Commun. *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. Nat. Struct. Mol. Biol. *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. Fertil. Steril. *99*, 1600–1609.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
IFI16	Novusbio	Cat# NBP1-83118; RRID:AB_11041205
AQP3	Antibodies-online	Cat# ABIN863208
NR2F2	Abcam	Cat# ab211776
GATA2	Sigma	Cat# WH0002624M1; RRID:AB_1841726
CDH1	Abcam	Cat# ab1416; RRID:AB_300946
GATA3	R&D systems	Cat# AF2605; RRID:AB_2108571
NANOG	R&D systems	Cat# AF1997; RRID:AB_355097
SOX17	R&D systems	Cat# AF1924; RRID:AB_355060
ZO1	Invitrogen	Cat# 40-2200; RRID:AB_2533456
NANOG	Thermofisher	Cat# PA1-097; RRID:AB_2539867
GATA4	Invitrogen	Cat# 14-9980-82; RRID:AB_763541
IFI16	Abcam	Cat# ab55328; RRID:AB_943797
OTX2	R&D	Cat# AF1979; RRID:AB_2157172
Donkey anti rabbit Alexa Fluor 488	Life Technologies	Cat#A21206; RRID:AB_2535792
Donkey anti mouse Alexa Fluor 488	Life Technologies	Cat#A21202; RRID:AB_141607
Donkey anti rabbit Alexa Fluor 568	Life Technologies	Cat#A10042; RRID:AB_2534017
Donkey anti mouse Alexa Fluor 568	Life Technologies	Cat#A10037; RRID:AB_2534013
Donkey anti goat Alexa Fluor 568	Life Technologies	Cat#A11057; RRID:AB_2534104
Donkey anti goat Alexa Fluor 647	Life Technologies	Cat#A21447; RRID:AB_2535864
Donkey anti rat DyLight 550	Thermofisher	Cat#SA5-10027; RRID:AB_2556607
Donkey anti mouse Alexa Fluor Plus 647	Life Technologies	Cat#A32787; RRID:AB_2762830
Biological samples		
Human pre-implantation embryos from 8 cells to B5 stage	Assisted Reproductive Technology unit of the University Hospital of Nantes	Authorized project: RE13-010 and RE18-010
Chemicals, peptides, and recombinant proteins		
RapidWarmBlast	Vitrolife	10120
RapidWarCleave	Vitrolife	10118
G2 plus	Vitrolife	10132
HSA solution	Vitrolife	10064
G-PGD	Vitrolife	10074
GTL	Vitrolife	10145
IVC1 medium	Cell guidance systems	M11-6
IVC2 medium	Cell guidance systems	M12-6
Deposited data		
scRNA-seq original data	This paper; ENA	https://www.ebi.ac.uk/ena/browser/view/PRJEB30442
scRNA-Seq alignment pipeline	This paper; GitLab	https://gitlab.univ-nantes.fr/E114424Z/SingleCell_Align
scRNA-Seq preprocessing and normalization	This paper; GitLab	https://gitlab.univ-nantes.fr/E114424Z/singlecellnormalize
WGCNA workflow	This paper; GitLab	https://gitlab.univ-nantes.fr/E114424Z/WGCNA
Monocle2 workflow, UMAP and cell clustering	This paper; GitLab	https://gitlab.univ-nantes.fr/E114424Z/monocle2_workflow

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Pseudotime User Interface source code	This paper; GitLab	https://gitlab.univ-nantes.fr/E114424Z/ptuihost
Software and algorithms		
Volocity	Quorum technologies	V6.3
Fiji	ImageJ	V1.53c
R	Bioconductor	v 4.0.3

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

The accession number for the original scRNA-seq data reported in this paper is ENA: PRJEB30442.

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

scRNaseq alignment pipeline:

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

scRNaseq 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).

METHOD 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 S4.

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 (https://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. Several embryos were photobleached in order to restain it (Figure 6C; Figures S6A and S6B). 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 embryos were restained with NR2F2 (Rabbit / 488) – GATA4 (Rat / 568) – IFI16 (Mouse / 647). Scale of Figure S1E was estimated with the diameter of the STRIPPER® tip.

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.

QUANTIFICATION AND STATISTICAL ANALYSIS

Single-cell RNA sequencing

Single-cell isolation and overall dataset are presented in Figure S1 and Tables S1 and 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. (2016) 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 S4C. Gene list for each module can be found in Table S3.

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 (Korotkevich et al., 2019; Sergushichev, 2016).

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:

- estimation of gene dispersion and select ordering genes

Normalized expression:

- projection of gene expression on pseudotime or UMAP (Figures 1C and 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. Cell annotations can be found in Table S2.

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 Figures 3D and 3E were calculated with the function *pca.velocity.plot*.

RNA velocity vectors were projected on an isometric representation of the UMAP (Figures 3F, 3G, 5G, and 5H) 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.

In the Figures 3F and 3G, the limit between the sector of early TE and EPI is the bisector of the angle formed by the cluster centroid of EPI, B1 & B2 and early TE in the UMAP.

In the Figures 5G and 5H, the sector of the circular diagram was oriented to face the centroid of the PrE cluster from the centroid of the EPI.PrE cluster.

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 S4D) was processed by taking the first component of a principal component analysis of genes from each enriched term.

Cell Stem Cell, Volume *28*

Supplemental Information

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

Dimitri Meistermann, Alexandre Bruneau, Sophie Loubersac, Arnaud Reignier, Julie Firmin, Valentin François-Campion, Stéphanie Kilens, Yohann Lelièvre, Jenna Lammers, Magalie Feyeux, Phillipe Hulin, Steven Nedellec, Betty Bretin, Gaël Castel, Nicolas Allègre, Simon Covin, Audrey Bihouée, Magali Soumillon, Tarjei Mikkelsen, Paul Barrière, Claire Chazaud, Joel Chappell, Vincent Pasque, Jérémie Bourdon, Thomas Fréour, and Laurent David

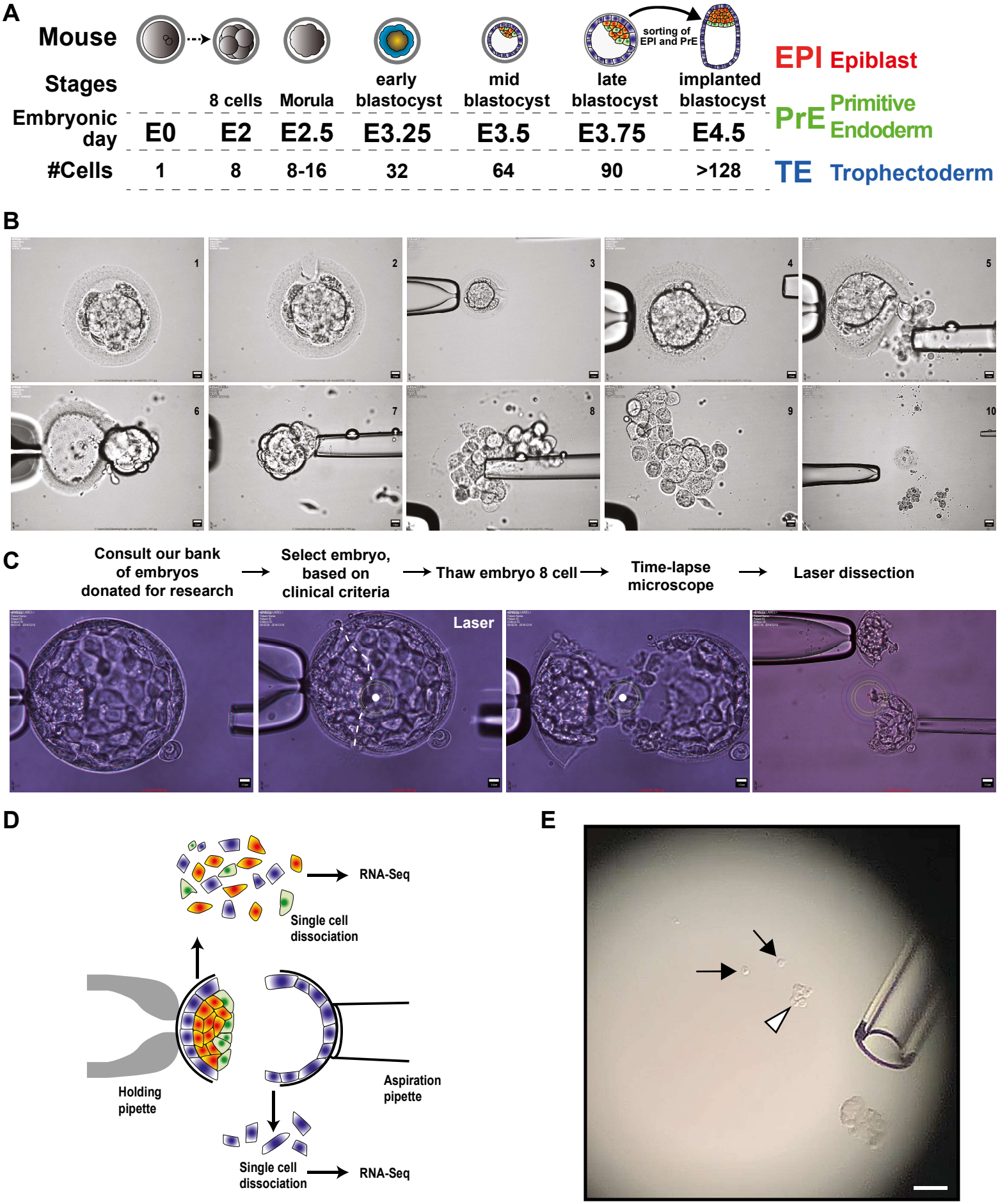
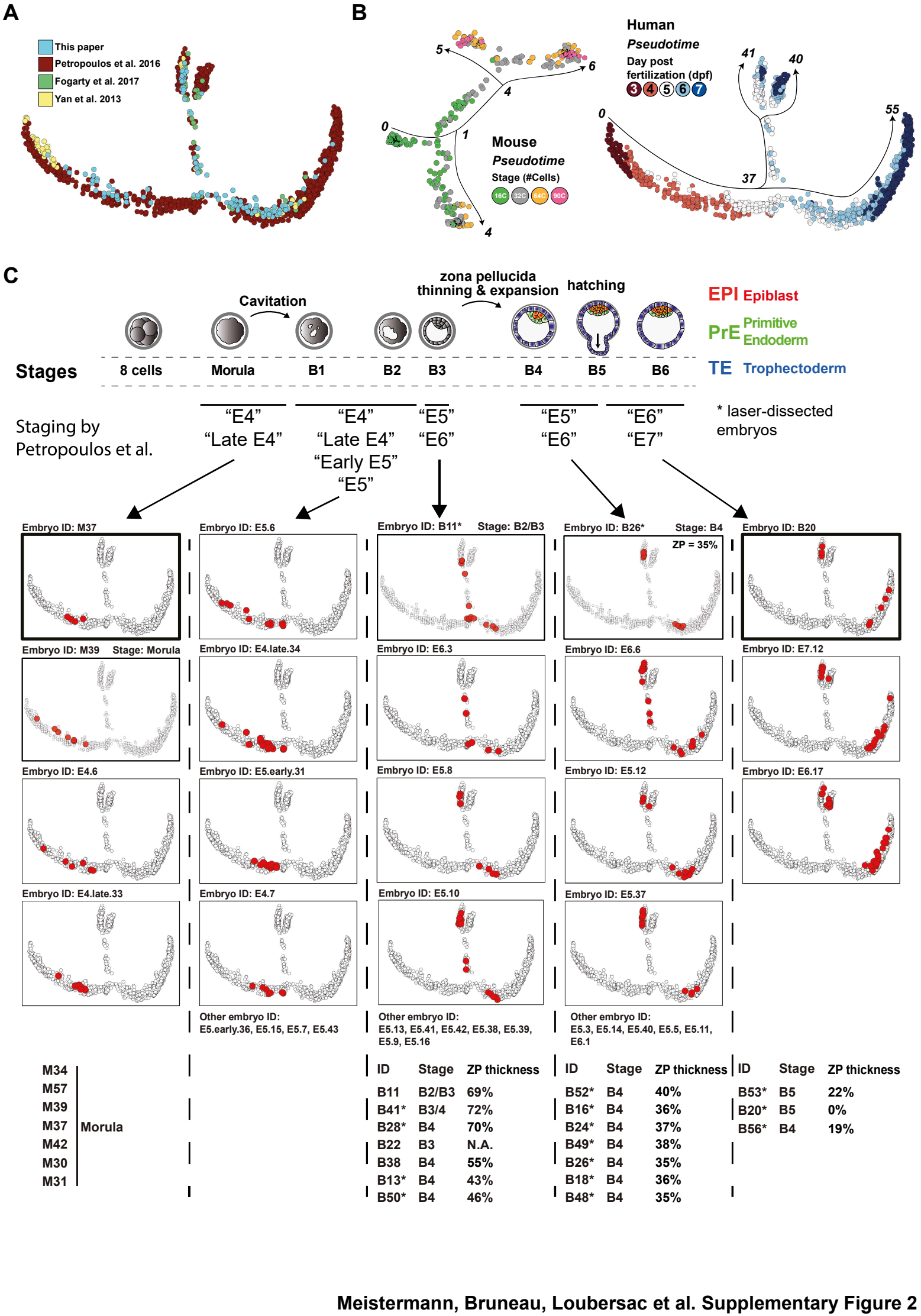


Figure S1 – related to figure 1. Context and sample preparation of the study
(A) Schematic of mouse pre-implantation development.
(B-E) Detailed manipulation of embryos before scRNA-seq. Morula are first incubated in de-compaction media, then the zona pellucida (ZP) is pierced by a laser. Using micropipettes, fragments are removed, and the morula is extracted from the ZP. Cells are then manipulated until dissociated (B). For blastocysts, embryos are thawed and monitored until desired developmental stage. Embryos are then laser-dissected (dot)(C), separating the polar side (containing TE, PrE and EPI cells) and the mural side (containing TE cells). Cells are dissociated and RNA-sequenced (D-E). For (E) : Arrow = single cells; Arrowhead = clumps. Estimated scale on (E) = 100µm



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%

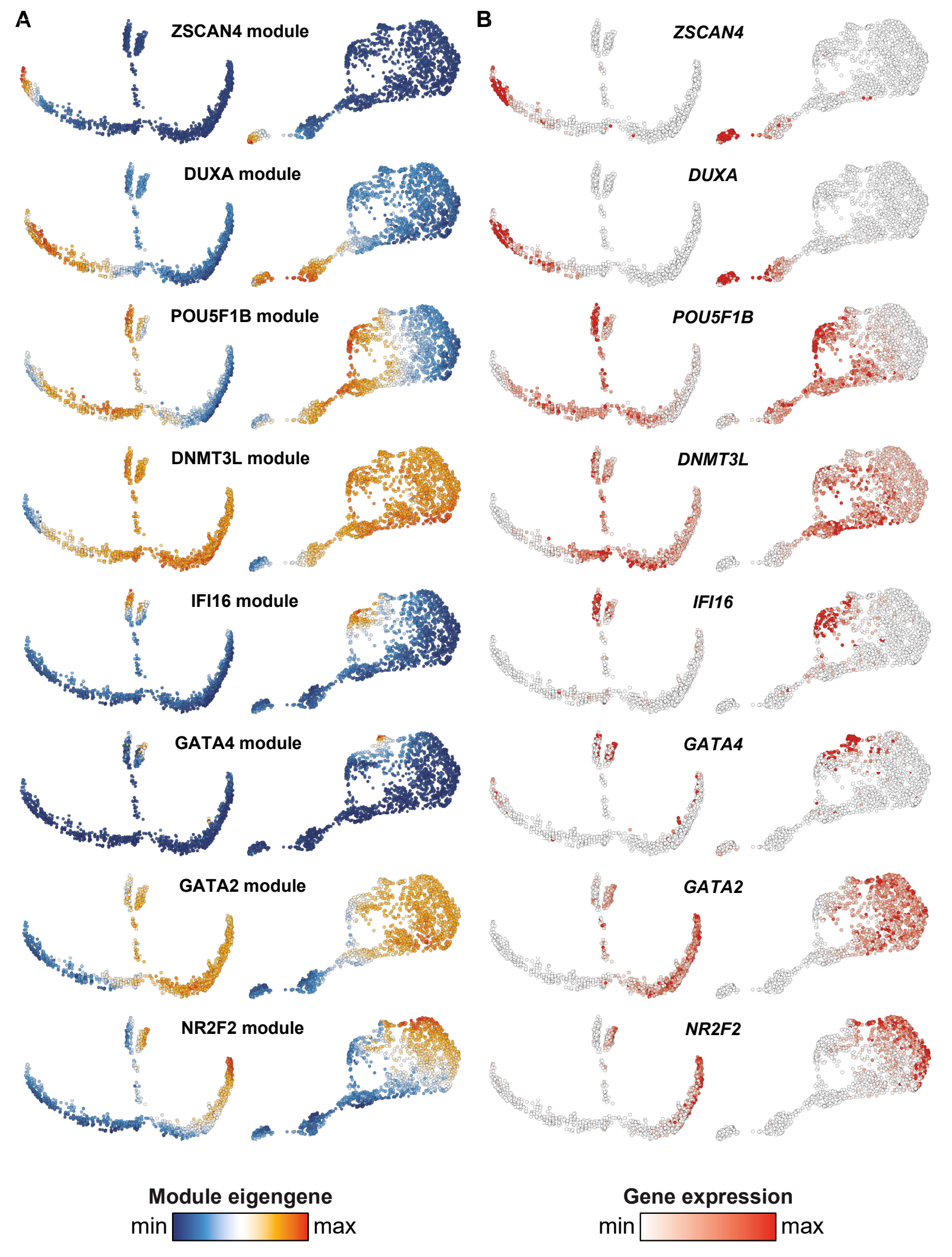
Meistermann, Bruneau, Loubersac et al. Supplementary Figure 2

Figure S2 – related to figure 1. Additional results projected on the pseudotime

(A) Projection of samples per dataset on the pseudotime.

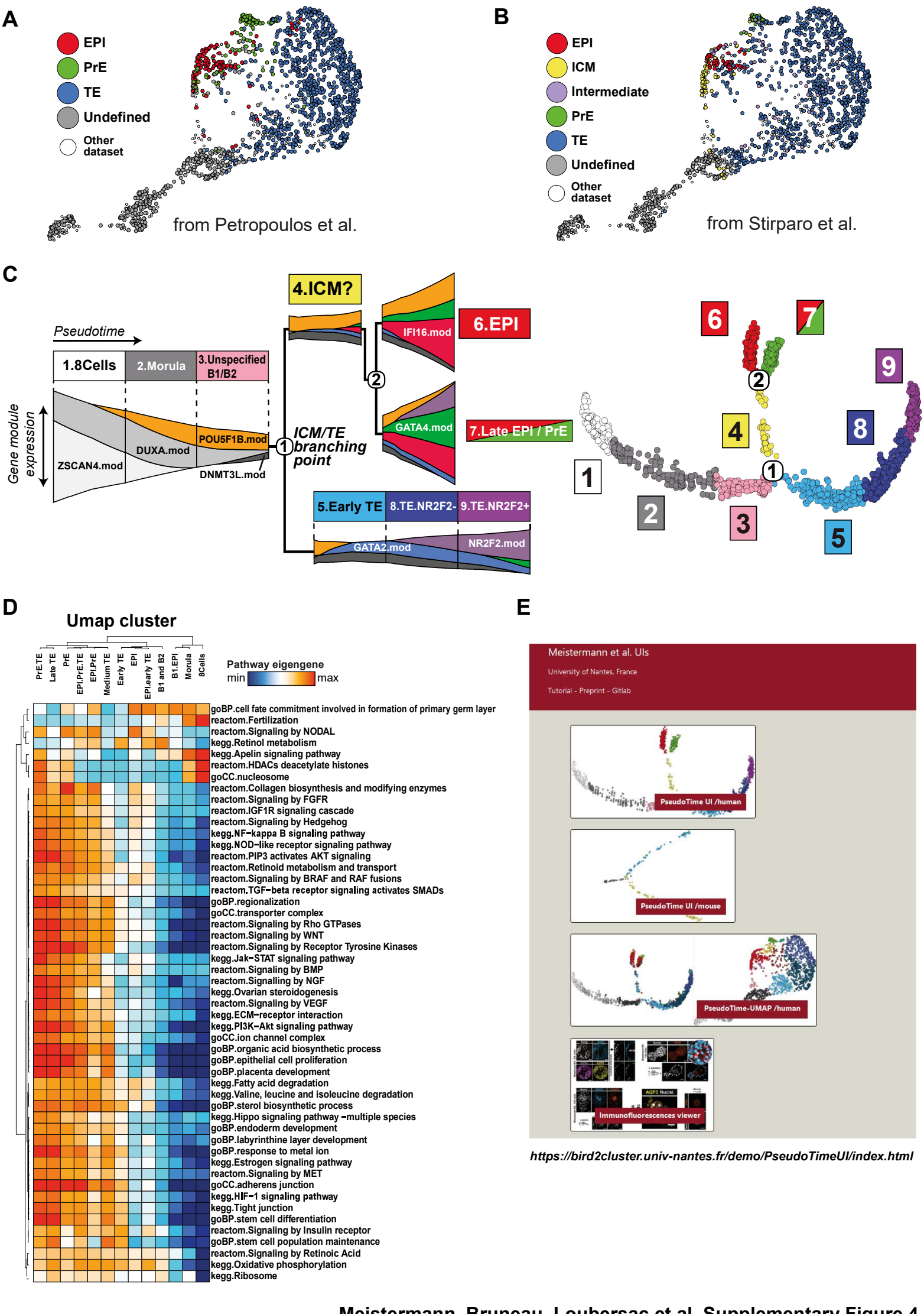
(B) Quantitative values of pseudotime (in italic) and development trajectories on the pseudotime of mouse (left) and human (right). Stages are indicated for both species, as number of cells in the embryo for mouse and day post fertilization in human. The pseudotime value indicates the quantity of transcriptomic changes from the root (pseudotime = 0) of the model.

(C) Refined annotation from all cells and embryos included in our study. Embryos are ranked based on the average positioning of all cells from each embryo included in the study. This highlights that embryos annotated “E5” can be found at the exit of morula, just before specification, just after specification or later on, interspersed with “E6”. This could be due to difficulty to annotate embryos based on a single observation, on the fact that multiple morphologies are present at day 5, or simply because the subset of cells sequenced from an embryo is not representative of the molecular state of the embryo. Our model allows us to bin embryos based on their molecular profile, allowing a finer analysis of events, pacing human pre-implantation development, especially at E5.



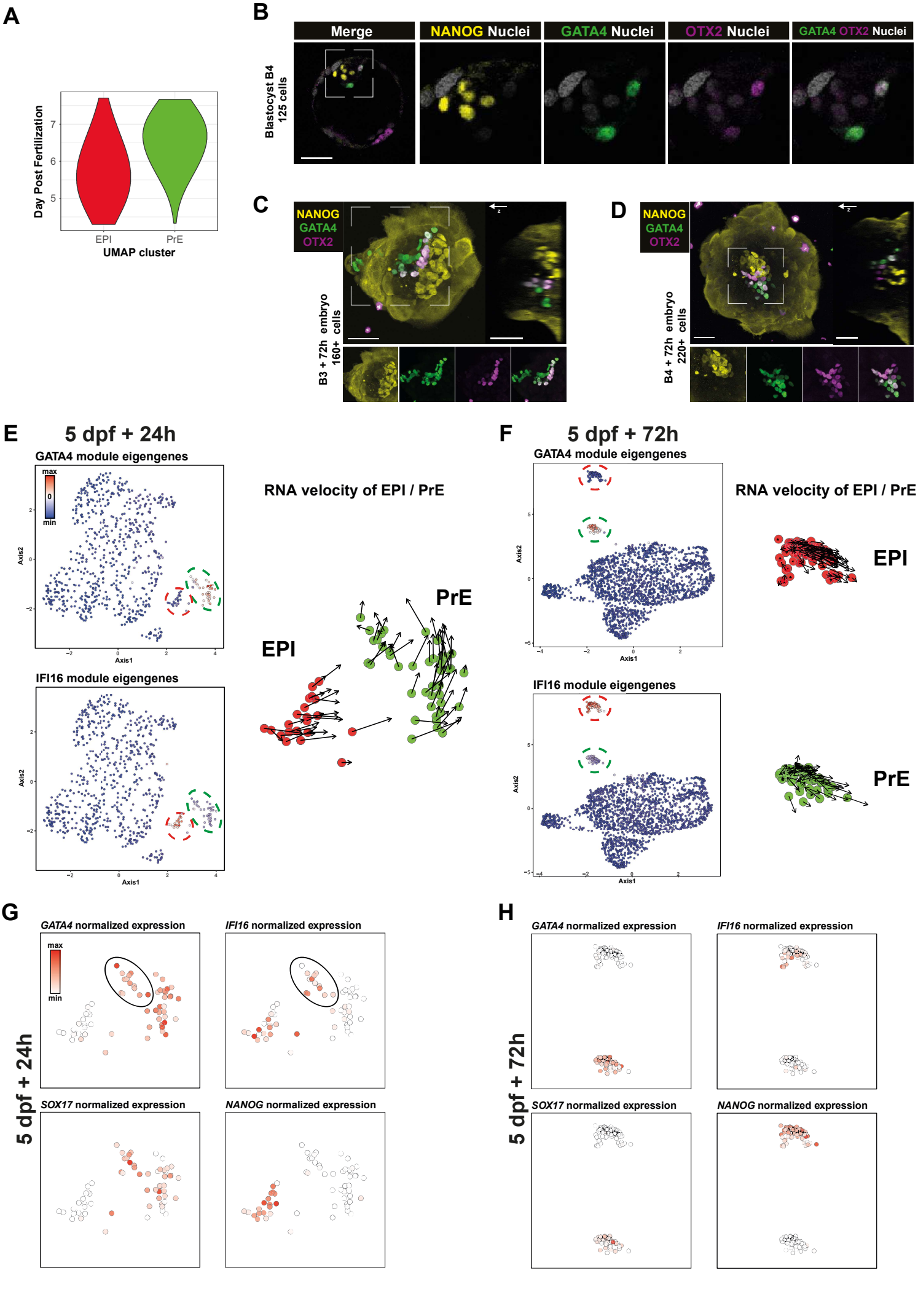
Meistermann, Bruneau, Loubersac et al. Supplementary Figure 3

Figure S3 – related to figure 2. Projection of gene modules on the pseudotime and the UMAP.
(A) Projection of module eigengene of each retained WGCNA module, on pseudotime and UMAP. A positive value indicates that genes of the module are globally more expressed than the mean expression, a negative value less expressed. The UMAP was performed by reducing this information into dimension 2.
(B) Projection of expression of the genes that named the modules, on pseudotime and UMAP.



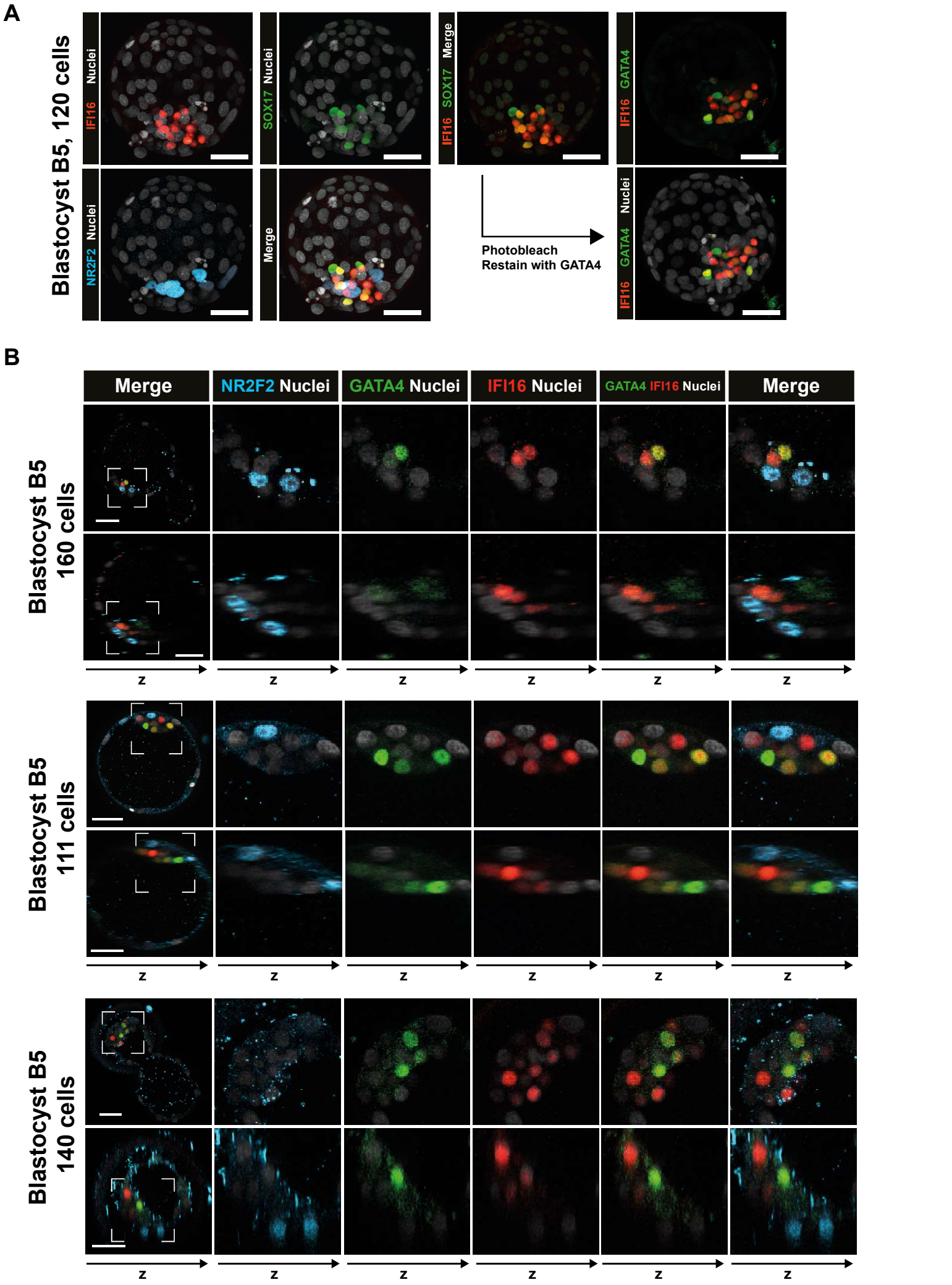
Meistermann, Bruneau, Loubersac et al. Supplementary Figure 4

Figure S4 – related to figure 2. Comparison with previous scRNAseq studies, gene modules dynamics on the pseudotime, functional enrichment and presentation of the web user interface.
(A) Projection of Petropoulos et al. cell lineage attribution on eigengenes UMAP
(B) Projection of Stirparo et al. cell lineage attribution on eigengenes UMAP
(C) Streamgraph of gene module expression along pseudotime states. Each module was named by one of its representative genes. Expression level of each module is represented by the thickness of its ribbon on the streamgraph and corresponds to the WGCNA eigengene metric. The pseudotime states are projected on the pseudotime (right).
(D) Heatmap representing the average of 51 pathway eigengene across the UMAP clusters (presented in Figure 2A). Pathway eigengene is defined as the first principal component of a principal component analysis with the genes of the pathway as initial dimensions. Enrichment was done in gene modules according to 5 databases: Gene ontology (GO Molecular Function, GO Cellular component, GO Biological process), KEGG and Reactom. All terms were significantly enriched in at least one gene module with an adjusted pvalue < 0.05.
(E) Screenshot of our web application “PseudoTime User Interface”. URLs are indicated below.



Meistermann, Bruneau, Loubersac et al. Supplementary Figure 5

Figure S5 – related to figure 5. Additional data on EPI and PrE specification.
(A) Violin plot of day post fertilization distribution per UMAP cluster EPI-PrE from Figure 2A.
(B) Immunofluorescence of NANOG (yellow), GATA4 (green) and OTX2 (magenta) at B4 stage. Nuclear counterstaining is in white (n=3).
(C and D) Immunofluorescence of NANOG (yellow), GATA4 (green) and OTX2 (magenta) at B3 + 72h (left) and B4 + 72h (right) in vitro culture stage. Scale bar = 47 µm (n=3).
(E-F) Projection of IFI16.mod and GATA4.mod eigengene on a dimension reduction (UMAP) of gene expression, at 6 dpf (E) or 8 dpf (F). Dashed ellipses indicate EPI (red) and PrE (green) manually picked clusters. RNA velocity vectors are projected on the UMAP on the right, for the EPI and PrE clusters. Data from Zhou et al. 2019.
(G-H) Projection of gene expression from representative markers of PrE (GATA4, SOX17) and EPI (IFI16, NANOG) on the UMAP of gene expression from Zhou et al. 2019, at 6 dpf (G) or 8 dpf (H). Ellipses indicate an aera with co-expression of IFI16 and GATA4.



Meistermann, Bruneau, Loubersac et al. Supplementary Figure 6

Figure S6 – related to Figure 5 and 6. Additional data on EPI and PrE specification and TE maturation.
(A) Immunofluorescence of NR2F2 (cyan), SOX17 (green) and IFI16 (red) at blastocyst B5 stage. Embryo was photobleached and restained with GATA4 (green) (see material and methods). Nuclear counterstaining is in white. (n=2 for SOX17 staining, n=4 for GATA4). GATA4 staining is an additionnal view of Figure 6C.
(B) Immunofluorescence of NR2F2 (cyan), GATA4 (green) and IFI16 (red) in three blastocysts at B5 stage. Nuclear counterstaining is in white (n=4).
Scale bar = 47 μm.

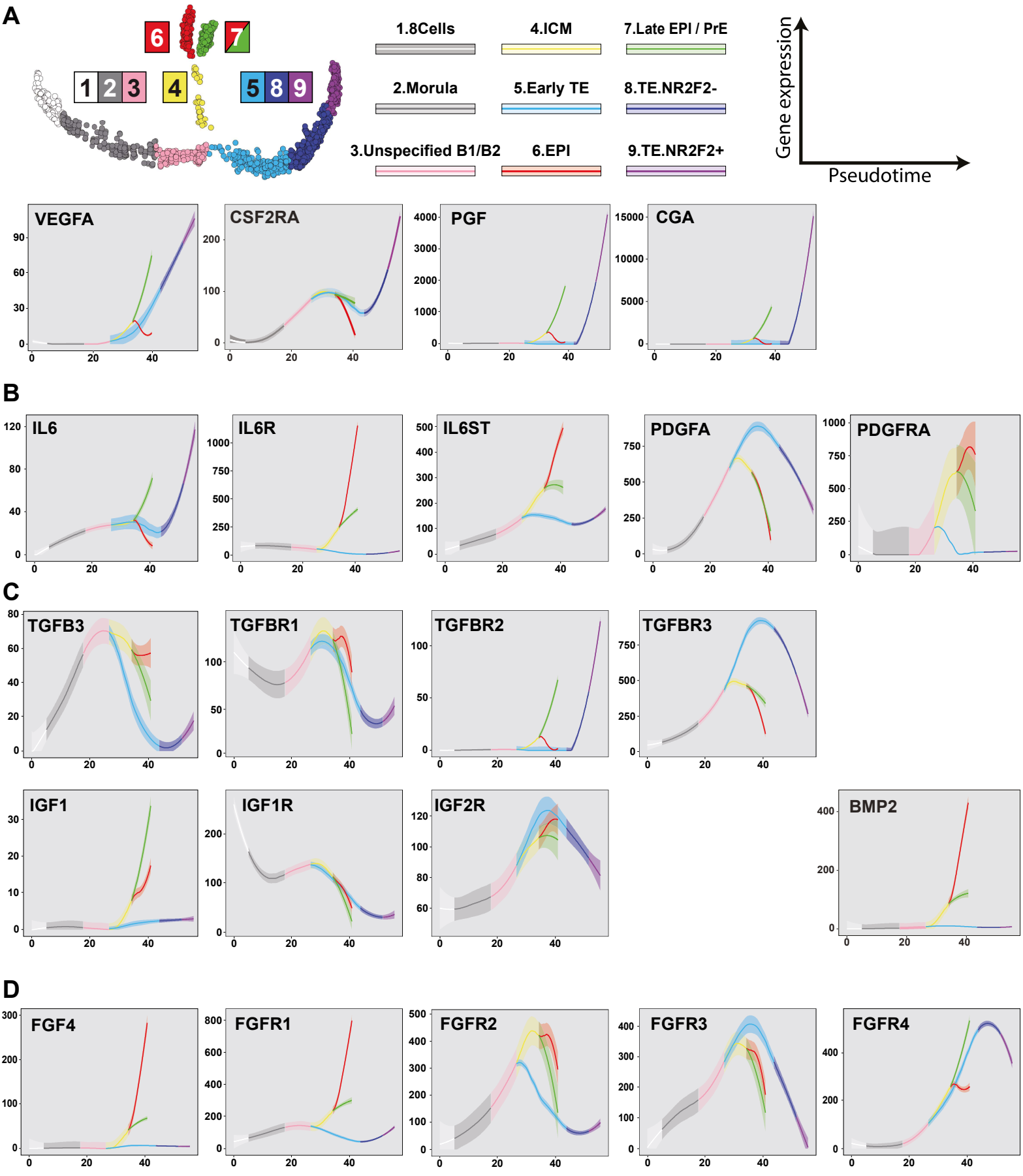


Figure S7 – related to figure 6. Expression profile of selected cytokines and receptors.
(A-D), Expression profiles signaling pathways components during pre-implantation development. Cytokines secreted by the TE targeting the uterine wall (A), cytokines expressed by the TE and targeting the ICM (B), cytokines expressed by the ICM and targeting the TE (C) and cytokines expressed by the ICM targeting the whole embryo (D).

Primary antibodies

Target	Species	Mono or polyclonal	Reactivity	Suplyer	Cat n°	Dilution	Related to
IFI16	**Rabbit**	Poly	Human	novusbio	Cat# NBP1-83118; RRID:AB_11041205	1/100	Figure 5A-B
AQP3	**Rabbit**	Poly	Human	ac on lines	Cat# ABIN863208	1/400	Figure 4B
NR2F2	**Rabbit**	Poly	Human	abcam	Cat# ab211776	1/100	Figure 6A-B-C-D, S6A, S6B
GATA2	**Mouse**	Mono	Human	Sigma	Cat# WH0002624M1; RRID:AB_1841726	1/50	Figure 5A-6B
CDH1	**Mouse**	Mono	Human	abcam	Cat# ab1416; RRID:AB_300946	1/500	Figure 4A
GATA3	**Goat**	Poly	Human	R&D	Cat# AF2605; RRID:AB_2108571	1/200	Figure 6B
NANOG	**Goat**	Poly	Human	R&D	Cat# AF1997; RRID:AB_355097	1/200	Figure 6A
SOX17	**Goat**	Poly	Human	R&D	Cat# AF1924; RRID:AB_355060	1/200	Figure S6A
ZO1	**Rabbit**	Poly	Human, dog, mouse rat	Invitrogen	Cat# 40-2200; RRID:AB_2533456	1/500	Figure 4A
NANOG	**Rabbit**	Poly	Human	Thermofisher	Cat# PA1-097; RRID:AB_2539867	1/200	Figure 5C-D, S5B-D
GATA4	**Rat**	Mono	Human, mouse	Invitrogen	Cat# 14-9980-82; RRID:AB_763541	1/400	Figure 5C-D, 6C-D, S5B-D, S6A, S6B
IFI16	**Mouse**	Mono	Human	Abcam	Cat# ab55328; RRID:AB_943797	1/200	Figure 5C-D, 6C-D, S6A, S6B
OTX2	**Goat**	Poly	Human	R&D	Cat# AF1979; RRID:AB_2157172	1/200	S5B-D

Secondary antibodies

Fluorochrom	Species	Reactivity	Supplyer	Cat n°	Dilution
Alexa-Fluor 488	donkey	**Rabbit**	Life Technologies	Cat#A21206; RRID:AB_2535792	1/300 or 1/1000
Alexa-Fluor 488	donkey	**Mouse**	Life Technologies	Cat#A21202; RRID:AB_141607	1/1000
Alexa-Fluor 568	donkey	**Rabbit**	Life Technologies	Cat#A10042; RRID:AB_2534017	1/1000
Alexa-Fluor 568	donkey	**Mouse**	Life Technologies	Cat#A10037; RRID:AB_2534013	1/1000
Alexa-Fluor 568	donkey	**Goat**	Life Technologies	Cat#A11057; RRID:AB_2534104	1/1000
Alexa-Fluor 647	donkey	**Goat**	Life Technologies	Cat#A21447; RRID:AB_2535864	1/1000
DyLight 550	donkey	**Rat**	Thermofisher	Cat#SA5-10027; RRID:AB_2556607	1/300
Alexa Fluor® Plus 647	donkey	**Mouse**	Life Technologies	Cat#A32787; RRID:AB_2762830	1/300

Supplemental Table 4 - Antibodies. Details of primary and secondary antibodies used for this study - related to figures 4, 5, 6 and supplemental figures S5-6

Embryo ID	Sex	# of cells	Stage	Grade
M30	M	10	M	N/A
M31	F	15	M	N/A
M32	F	9	M	N/A
M34	F	5	M	N/A
M37	M	15	M	N/A
M39	M	8	M	N/A
M42	F	11	M	N/A
M57	M	14	M	N/A
B11	M	19	B3	N/A
B22	M	15	B3	N/A
B41	M	6	B3	N/A
B24	M	16	B4	N/A
B13	M	13	B4	BA
B16	F	2	B4	CB
B18	M	5	B4	AA
B20	M	11	B4	BC
B26	F	16	B4	BB
B28	F	8	B4	CC
B38	F	19	B4	CC
B52	F	25	B4	AA
B53	F	32	B4	BB
B56	M	19	B4	BC
B48	M	20	B5	AB
B49	F	11	B5	BC
B50	F	11	B5	AA

Supplemental Table 5 - 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 - related to figure 1.

1.1.3 Discussion

Dans cet article, nous proposons une analyse intégrée du développement embryonnaire humain de J3 à J7. Nous avons combiné des données morphologiques cliniques, bioinformatiques, et protéiques d'embryons humains afin de situer dans le temps et l'espace certains moments clés du développement.

Dans les prochains paragraphes, nous discuterons de notre démarche et des résultats obtenus.

Valider l'approche intégrée chez la souris

Afin de convaincre de la qualité de notre approche bioinformatique, il nous avait paru nécessaire de l'appliquer à un jeu de données murin (Posfai et al., 2017). Cette étude se concentre essentiellement sur l'émergence du TE et de l'ICM, les cellules provenant d'embryons entre 16 et 64 cellules, soit entre le stade morula et le stade blastocyste en cours d'expansion. Au stade 64 cellules (64C) à E3.5, la spécification EPI/PrE est encore en cours dans la plupart des cellules, ce qui rend difficile l'identification de populations EPI ou PrE (Chazaud and Yamanaka, 2016). Nous avons donc dû compléter notre analyse par des données de sc-qPCR obtenues sur des embryons entre le stade 32C et 90C (Allègre et al., 2022). Nous avons obtenu un pseudotemps à cinq branches, commençant par le stade morula et se terminant par trois branches terminales que nous avons annotées EPI, TE, PrE car exprimant respectivement *Sox2*, *Cdx2*, et *Sox17*. Nous étions confiants dans notre analyse car plusieurs stades développementaux connus, tant morphologiques que protéiques, étaient retranscrits dans notre pseudotemps murin. Par exemple, nous détectons l'expression de *Cdx2* au stade morula-16C, sachant que la protéine CDX2 est restreinte aux cellules externes de la morula-16C. Ce stade coincide précisément avec l'émergence de deux lignées

dans le pseudotemps : l'une étant le TE et maintenant une expression robuste de *Cdx2*, et l'autre l'ICM. Ce n'est qu'après, au stade 32C, que l'embryon de souris passe au stade blastocyste. Ainsi chez la souris, la première spécification est initiée et achevée avant la blastulation.

Résoudre le problème de l'annotation journalière de l'embryon humain

Chez la souris, chaque jour de développement (ex : E4.5) est associé à une morphologie et une combinaison de marqueurs en immunofluorescence bien définis.
Chez l'Homme en revanche, plusieurs embryons au même jour de développement n'auront pas le même aspect morphologique ni le même profil de marqueurs (Niakan et Eggan, 2013). Ceci est dû à des variabilités biologiques inhérentes, mais également aux conditions de culture qui peuvent varier d'un laboratoire à un autre (Kirkegaard et al., 2016). En clinique, les embryons humains ne sont pas tous fécondés au même moment. Ils ne sont pas non plus congelés et décongelés à la même heure. Enfin, à la décongélation, certains embryons peuvent reprendre leur développement plus ou moins vite que d'autres. Ainsi, au 5ème jour de développement, un embryon peut commencer à caviter quand d'autres sont en cours d'expansion (Meistermann et al., 2021, Figure 1A). Avoir recourt aux jours embryonnaires ne nous semblait pas suffisamment précis pour situer les embryons dans leur développement. Les travaux de Petropoulos et Lanner illustrent cette hypothèse. En effet, ils n'ont pas pu distinguer de manière fine leurs embryons J5 les uns des autres, ce qui les a amenés à conclure que l'émergence des lignées EPI, TE, PrE était simultanée.
Pour définir précisément une hiérarchie moléculaire basée sur le développement de l'embryon, nous avons séquencé les cellules d'embryons cultivés en incubateurs timelapse. Les vidéos ont été analysées en temps réel et les embryologistes ont gradé les embryons selon la classification morphologique de Gardner et Schoolgraft

(Gardner et al., 2000). L'utilisation de cette classification nous a permis d'associer à une morphologie précise un transcriptome donné. Nous avons combiné nos données de scRNAseq à celles de jeux de données disponibles (Fogarty et al., 2017; Petropoulos et al., 2016; Yan et al., 2013), résultant un ensemble de 1751 cellules et 128 embryons entre les jours 3 et 7 de développement. L'intérêt du couplage morphologie-transcriptome a été de pouvoir réannoter les autres jeux de données, et associer un stade transcriptomique précis à des embryons simplement étiquetés « J5 ».

Spécificités de la première spécification humaine

Nous avons appliqué une stratégie d'analyse de pseudotemps à nos données de scRNAseq humaines, nous permettant de reconstituer un continuum transcriptomique des premières cellules à J3 jusqu'aux dernières disponibles à J7. Contrairement à la souris qui spécifie son ICM et TE avant la blastulation, nous observons une bifurcation des branches entre les stades B2 et B3 chez l'Homme. A partir du stade B3, toutes les cellules des embryons se situent soit dans la branche TE, soit dans la branche EPI (Meistermann et al., Figure 1F-G-H). Nous avons appliqué une analyse de vélocité des ARNm non-épissés (« RNA velocity », La Manno et al., 2018) sur la population de cellules B1-B2, juste avant la bifurcation. Bien qu'elle soit composée à instant t de cellules transcriptomiquement proches, il existe une hétérogénéité à l'intérieur même de cette population. Certaines cellules semblent se destiner à un destin EPI (n=44, 26%), et d'autres à un destin TE (n=112, 65%) comme le suggèrent les vecteurs de vélocité (Meistermann et al., 2021, Figure 3 F-G). Il n'est pas surprenant de voir davantage de cellules se diriger vers le TE, sachant que les cellules externes des

blastocystes B1-B2 sont trois fois plus nombreuses que les cellules internes (Regin et al., 2022). Afin de comprendre ce qui oriente les cellules vers le TE ou l'EPI, nous avons étudié le stade morula, car c'est à ce moment chez la souris que s'effectue le choix des destins cellulaires TE et ICM.

Nous montrons d'abord que *CDX2* n'est pas exprimé au stade morula (Figure 14), ce qui explique l'absence de CDX2 en immunofluorence à ce stade (Niakan et Eggan, 2013). En revanche, nous y détectons une expression hétérogène de *GATA3,* ce qui rappelle celle chez la souris où GATA3 (et CDX2) sont exprimés uniquement dans les cellules externes de la morula, et pas dans les cellules internes. Nous avons donc initié une étude moléculaire de la première spécification, et avons montré une expression précoce de GATA3 dès le stade morula, uniquement dans les cellules externes, qui se maintient dans le TE du blastocyste (Figure 14). Nous avons de plus marqué des morulas et blastocystes B1/B2 pour des marqueurs de polarité. Aussi, nous confirmons que la polarité se met en place dès le stade morula avec une expression baso-latérale de CDH1, et latérale de ZO-1 (Meistermann et al., 2021, Figure 4A). Nous avons dès lors émis l'hypothèse que chez l'Homme, *GATA3* est le marqueur le plus précoce de l'initiation du destin TE. Nous avons exploré cette hypothèse en collaboration avec Claudia Gerri et Kathy Niakan (voir deuxième partie des résultats). Des expériences fonctionnelles d'inhibition de la polarité ont permis de démontrer que la polarité est nécessaire à l'expression de *GATA3* et à la formation d'un TE chez l'embryon humain (Gerri et al., 2020).

Pour résumer, nous proposons que la première spécification chez l'Homme s'inscrit comme un processus dynamique initié dès la polarisation au stade morula avec l'expression de *GATA3* dans les cellules externes, et s'achevant après le stade B2-B3 où une identité transcriptomique propre au TE est détectable.

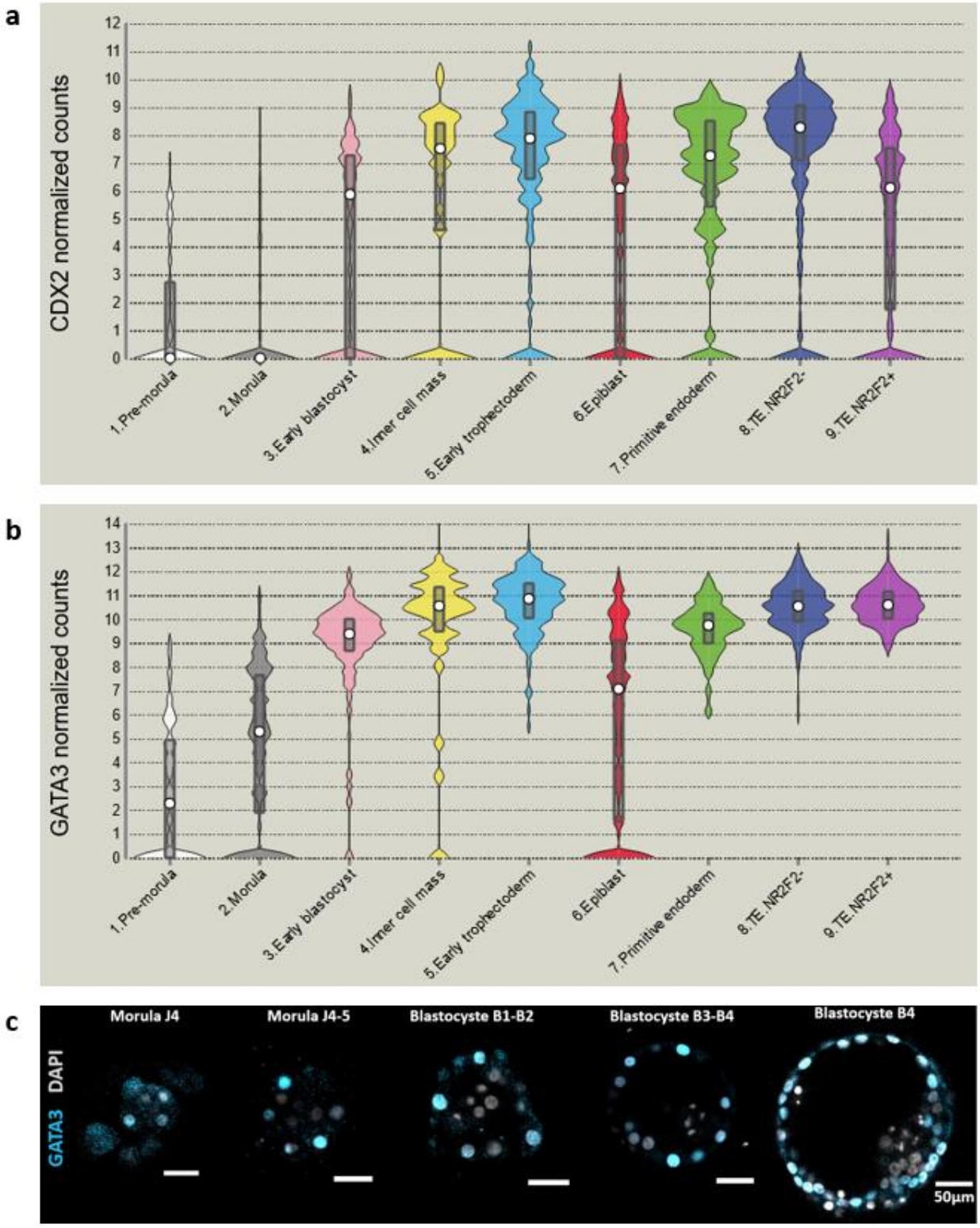


Figure 14 : GATA3/*GATA3* est un marqueur précoce du trophectoderme humain

(a-b) Violin plot de l'expression normalisée en log2 de *CDX2* et *GATA3* dans les différents groupes de cellules obtenus par réduction de dimension UMAP. (c) Cinétique d'expression de GATA3 entre le stade morula et blastocyste B4. GATA3 est détectable dès le stade morula, et uniquement dans les cellules externes.

La maturation du trophectoderme humain

Ce qui est remarquable sur la visualisation en pseudotemps, c'est la taille de branche du TE. La métrique de pseudotemps, qui correspond à une quantité de changements transcriptomiques, suggère que le TE passe par davantage d'états transcriptomiques que l'EPI ou le PrE entre J5 et J7 (Meistermann et al., 2021, Figure S2B). Grâce à l'analyse des réseaux de gènes co-exprimés par WGCNA (Langfelder and Horvath, 2008), nous avons pu raffiner la signature moléculaire des cellules qui composent le TE au cours de son développement. Nous trouvons trois identités transcriptomiques successives que sont : le TE précoce « 5.Early TE », exprimant encore des gènes de pluripotence comme *POU5F1B/OCT4,* le TE intermédiaire « 8.TE.NR2F2-», et un TE tardif « 9.TE.NR2F2+ » (Meistermann et al. Figure S4C). Dans le TE précoce, l'expression de gènes de pluripotence dans le TE supporte l'hypothèse d'un TE plastique, potentiellement toujours capable de s'engager dans un destin pluripotent. Les cellules TE de jeunes blastocystes, réagrégées, sont d'ailleurs capables de reformer des blastocystes, avec une ICM exprimant NANOG (De Paepe et al., 2013). Aussi, la plasticité du TE pourrait expliquer pourquoi des blastocystes avec une ICM gradée « C » (pas/peu de cellules) peuvent tout de même permettre une grossesse (Zhao et al., 2018). Il est envisageable qu'un mécanisme de compensation puisse permettre à certaines cellules du TE de se différencier en EPI si ce dernier n'est pas ou peu formé.

Le module *NR2F2* est le dernier module de gènes de la branche TE. Il est exprimé dans le TE des embryons aux 5-6-7ème jours, aux stades B4-B5-B6, soit immédiatement avant l'implantation. *NR2F2,* aussi appelé *COUP-TFII,* est un facteur de transcription. C'est un récepteur nucléaire orphelin, impliqué dans de nombreux

contextes biologiques comme l'angiogenèse, la formation du cœur, l'homéostasie cellulaire, l'infertilité masculine et la formation du placenta (Hubert et al., 2010; Polvani et al., 2019; Wu et al., 2016). Par immunofluorescence, nous avons été surpris de constater que NR2F2 est localisée uniquement dans le TE polaire, soit le TE en contact avec l'ICM (Meistermann et al. Figure 6A-B-C). C'est la première fois à notre connaissance qu'il est montré la présence d'un régulateur transcriptionnel dans le TE polaire de l'embryon humain. Les cellules NR2F2 polaires co-expriment GATA3 et GATA2, et notre analyse transcriptomique suggère qu'elles perdent *CDX2*. Nous montrons effectivement qu'entre le stade B4 et B6, l'apparition des cellules NR2F2 coïncide avec une perte de CDX2 dans ces mêmes cellules (Figure 15). Ce switch CDX2-NR2F2 au stade B4 illustre le passage de l'identité intermédiaire « 8.TE.NR2F2- » au TE tardif « 9.TE.NR2F2+ ».

La maturation du TE par le coté polaire suggère une communication EPI/TE dans le blastocyste tardif. Nous avons en effet mis en évidence l'activité de nombreuses voies de signalisation potentiellement impliquées dans ce dialogue moléculaire comme IL-6/IL6-R et TGF/TGFBR. Nous pensons que cette boucle de régulation EPI/TE permet la formation d'un EPI et d'un TE compétents pour la poursuite du développement. Cette hypothèse a par la suite été testée dans des blastoïdes humains traités avec des inhibiteurs de la voie IL-6 : ces derniers sont dépourvus de masse cellulaire pluripotente et ne peuvent pas s'implanter *in vitro* (Kagawa et al., 2021).

In vivo comme *in vitro,* l'embryon humain s'implante du côté polaire contrairement à la souris qui s'implante du côté mural (Aberkane et al., 2018; Hertig et al., 1956; Siriwardena and Boroviak, 2022). Autrement dit, les premières cellules de l'embryon humain à entrer en contact avec l'endomètre expriment probablement *NR2F2.*

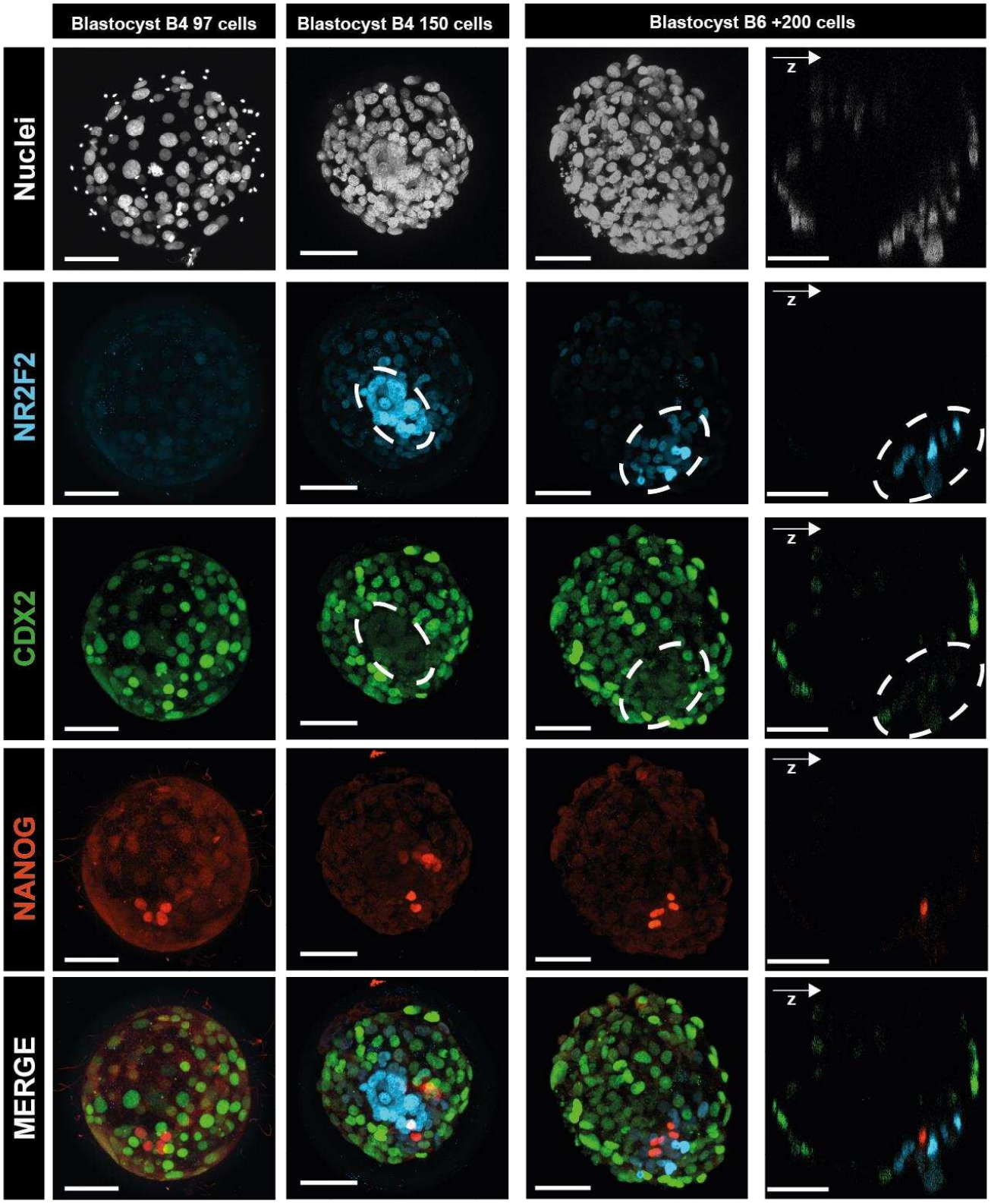


Figure 15 : dynamique d'expression de CDX2 et NR2F2 dans le trophectoderme humain
Dans un jeune blastocyste B4/J5, CDX2 est exprimé dans toutes les cellules du TE. NR2F2 est détecté à partir du stade B4 dans le TE polaire qui n'exprime plus CDX2.

De manière cohérente, nos analyses montrent que le TE polaire *NR2F2* à J7 exprime des facteurs comme le Placental Growth Factor (PIGF) ou le Vascular Endothelial Growth Factor (VEGF), deux protéines nécessaires à une placentation normale (Binder et al., 2016; Chau et al., 2017).

L'ensemble de ces données suggère que *NR2F2* n'est pas uniquement un marqueur du TE polaire, mais aussi un régulateur nécessaire à la maturation du trophoblaste et du développement du placenta. Ces hypothèses pourraient être explorées avec des inhibiteurs de l'activité de NR2F2 (Le Guével et al., 2017) ou par *knock-out* dans des modèles d'embryons humains. Il serait intéressant de constater que les embryons ou blastoïdes traités seraient alors incapables de s'implanter *in vitro,* ou de différencier leur TE dans la lignée trophoblastique.

La seconde spécification dans l'embryon humain préimplantatoire

Nous avons ensuite étudié l'établissement de l'EPI du PrE dans le blastocyste humain. Nous avons tout d'abord mis en évidence le module de gène *IFI16,* exprimé quand l'embryon atteint le stade B4/jour 5. IFI16 est un élément de réponse à l'interferon-β. Il est impliqué dans plusieurs processus biologiques comme la prolifération, la différenciation, la sénescence, et la réponse immunitaire, notamment face à de l'ADN exogène (Dell'Oste et al., 2015; Unterholzner et al., 2010). Son rôle dans le développement n'est pas documenté, mais des analyses récentes dans l'embryon humain et de ouistiti ont montré qu'il fait partie des réseaux de gènes liés à la pluripotence (Boroviak et al., 2018). Dans nos analyses, *IFI16* a un profil transcriptomique proche du marqueur de l'épiblaste *NANOG.* Nous avons effectué des

marquages d'embryons du stade morula au stade B5, et montrons la présence de IFI16 à partir du stade B4 exclusivement dans l'ICM (Meistermann et al., 2021, Figure 5A-B). IFI16 est détectable soit dans des cellules NANOG, soit dans des cellules SOX17/GATA4 (PrE). Nous n'avons pas détecté de cellules IFI16+/NANOG-, et environ 40% des cellules GATA4+ du blastocyste que nous avons marquées expriment IFI16 à des niveaux variables. Seules 5% des cellules GATA4+ n'expriment pas IFI16 (Meistermann et al. Figure 5 E-F). Nous pensons que les cellules NANOG+/IFI16+ correspondent à de l'EPI spécifié, et que les cellules GATA4+/IFI16+ correspondent à une population qui a initié la différenciation en PrE. Cette hypothèse est supportée par notre analyse de clustering UMAP qui détecte une population intermédiaire « EPI.PrE ». Cette population est principalement constituée de cellules de blastocystes à J5-6, et pas/peu à J7. L'analyse par vélocité ARNm montre que 64% (n=18) des cellules du cluster pointent en direction du PrE, contre 36% (n=10) dirigées vers l'EPI (Meistermann et al., 2021, Figure 5G-H).

En 2012, Roode et collègues considéraient la spécification du PrE comme achevée à jour 6-7 dans les cellules NANOG et GATA4 exclusives (Roode et al., 2012). Cependant, même à J7, la plupart des cellules SOX17/GATA4 expriment toujours le facteur de pluripotence OCT4. Dans notre analyse, l'EPI est spécifié au stade B3 avant le PrE au stade B4, et certaines cellules de l'EPI semblent s'orienter vers le PrE (Meistermann et al., 2021, Figure S5E). Nous en concluons que la spécification du PrE est plus tardive que celle de l'EPI. Nous pensons que la population intermédiaire EPI.PrE peut correspondre à de l'EPI en cours de différenciation vers le PrE. Ces cellules peuvent être amenées à perdre IFI16 pour devenir du PrE définitif. En utilisant des données de scRNAseq d'embryons à J8 (Zhou et al., 2019), nous montrons l'existence de deux clusters EPI et PrE dont les vecteurs de vélocité ne pointent plus vers l'un ou l'autre. Cette observation est conforme avec des marquages d'embryons

humains à J8 montrant des cellules GATA6 (PrE) n'exprimant pas OCT4 (EPI) (Deglincerti et al., 2016).

Pour résumer, nous émettons l'hypothèse que la spécification de l'EPI et du PrE n'est pas achevée dans toutes les cellules de l'embryon humain préimplantatoire, et qu'elle se poursuit dans l'embryon postimplantatoire. Nous pensons que IFI16 est un marqueur dynamique de la seconde spécification, et qu'il permet de localiser les cellules intermédiaires entre l'EPI et le PrE.

Notre analyse bioinformatique n'a pas permis de définir avec certitude une identité transcriptomique ICM d'où émergerait l'EPI et le PrE. Cette problématique était jusqu'à présent partagée par la communauté (Petropoulos et al., 2016; Weltner and Lanner, 2021).

De récentes analyses ont permis d'isoler une sous-population de cellules présentes au stade blastocyste J5, transitoire, en aval du stade morula et avant la spécification de l'EPI, qui pourrait correspondre à l'ICM. (Radley et al., 2022). Cette population co-exprime OCT4 et la protéine membranaire LAMA4. A J6-J7, dans les cellules exclusives pour SOX17 ou OCT4, LAMA4 n'est plus détectable. Les auteurs ne montrent ces marquages que dans des blastocystes à partir de B2-B3. Il serait intéressant d'effectuer des marquages LAMA4 dans des morulas et blastocystes B1-B2. Ceci pourrait permettre de situer précisément dans le temps l'émergence de la population ICM.

La maturation postimplantatoire des lignées embryonnaires humaines

Nos analyses ont montré la présence de deux types de TE préimplantatoire à J7 : un polaire NR2F2+/CDX2-, et l'autre mural NR2F2-/CDX2+. La continuité de ces types cellulaires dans notre modèle de pseudotemps et l'absence de bifurcation dans la branche TE suggère que le TE polaire est le premier à maturer, et que cette maturation

est amenée à se propager au TE mural après J7. Aussi, la présence d'IFI16 dans certaines cellules GATA4/SOX17, et la co-expression OCT4/SOX17/GATA4 dans les cellules de l'ICM à J5-J6-J7 suggère que la seconde spécification EPI/PrE n'est pas achevée dans toutes les cellules de l'embryon préimplantatoire. Collectivement, ces observations suggèrent que les lignées EPI, TE, et PrE de l'embryon humain continuent à se spécifier et/ou maturer dans l'embryon postimplantatoire.

Pour tester ces hypothèses, nous avons cultivé, fixé, et marqué des embryons humains jusqu'au 8ème jour de développement (Deglincerti et al., 2016; Shahbazi et al., 2016). Dans ces conditions de culture, nous ne pouvions pas utiliser la classification morphologique de Gardner pour grader et estimer la qualité embryonnaire. Nous nous sommes donc basés sur trois critères pour estimer la survie de l'embryon :

- Son attachement à jour 7-8
- Sa taille et sa croissance
- Son pourcentage de lyse apparent

Nous avons attribué une note de 1 à 3 pour grader les embryons. 1 étant un embryon avec de bons critères morphologiques, 2 un embryon à l'aspect intermédiaire, et 3 un embryon lysé. Nous avons considéré les embryons gradés 1 et 2 comme vivants et en cours de développement.

En se basant sur ces trois critères, nous estimons le taux de survie des embryons à J8 entre 60-70%. A J8, l'embryon attaché *in vitro* a la forme d'un « d'œuf au plat », avec des bords fins de l'ordre de 40µm d'épaisseur et un centre plus épais en moyenne à 120µm d'épaisseur. Dans la plupart des cas, nous avons observé des cellules exprimant NANOG au centre et en bas de l'embryon, signe que l'embryon s'est attaché à la plaque de culture du côté de l'ICM, par le TE polaire. Nous avons

systématiquement marqué les embryons par au moins un marqueur de l'EPI. La présence de l'EPI à jour 8 est notre critère moléculaire de survie embryonnaire.

Par immunofluorescence, nous avons observé que NR2F2 était progressivement exprimée dans tout le trophoblaste de l'embryon entre J7 et J8 (Figure 16A, B). Aussi, nous détectons l'hormone β-HCG dans les cellules externes du trophoblaste à J8, suggérant une différenciation du cytotrophoblaste en syncytiotrophoblaste (Figure 16 B). Entre J9-J10, nous observons que NR2F2 est plus fortement exprimé au centre de l'embryon qu'en périphérie (Figure 16C). Nos analyses transcriptomiques postimplantatoires suggèrent que NR2F2 est perdu lors de la différenciation du CT en ST et EVT (données non publiées). Nous aurions aussi pu confirmer la différenciation en ST avec des marquages membranaires qui auraient montré des cellules multinucléés (Castel et al., 2020; Deglincerti et al., 2016). Dans les embryons attachés à J8, nous montrons que la quasi-totalité des cellules NANOG+ de l'épiblaste expriment IFI16. IFI16 est encore détectable dans 25% des cellules GATA4, et la proportion de cellules GATA4+/IFI16- est passée à 19% à J8 contre 5% en postimplantation. Il est envisageable que l'expression de IFI16 diminue progressivement dans le PrE après J8 au fur et à mesure du développement. Nous avons de plus montré qu'OTX2 est exclusivement exprimé dans les cellules GATA4 du PrE péri-implantatoire. OXT2 est faiblement exprimé au stade blastocyste B4 et fortement exprimé dans le PrE à J8 (Meistermann et al. 2021, Figure S5 B-C-D). Curieusement, il semble qu'OTX2 soit d'abord exprimé dans les cellules du PrE les plus proches de l'EPI, puis se répand à toutes les cellules du PrE. Cette observation semble également possible dans les immunofluorescences de Zhou et collègues qui ne précisent cependant pas le jour de développement de leurs embryons (Zhou et al., 2019, Figure 2B). Aussi, la progression d'OTX2 aux cellules GATA4 suggère une maturation postimplantatoire du PrE. Il serait intéressant de constater que la

progression d'OTX2 dans le PrE coïncide avec la perte d'IFI16 dans ces mêmes cellules.

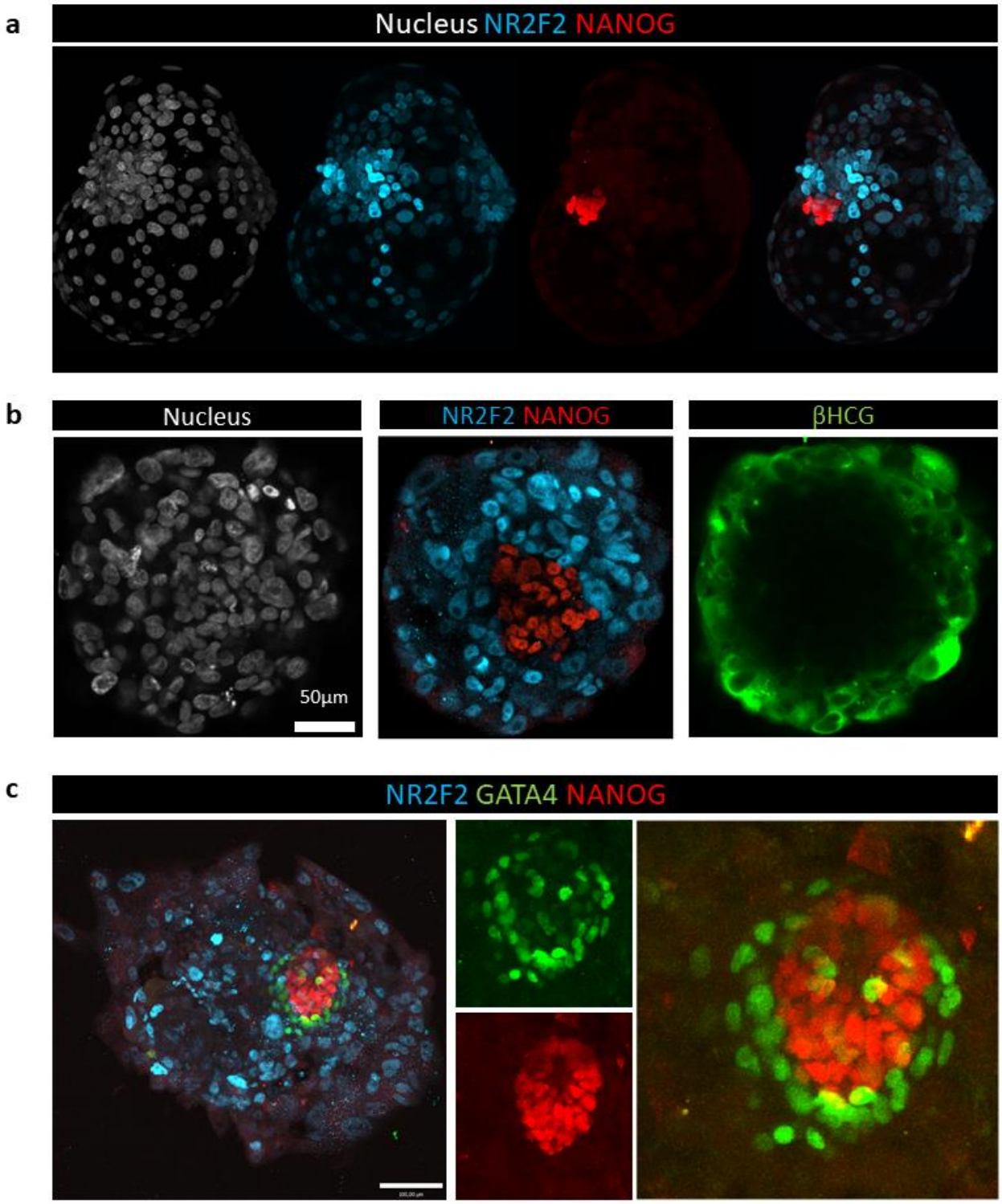


Figure 16 : dynamique d'expression de NR2F2 dans l'embryon humain postimplantatoire

(a) Blastocyste B6 à J7. NR2F2 est fortement exprimé dans le TE polaire face à l'épiblaste NANOG+, et commence à se répandre au trophectoderme mural
(b) Embryon attaché à J8 cultivé en milieu IVC. NR2F2 est exprimé dans tout le trophoblaste qui exprime la β-HCG
(c) Embryon attaché à J10 cultivé en milieu CMRL. La cavité amniotique est formée dans les cellules de l'épiblaste NANOG+.

1.2 Manuscrit secondaire #1 : “Initiation of a conserved trophectoderm program in human, cow and mouse embryos”. Gerri et al. 2020

1.2.1 Contexte de l’étude

L’initiation de la première spécification chez la souris est largement documentée. Il est admis que suite à la compaction des cellules au stade morula, l’inactivation de la voie HIPPO dans les cellules externes polarisées active l’expression de gènes régulateurs du TE comme *Cdx2* et *Gata3*. Les cellules internes apolaires maintiennent l’expression de OCT4/SOX2 et s’orientent par défaut vers un destin ICM.

Chez l’Homme, ce mécanisme de spécification n’a pas été étudié. Notre analyse transcriptomique nous a permis d’identifier que le premier marqueur du TE exprimé au stade morula est *GATA3,* et non *CDX2.*

Suite à une discussion avec le laboratoire de Kathy Niakan, nous avons décidé de joindre nos analyses à leur étude mécanistique comparative de la compaction humaine, bovine et murine.

1.2.2 Manuscrit Gerri et al. 2020

Gerri C, Afshan McCarthy, Alanis-Lobato G, Demtschenko A, **Bruneau A**, Loubersac S, Fogarty NME, Hampshire D, Elder K, Snell P, Christie L, David L, Van de Velde H, Fouladi-Nashta AA, Niakan KK. 2020a. Initiation of a conserved trophectoderm program in human, cow and mouse embryos. *Nature* **587**:443–447. doi:10.1038/s41586-020-2759-x

Article

Initiation of a conserved trophectoderm program in human, cow and mouse embryos

https://doi.org/10.1038/s41586-020-2759-x

Received: 10 May 2019

Accepted: 21 August 2020

Published online: 23 September 2020

Check for updates

Claudia Gerri[1], Afshan McCarthy[1], Gregorio Alanis-Lobato[1], Andrej Demtschenko[2], Alexandre Bruneau[3], Sophie Loubersac[3,4], Norah M. E. Fogarty[1,9], Daniel Hampshire[5], Kay Elder[6], Phil Snell[6], Leila Christie[6], Laurent David[3,7], Hilde Van de Velde[2,8], Ali A. Fouladi-Nashta[5] & Kathy K. Niakan[1,10]✉

Current understandings of cell specification in early mammalian pre-implantation development are based mainly on mouse studies. The first lineage differentiation event occurs at the morula stage, with outer cells initiating a trophectoderm (TE) placental progenitor program. The inner cell mass arises from inner cells during subsequent developmental stages and comprises precursor cells of the embryo proper and yolk sac[1]. Recent gene-expression analyses suggest that the mechanisms that regulate early lineage specification in the mouse may differ in other mammals, including human[2–5] and cow[6]. Here we show the evolutionary conservation of a molecular cascade that initiates TE segregation in human, cow and mouse embryos. At the morula stage, outer cells acquire an apical–basal cell polarity, with expression of atypical protein kinase C (aPKC) at the contact-free domain, nuclear expression of Hippo signalling pathway effectors and restricted expression of TE-associated factors such as GATA3, which suggests initiation of a TE program. Furthermore, we demonstrate that inhibition of aPKC by small-molecule pharmacological modulation or Trim-Away protein depletion impairs TE initiation at the morula stage. Our comparative embryology analysis provides insights into early lineage specification and suggests that a similar mechanism initiates a TE program in human, cow and mouse embryos.

The length of pre-implantation development varies between human, cow and mouse embryos. We therefore initially performed a morphokinetic analysis to benchmark the initiation and duration of key morphological events (Extended Data Fig. 1, Supplementary Videos 1–3, Supplementary Table 1). We next sought to determine whether TE-associated genes may be expressed before blastocyst formation, which would suggest initiation of a differentiation program. We mined published human pre-implantation single-cell RNA-sequencing (scRNA-seq) datasets[4,5,7], and observed heterogeneous expression of *GATA3* in human morula cells (Fig. 1a, b). Re-analysis of a human pre-implantation chromatin accessibility dataset[8] revealed enrichment of GATA and TEAD motifs in open chromatin regions at the morula stage, whereas at the 8-cell stage we observed motif enrichment of genes involved in embryonic genome activation, such as *DUXA*, *DUX4*[9] and *ZSCAN4*[10] (Fig. 1c). In the mouse, the transcription factor GATA3 functions to promote TE differentiation in outer cells at the morula stage, downstream of the Hippo signalling transcription factor TEA-domain family member 4 (TEAD4)[11]. Using immunofluorescence analysis, we confirmed that TEAD4 is detected in all nuclei of human morula stage embryos (Extended Data Fig. 2a–d), similar to the mouse[12]. We also observed colocalization of GATA3 and the TEAD4 cofactor Yes-associated protein 1 (YAP1) in outer cells at the morula stage in human, cow and mouse embryos (Fig. 1d–f, Extended Data Fig. 2e–k), consistent with previous findings in the mouse[11,13,14]. Similar to the mouse[13], we observed overlapping nuclear expression of YAP1 and another TEAD4 transcriptional cofactor, WW-domain-containing transcription regulator protein 1 (WWTR1), in outer and TE cells in the human morula and blastocyst stages, respectively (Extended Data Fig. 3a, b). GATA2 is considered a TE marker in human blastocysts[3–5,15]. Importantly, GATA2 was not detected at the morula stage (Extended Data Fig. 2l), despite its restriction to TE cells at the blastocyst stage (Extended Data Fig. 2m). Altogether, this suggests that GATA3 and Hippo signalling components are conserved in distinguishing cells that initiate a TE program in human, cow and mouse morula stage embryos.

We next sought to identify genes that were co-expressed with *GATA3* in human morula cells using scRNA-seq datasets[4,5,7]. We identified 22 TE-enriched genes that showed a positive correlation (Pearson's $r > 0.25$) with *GATA3* when comparing all human morula cells (Supplementary Table 2). Genes related to epithelial cell formation (*KRT18*, *CLDN4*, *RAB20* and *RAB25*) and placenta morphogenesis (*PTGES*, *TFEB*

[1]Human Embryo and Stem Cell Laboratory, The Francis Crick Institute, London, UK. [2]Department of Reproduction and Immunology, Vrije Universiteit Brussel, Brussels, Belgium. [3]Université de Nantes, CHU Nantes, Inserm, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ITUN, Nantes, France. [4]Service de Biologie de la Reproduction, CHU Nantes, Université de Nantes, Nantes, France. [5]Department of Comparative Biomedical Sciences, Royal Veterinary College, London, UK. [6]Bourn Hall Clinic, Cambridge, UK. [7]Université de Nantes, CHU Nantes, Inserm, CNRS, SFR Santé, Inserm UMS 016, CNRS UMS 3556, Nantes, France. [8]Center for Reproductive Medicine, UZ-Brussel, Brussels, Belgium. [9]Present address: Centre for Stem Cells and Regenerative Medicine, King's College London, London, UK. [10]Present address: The Centre for Trophoblast Research, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. ✉e-mail: kathy.niakan@crick.ac.uk

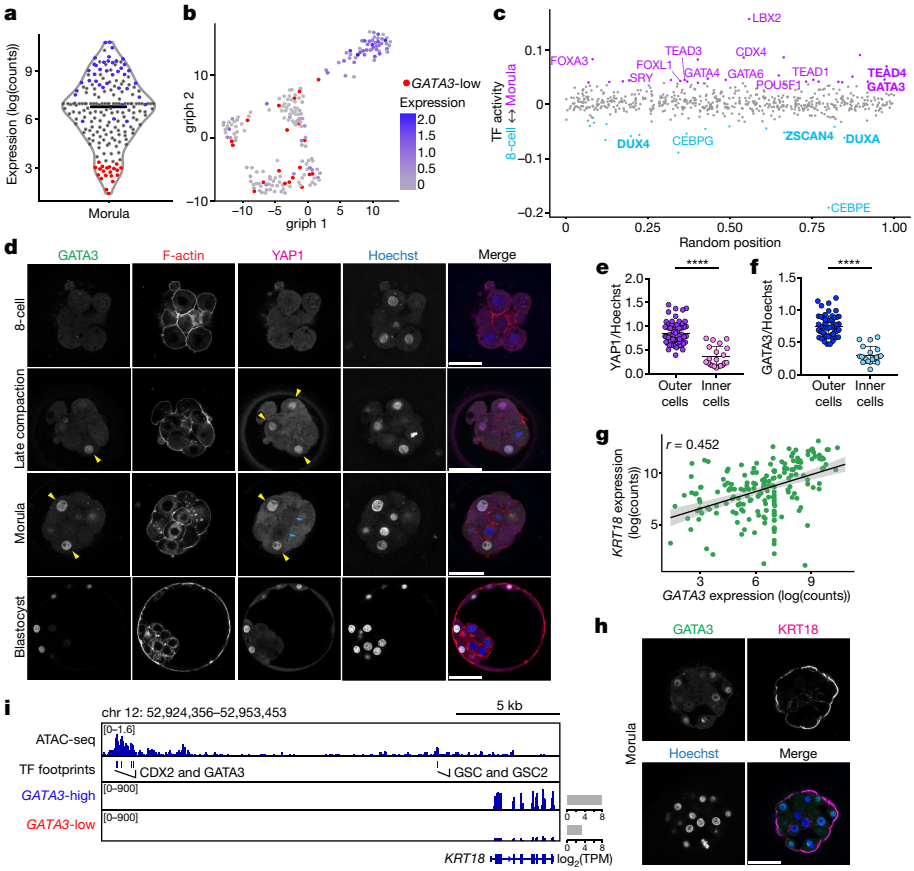


Fig. 1 | Transcriptional and protein expression differences between cells in human morula embryos. a, **b**, log-transformed size-factor-normalized expression of *GATA3* in human morula cells. $n = 197$ cells. **a**, Violin plot of cells with lowest (red) versus highest (blue) *GATA3* expression. Black line corresponds to the median. **b**, Dimensionality reduction analysis using graph inference of population heterogeneity (griph). Cells coloured according to *GATA3* expression. **c**, Chromatin accessibility from assay for transposase-accessible chromatin using sequencing (ATAC-seq) data of human morula (purple) versus 8-cell (cyan) stage embryos. Coloured dots represent transcription factors (TF) with a significant change in activity score (false discovery rate (FDR)-corrected $P < 0.05$; two-tailed Student's *t*-test). A few representative transcription factors are noted and, those highlighted in bold, are discussed in the text. **d**, Immunofluorescence analysis of human embryos at 8-cell ($n = 5$), late compaction ($n = 5$), morula ($n = 10$) and expanded blastocyst ($n = 4$) stages. Outer (yellow arrowheads) versus inner (cyan arrows) cells show differences in YAP1 and GATA3 expression. **e**, **f**, Quantification of YAP1 (**e**) and GATA3 (**f**) normalized fluorescence intensity in human morula stage embryos ($n = 95$ cells for YAP1, $n = 80$ cells for GATA3 from 10 embryos). Data are presented as mean ± s.d. Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. **g**, Scatter plot showing positive correlation of *GATA3* and *KRT18* expression in human morula cells. $n = 197$ cells. *r*, Pearson correlation coefficient. Values displayed as log-transformed size-factor-normalized counts. The black line corresponds to a linear regression model fitted to the data with 95% confidence bands. **h**, Immunofluorescence analysis of human morula embryos ($n = 3$). **i**, Genome browser view of the ATAC-seq signal at the *KRT18* locus. High-confidence peaks (FDR-corrected $P < 0.001$; MACS2 dynamic Poisson model) were used to identify transcription factor motifs. The average expression of *KRT18* in *GATA3*-high- and *GATA3*-low-expressing cells at the morula stage is shown, and the transcripts per million (TPM) units indicated. Scale bars, 50 µm. More details about statistics and reproducibility are provided in Methods.

and *PLAC8*), and genes encoding transporter subunits (*ATP6V1B1*, *ATP6V1C2*, *FXYD4*, *ATP6V0A4* and *SLC7A2*), positively correlate with *GATA3* (Fig. 1g, Extended Data Fig. 4a–c, f, Supplementary Table 3). We observed that vestigial-like protein 4 (*VGLL4*), a transcriptional cofactor and regulator of TEAD transcriptional activity[16,17], also showed positive correlation with *GATA3* (Extended Data Fig. 4d, f, Supplementary Table 3). Immunofluorescence analysis confirmed specific expression of KRT18 in outer and TE cells in human morula and blastocyst stage embryos, respectively (Fig. 1h, Extended Data Fig. 3c), as previously described[18]. Furthermore, we identified enrichment of GATA3-binding motifs at the *KRT18* locus in open chromatin regions (Fig. 1i), which suggests that *KRT18* may be a direct target of GATA3 in human embryos. We detected positive correlation with grainyhead-like transcription factor 2 (*GRHL2*) (Extended Data Fig. 4a, f, Supplementary Table 3), a gene that is important for epithelial morphogenesis and trophoblast branching in mouse embryos[19], and enrichment of GATA3-binding motifs upstream of the *GRHL2* locus (Extended Data Fig. 3d). Using immunofluorescence analysis, we observed that GRHL2 was expressed in both outer and inner cells at the morula stage (Extended Data Fig. 3e), then specifically restricted to the TE in blastocysts (Extended Data Fig. 3f). Next, we analysed scRNA-seq datasets to identify genes that exhibit an anti-correlated expression pattern to *GATA3* in human morula cells, and which could be putative inner-cell-associated markers. Genes involved in embryonic stem cell pluripotency and/or genes enriched in the epiblast or inner cell mass (ICM) of the blastocyst—such as *ZNF207*[20] and *KLF17*[4,21]—were transcriptionally negatively correlated with *GATA3* (Fig. 2a, Extended Data Fig. 4e, f, Supplementary Tables 4, 5). Altogether, our analysis suggests transcriptional differences between inner and outer cells in human morula stage embryos, with outer cells initiating a TE program.

In the mouse morula, the transcription factor SOX2 is specifically restricted to inner cells by members of the Hippo pathway and is considered the first marker of ICM pluripotency[22] (Extended Data Fig. 5a, b). At the 8-cell stage, SOX2 was detected in a few blastomeres in cow

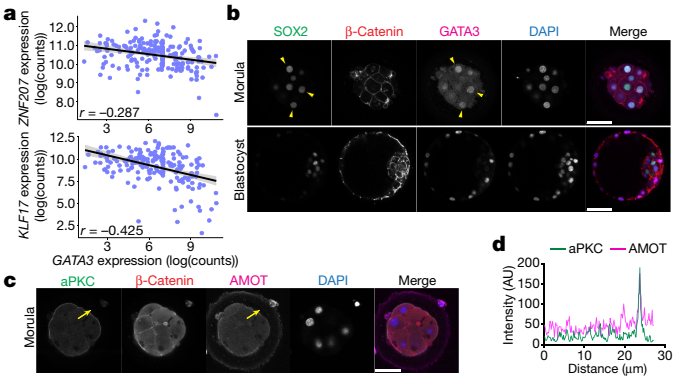


Fig. 2 | Apical expression of aPKC and AMOT in outer cells in human morula embryos, in which SOX2 expression is retained. a, Scatter plots showing negative correlation of *GATA3* versus *ZNF207* or *KLF17* expression in human morula cells. $n = 197$ cells. *r*, Pearson correlation coefficient. Values are displayed as log-transformed size-factor-normalized counts. The black line corresponds to a linear regression model fitted to the data with 95% confidence bands. **b**, Immunofluorescence analysis of human embryos at the morula ($n = 6$) or expanded blastocyst ($n = 5$) stages. Outer cells expressing SOX2 and GATA3 (yellow arrowheads). **c**, Immunofluorescence analysis of human morula stage embryos ($n = 10$). **d**, Fluorescence intensity profile of aPKC and AMOT shown along the yellow arrows in the human morula stage embryo shown in **c**. AU, arbitrary units. Scale bars, 50 μm. More details about statistics and reproducibility are provided in Methods.

embryos, whereas in human embryos SOX2 was expressed in all nuclei (Extended Data Fig. 5c, d). Expression of SOX2 in all nuclei continues in both human and cow embryos up to the formation of an early blastocyst, which is in contrast to an earlier restriction in the mouse (Fig. 2b, Extended Data Fig. 5c, e–g). SOX2 was eventually restricted to cells of the ICM in expanded blastocysts (Fig. 2b, Extended Data Fig. 5c, g), confirming previously published data in human embryos[18]. Altogether, these data indicate that SOX2—the specific inner cell marker in the mouse—displays a different expression pattern in human and cow embryos, which suggests that there are differences in the regulation of the ICM between these species.

Given the restricted localization of YAP1 and GATA3 at the morula stage in these three species, we next sought to investigate upstream regulators of this pathway. We analysed the expression pattern of aPKC and angiomotin (AMOT; a modulator of the Hippo pathway), both of which influence YAP1 nuclear localization in mouse outer cells[23]. We confirmed that, in the mouse morula, aPKC and AMOT were expressed at the apical membrane of outer cells, whereas in inner cells AMOT and E-cadherin were enriched at the cell–cell contact sites (Extended Data Fig. 6a–c). We detected a similar expression pattern in human and cow morula stage embryos, in which aPKC and AMOT were colocalized at the apical domain of outer cells and β-catenin was expressed at the basolateral domain (Fig. 2c, d, Extended Data Fig. 6d–g). Moreover, we observed colocalization of aPKC and its partner PARD6B at the apical domain of cells in human embryos (Extended Data Fig. 6h, i). These data reveal differential cell polarization between outer and inner cells in human and cow embryos, with apical and basolateral proteins showing an expression pattern similar to that of the mouse[24,25].

Our data suggest a functional link between cell polarity and TE initiation at the morula stage in human and cow embryos. To test this hypothesis, we used the aPKC inhibitor CRT0276121—this inhibitor is a derivative of CRT0103390, which has previously been shown to specifically inhibit aPKC in various biological and cellular contexts[26–28]. Following dose–response experiments (Extended Data Fig. 7a, b, Supplementary Table 6), we observed that YAP1 was restricted to the cytoplasm of outer cells in mouse embryos treated with the aPKC inhibitor (Fig. 3a, b), consistent with previous aPKC knockdown and knockout studies[23,29,30]. We also observed that inhibition of aPKC led to reduced GATA3 expression at the morula stage (Fig. 3a, c), which indicates that aPKC is required to initiate TE-associated gene expression in outer cells. Moreover, SOX2 was ectopically expressed in outer cells at the morula stage following aPKC inhibition in the mouse (Extended Data Fig. 8a–c). These data phenocopy the effects observed in the *Yap1;Wwtr1*-null mutants, and the effects of ROCK inhibition in mouse embryos[31]. Using morphokinetic analysis, we could not detect differences in cleavage rates between control and inhibitor-treated embryos (Extended Data Fig. 8d, e). Whereas control mouse embryos treated with dimethyl sulfoxide (DMSO) developed to the blastocyst stage, embryos treated with the aPKC inhibitor did not cavitate and underwent developmental arrest at the morula stage (Extended Data Fig. 8f, g), phenocopying aPKC-null mutant embryos[23,30].

We next analysed the effects of aPKC inhibition on cow and human embryos by treating them with the aPKC inhibitor from precompaction until the morula stage, following dose–response experiments (Extended Data Fig. 7c–f, Supplementary Tables 7, 8). Inhibition of aPKC in cow embryos resulted in a reduction of nuclear YAP1 and GATA3 expression in outer cells at the morula stage (Fig. 3d–f). Similarly, human embryos at the morula stage exhibited reduced expression of YAP1 and GATA3 in outer cells after aPKC inhibition (Fig. 3g–i). Moreover, when both cow and human embryos were allowed to develop to the blastocyst stage, the DMSO-treated embryos were able to form expanded blastocysts, whereas a considerable number of the embryos treated with the aPKC inhibitor arrested at cavitation (Extended Data Fig. 8h–k). aPKC inhibition had no discernible effect on SOX2 expression (Extended Data Fig. 8l–q), which was in notable contrast to the mouse. As expected, TEAD4 expression was unchanged in both mouse and human embryos treated with the aPKC inhibitor (Extended Data Fig. 9a–f).

We sought to further test our hypothesis by applying a Trim-Away protein depletion method. Trim-Away has recently been reported to induce rapid and efficient degradation of proteins of interest in mouse oocytes and embryos[32–34]. First, we optimized electroporation of *mCherry-TRIM21* mRNA with an antibody against aPKC in mouse embryos at the 4-cell stage (Extended Data Figs. 9g–j, 10a, Supplementary Table 9). We confirmed that expression of endogenous aPKC was reduced following *mCherry-TRIM21* mRNA and aPKC antibody electroporation, as compared to the control mouse embryos in which only *mCherry-TRIM21* mRNA was provided (Extended Data Fig. 10b, c). Similar to the effect observed with the aPKC inhibitor, we detected a reduction of YAP1 and GATA3 protein expression in embryos electroporated with *mCherry-TRIM21* mRNA and aPKC antibody (Extended Data Fig. 10d–f). We also observed ectopic induction of SOX2 in the outer cell of mouse morula stage embryos (Extended Data Fig. 10g, h). Moreover, TEAD4 and E-cadherin expression were unaffected in aPKC Trim-Away mouse embryos (Extended Data Fig. 10b, i, j), consistent with inhibitor treatment (Extended Data Fig. 9a, b) and knockdown of *Pard6b*[35], respectively. Altogether, this suggests that although the effect of aPKC Trim-Away was not as pronounced as that of the inhibitor, the results were consistent and there was no obvious toxicity. Electroporation of 4-cell stage human embryos led to a similar reduction of aPKC, YAP1 and GATA3 protein expression at the morula stage compared to control embryos (Fig. 3j–l, Extended Data Fig. 11a, c, d), confirming the effect seen with the aPKC inhibitor. Despite attempts to optimize electroporation in cow embryos, we were unable to identify a concentration of Trim-Away components that affected YAP1 and GATA3 protein expression without affecting embryo viability (Extended Data Fig. 11b, e–h, Supplementary Table 10).

Altogether, we propose that cell polarity—through aPKC activity—initiates a TE program at the morula stage in the outer cells of human and cow embryos (Extended Data Fig. 11i), similar to the mouse[23,29,30]. Our data suggest that aPKC sequesters AMOT at the

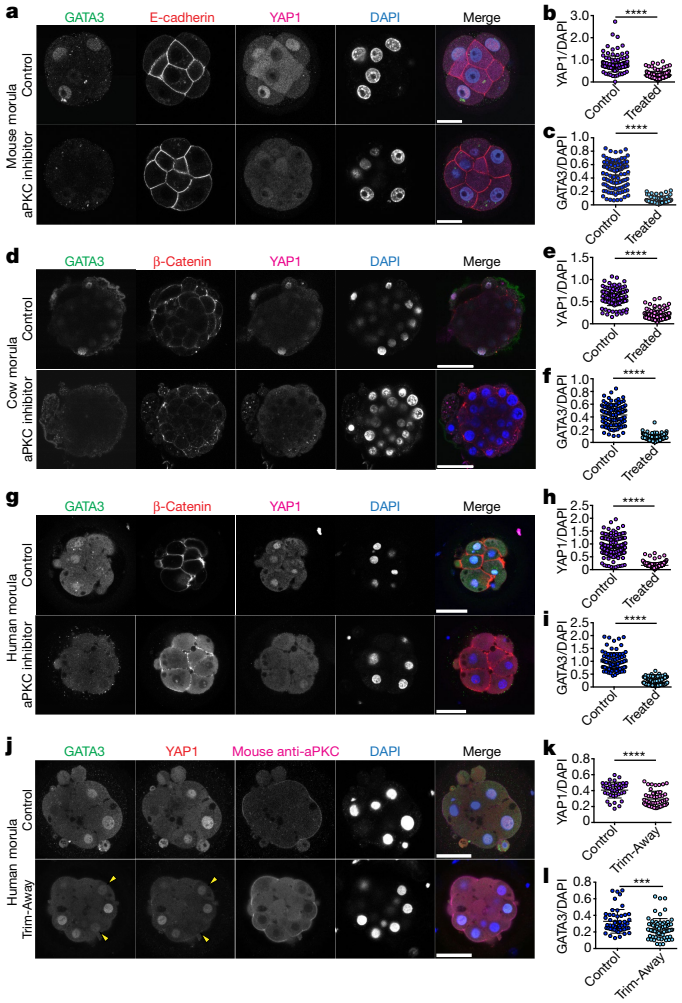


Fig. 3 | aPKC activity is required for YAP1 and GATA3 expression in mouse, cow and human morula embryos. a, Immunofluorescence analysis of control and aPKC-inhibitor-treated mouse morula stage embryos. **b**, **c**, Quantification of YAP1 (**b**) and GATA3 (**c**) normalized fluorescence intensity in outer cells of embryos shown in **a** ($n = 243$ cells for YAP1 from 28 embryos, $n = 191$ cells for GATA3 from 26 embryos). Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. **d**, Immunofluorescence analysis of control and aPKC-inhibitor-treated cow morula embryos. **e**, **f**, Quantification of YAP1 (**e**) and GATA3 (**f**) normalized fluorescence intensity in outer cells of embryos shown in **d** ($n = 209$ cells for YAP1 from 19 embryos, $n = 218$ cells for GATA3 from 21 embryos). Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. **g**, Immunofluorescence analysis of control and aPKC-inhibitor-treated human morula embryos. **h**, **i**, Quantification of YAP1 (**h**) and GATA3 (**i**) normalized fluorescence intensity in outer cells of embryos shown in **g** ($n = 406$ cells for YAP1 from 37 embryos, $n = 218$ cells for GATA3 from 21 embryos). Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. **j**, Immunofluorescence analysis of human morula stage embryos in control and Trim-Away conditions. $n = 2$ independent experiments for YAP1 analysis. Decreasing YAP1 and GATA3 expression (yellow arrowheads). **k**, **l**, Quantification of YAP1 (**k**) and GATA3 (**l**) normalized fluorescence intensity in outer cells of embryos shown in **j** ($n = 88$ cells for YAP1 from 8 embryos, $n = 116$ cells for GATA3 from 11 embryos). Two-tailed Student's *t*-test, $^{***}P = 0.00039$ for GATA3 distribution, $^{****}P < 0.0001$ for YAP1 distribution. Data are presented as mean ± s.d. Scale bars, 25 µm (mouse embryos), 50 µm (cow and human embryos). More details about statistics and reproducibility are provided in Methods.

apical domain, thus keeping the Hippo signalling pathway in an inactive state. YAP1 subsequently translocates to the nucleus, where, together with TEAD4, it promotes the transcriptional activation of a TE program to support cavitation and the formation of a blastocyst (Extended Data Fig. 11i). Additional molecular characterization and functional analyses are needed in human and cow embryos to elucidate how differences in cell polarity leads to differential Hippo signalling in outer and inner cells to drive a placental progenitor program, and whether this involves mechanisms that are conserved or divergent compared to the mouse. Although our data reveal a molecular cascade that leads to the initiation of a TE program, cells are unlikely to be committed at this stage—which is supported by studies that suggest that determination of cell fate determination occurs later[5,36].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2759-x.

1. Cockburn, K. & Rossant, J. Making the blastocyst: lessons from the mouse. *J. Clin. Invest.* **120**, 995–1003 (2010).
2. Niakan, K. K. & Eggan, K. Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Dev. Biol.* **375**, 54–64 (2013).
3. Fogarty, N. M. E. et al. Genome editing reveals a role for OCT4 in human embryogenesis. *Nature* **550**, 67–73 (2017)
4. Blakeley, P. et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3151–3165 (2015).
5. Petropoulos, S. et al. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* **165**, 1012–1026 (2016).
6. Berg, D. K. et al. Trophectoderm lineage determination in cattle. *Dev. Cell* **20**, 244–255 (2011).
7. Yan, L. et al. Single-cell RNA-seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139 (2013)
8. Liu, L. et al. An integrated chromatin accessibility and transcriptome landscape of human pre-implantation embryos. *Nat. Commun.* **10**, 364 (2019).
9. Hendrickson, P. G. et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
10. Madissoon, E. et al. Characterization and target genes of nine human PRD-like homeobox domain genes expressed exclusively in early embryos. *Sci. Rep.* **6**, 28995 (2016).
11. Ralston, A. et al. Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* **137**, 395–403 (2010).
12. Nishioka, N. et al. Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech. Dev.* **125**, 270–283 (2008).
13. Nishioka, N. et al. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev. Cell* **16**, 398–410 (2009).
14. Hirate, Y., Cockburn, K., Rossant, J. & Sasaki, H. Tead4 is constitutively nuclear, while nuclear vs. cytoplasmic Yap distribution is regulated in preimplantation mouse embryos. *Proc. Natl Acad. Sci. USA* **109**, E3389–E3390 (2012).
15. Kilens, S. et al. Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nat. Commun.* **9**, 360 (2018).
16. Lin, K. C., Park, H. W. & Guan, K.-L. Regulation of the Hippo pathway transcription factor TEAD. *Trends Biochem. Sci.* **42**, 862–872 (2017).
17. Soncin, F. et al. Comparative analysis of mouse and human placentae across gestation reveals species-specific regulators of placental development. *Development* **145**, dev156273 (2018).
18. Cauffman, G., De Rycke, M., Sermon, K., Liebaers, I. & Van de Velde, H. Markers that define stemness in ESC are unable to identify the totipotent cells in human preimplantation embryos. *Hum. Reprod.* **24**, 63–70 (2009).
19. Walentin, K. et al. A Grhl2-dependent gene network controls trophoblast branching morphogenesis. *Development* **142**, 1125–1136 (2015).
20. Fang, F. et al. A distinct isoform of ZNF207 controls self-renewal and pluripotency of human embryonic stem cells. *Nat. Commun.* **9**, 4384 (2018).
21. Guo, G. et al. Epigenetic resetting of human pluripotency. *Development* **144**, 2748–2763 (2017).
22. Wicklow, E. et al. HIPPO pathway members restrict SOX2 to the inner cell mass where it promotes ICM fates in the mouse blastocyst. *PLoS Genet.* **10**, e1004618 (2014).
23. Hirate, Y. et al. Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos. *Curr. Biol.* **23**, 1181–1194 (2013).
24. Vinot, S. et al. Asymmetric distribution of PAR proteins in the mouse embryo begins at the 8-cell stage during compaction. *Dev. Biol.* **282**, 307–319 (2005).
25. Hirate, Y. et al. Par-aPKC-dependent and -independent mechanisms cooperatively control cell polarity, Hippo signaling, and cell positioning in 16-cell stage mouse embryos. *Dev. Growth Differ.* **57**, 544–556 (2015).
26. Kjær, S. et al. Adenosine-binding motif mimicry and cellular effects of a thieno[2,3-*d*] pyrimidine-based chemical inhibitor of atypical protein kinase C isoenzymes. *Biochem. J.* **451**, 329–342 (2013).
27. Rodriguez, J. et al. aPKC cycles between functionally distinct PAR protein assemblies to drive cell polarity. *Dev. Cell* **42**, 400–415.e9 (2017).

28. Aguilar-Aragon, M. et al. Pak1 kinase maintains apical membrane identity in epithelia. *Cell Rep*. **22**, 1639–1646 (2018).
29. Plusa, B. et al. Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. *J. Cell Sci*. **118**, 505–515 (2005).
30. Korotkevich, E. et al. The apical domain is required and sufficient for the first lineage segregation in the mouse embryo. *Dev. Cell* **40**, 235–247.e7 (2017).
31. Frum, T., Murphy, T. M. & Ralston, A. HIPPO signaling resolves embryonic cell fate conflicts during establishment of pluripotency in vivo. *eLife* **7**, e42298 (2018).
32. Clift, D. et al. A method for the acute and rapid degradation of endogenous proteins. *Cell* **171**, 1692–1706 (2017).
33. Israel, S., Casser, E., Drexler, H. C. A., Fuellen, G. & Boiani, M. A framework for TRIM21-mediated protein depletion in early mouse embryos: recapitulation of Tead4 null phenotype over three days. *BMC Genomics* **20**, 755 (2019).
34. Chi, F., Sharpley, M. S., Nagaraj, R., Roy, S. S. & Banerjee, U. Glycolysis-independent glucose metabolism distinguishes TE from ICM fate during mammalian embryogenesis. *Dev. Cell* **53**, 9–26.e4 (2020).
35. Alarcon, V. B. Cell polarity regulator PARD6B is essential for trophectoderm formation in the preimplantation mouse embryo. *Biol. Reprod*. **83**, 347–358 (2010).
36. De Paepe, C. et al. Human trophectoderm cells are not yet committed. *Hum. Reprod*. **28**, 740–749 (2013).

Methods

Ethics statement

All human embryo experiments followed all relevant institutional and national guidelines and regulations.

Experiments performed in the UK. This study was approved by the UK Human Fertilisation and Embryology Authority (HFEA): research licence number R0162, and the Health Research Authority's Research Ethics Committee (Cambridge Central reference number 19/EE/0297).

The process of licence approval entailed independent peer review along with consideration by the HFEA Licence and Executive Committees. Our research is compliant with the HFEA code of practice and has undergone inspections by the HFEA since the licence was granted. Research donors were recruited from patients at Bourn Hall Clinic, Homerton University Hospital, The Bridge Centre and IVF Hammersmith.

Informed consent was obtained from 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, an opportunity to receive counselling and the conditions that apply within the licence and the HFEA code of practice. Donors were informed that, in the experiments, embryo development would be stopped before 14 days post-fertilization, and that subsequent biochemical and genetic studies would be performed. Informed consent was also obtained from donors for all the results of these studies to be published in scientific journals. No financial inducements were offered for donation. Consent was not obtained to perform genetic tests on patients and no such tests were performed. The patient information sheets and consent document provided to patients are publicly available (https://www.crick.ac.uk/research/a-z-researchers/researchers-k-o/kathy-niakan/hfea-licence/). Donated embryos surplus to the IVF treatment of the patient were cryopreserved, and were transferred to the Francis Crick Institute where they were thawed and used in the research project.

Experiments performed in France. 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. All human pre-implantation embryos used were obtained from and cultured at the Assisted Reproductive Technology unit of the University Hospital of Nantes, which is 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 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 the opportunity to receive counselling. No financial inducements were 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[37].

Experiments performed in Belgium. The use of human embryos donated to research was allowed by the Local Ethical Committee of UZ Brussel (BUN 143201526417) and the Belgian Federal Committee for research on human embryos (AdV057). The embryos were surplus after IVF treatment in the Centre for Reproductive Medicine at UZ Brussel. The embryos were cryopreserved and donated to research following informed consent and after the legally determined period of five years of cryopreservation.

Human embryo thawing

For embryos thawed in the UK, slow-frozen human cleavage stage embryos were thawed using Quinn's Advantage thaw kit (Origio; ART-8016). In brief, with Quinn's Advantage thaw kit, after thawing the embryos were transferred to 0.5% sucrose thawing medium and incubated for 5 min at 37 °C, followed by 0.2% sucrose thawing medium for 10 min at 37 °C. The embryos were then washed through seven drops of diluent solution before culture. For vitrified cleavage and blastocyst stages, embryos were thawed using vitrification thaw kit (Irvine Scientific; 90137-SO). With vitrification, embryos were thawed in TS thawing solution for 1 min and transferred to DS thawing solution for 4 min at room temperature. Then, embryos were washed twice in washing solution for 4 min at room temperature before culture. Slow-frozen human blastocysts were thawed using a Blast thaw kit (Origio; 10542010A). In brief, upon thawing, blastocysts were incubated in vial 1 thawing medium for 10 min at room temperature in the dark. Then, embryos were moved to vial 2 thawing medium for 10 min at room temperature in the dark before culture. For embryos donated in France, the embryos were thawed as previously described[38]. For embryos donated in Belgium, the embryos were thawed as previously described[39].

Mouse zygote collection

Four- to eight-week-old (C57BL6 × CBA) F_1 female mice were superovulated using injection of 5 IU of pregnant mare serum gonadotrophin (PMSG) (Sigma-Aldrich). Forty-eight hours after PMSG injection, 5 IU of human chorionic gonadotrophin (HCG) (Sigma-Aldrich) was administered. Superovulated female mice were set up for mating with eight-week-old or older (C57BL6 × CBA) F_1 male mice. Mice were maintained on a 12-h light–dark cycle. Mouse zygotes were isolated in FHM under mineral oil (Origio; ART-4008-5P) and cumulus cells were removed with hyaluronidase (Sigma-Aldrich; H4272). This study was approved by the Animal Ethics Committee and by the UK Home Office licence number 70/8560 and followed all relevant institutional and national guidelines and regulations.

Human and mouse embryo culture

Mouse or human embryos were cultured in drops of pre-equilibrated Global medium (LifeGlobal; LGGG-20) supplemented with 5 mg/ml protein supplement (LifeGlobal; LGPS-605) and overlaid with mineral oil (Origio; ART-4008-5P). Pre-implantation embryos were incubated at 37 °C and 5.5% CO_2 in an EmbryoScope+ time-lapse incubator (Vitrolife) and cultured up to the day of analysis. Human embryos donated in France were cultured as previously described[38]. Human embryos donated in Belgium were cultured as previously described[39].

Cow embryo generation and culture

Cow ovaries were obtained from the abattoir. Frozen bull sperm was obtained from Genus ABS UK and UK Sire Services. In vitro-matured oocytes were fertilized as previously described[40], with some modifications. In brief, motile spermatozoa were prepared after 30 min swim-up in calcium-free medium, followed by 10 min centrifugation (300*g*) at room temperature and resuspension of the pellet in fertilization medium (1×10^6 spermatozoa per ml). Cumulus–oocyte complexes (COC) were gently pipetted to remove adhering granulosa cells and break up aggregated COC. Disaggregated COC were then washed once in oocyte wash medium and transferred into 45-µl microdrops of spermatozoa and cultured for 24 h at 38.5 °C in a humidified incubator in an atmosphere of 5% CO_2 and atmospheric O_2. After 24 h, all presumptive zygotes were denuded from cumulus cells and cultured in embryo synthetic oviductal medium (SOF)[41] at 38.5 °C in a humidified incubator in an atmosphere of 5% O_2 and 5% CO_2. The medium was renewed every 2 days up to the day of analysis. This works was approved by the Ethics Committee at the Royal Veterinary College and followed all relevant institutional and national guidelines and regulations.

Morphokinetic analysis and embryo staging

Time-lapse imaging was performed using an EmbryoScope+ (for mouse and human embryos) and an Embryoscope (for cow embryos) time-lapse incubator (Vitrolife). The 8-cell stage was considered to begin when the embryos displayed eight obvious blastomeres. We observed an elongated 8-cell-to-compaction transition with multiple cell divisions in cow and human embryos (Extended Data Fig. 1b–d), which we term 'precompaction'. The beginning of compaction was defined when blastomeres started to flatten and adhere to each other. The morula stage starts when the embryos appear as a compacted group of cells, until the formation of small microlumens. We considered the blastocyst stage to start when embryos showed a single dominant blastocoel cavity. Mouse embryos exhibited a long morula stage, whereas cow and—in particular—human embryos showed a comparatively rapid transition between compaction and cavitation to initiate the formation of a blastocyst (Extended Data Fig. 1b–d, Supplementary Table 1). In our analysis, cavitation was considered from the end of the morula stage until the formation of an expanded blastocyst, when TE cells touch and start to stretch the zona pellucida, causing thinning of the zona pellucida (see Extended Data Fig. 1a for representative images of stages). The hatching blastocyst was defined as when part of the blastocyst was hatched outside of the zona pellucida. In addition, from time-lapse imaging, we manually counted the number of cell divisions that occurred between the 8-cell stage and the beginning of compaction (Extended Data Fig. 1e). We also performed quantification of the number of inner and outer cells at the morula stage, which revealed that cow and mouse embryos have a similar percentage of inner cells whereas human embryos show a notably reduced proportion of inner cells (Extended Data Fig. 1g). Similarly, we observed a difference in the number of ICM cells in mouse and cow compared to human blastocyst stage embryos (Extended Data Fig. 1h).

Inner and outer cell, and ICM and TE cell number, quantification at morula and blastocyst stages

The quantification of inner and outer cells was performed with Imaris 8.2 3D viewer (Bitplane) by counting the number of cells in an inside or outside position in *z*-stack confocal images of morula stage embryos from this publication. The quantification of ICM and TE cell number was obtained by reanalysing data from previous publications of mouse[29], cow[42–45] and human[2] blastocysts.

Immunofluorescence

Owing to the high background signal in the zona pellucida of cow embryos, before fixation the zona pellucida was removed with acidic Tyrode (Sigma) with 1 mg/ml polyvinyl alcohol (Sigma). Embryos were fixed with freshly prepared 4% paraformaldehyde in PBS that was pre-chilled at 4 °C. Embryo fixation was performed for 20 min at room temperature, and then the embryos were transferred through 3 washes of 1× PBS with 0.1% Tween-20 to remove residual paraformaldehyde. Embryos were permeabilized with 1× PBS with 0.5% Triton X-100 and then blocked in blocking solution (3% BSA in 1× PBS with 0.2% Triton X-100) for 2 h at room temperature on a rotating shaker. Then, embryos were incubated with primary antibodies diluted in blocking solution overnight at 4 °C on rotating shaker. The following day, embryos were washed in 1× PBS with 0.2% Triton X-100 for 20 min at room temperature on a rotating shaker, and then incubated with secondary antibodies diluted in blocking solution for 1 h at room temperature on a rotating shaker in the dark. Next, embryos were washed in 1× PBS with 0.2% Triton X-100 for 20 min at room temperature on a rotating shaker. Finally, embryos were placed in 1× PBS with 0.1% Tween-20 with Vectashield and DAPI mounting medium (Vector Lab; H-1200) (1:30 dilution). Phalloidin staining was performed after secondary antibody incubation diluted in blocking solution for 1 h at room temperature on a rotating shaker in the dark. Embryos were placed on µ-Slide 8-well dishes (Ibidi; 80826) for confocal imaging. The antibodies and concentrations used are reported in Supplementary Table 11.

For human embryos donated in France, the embryos were stained and imaged as previously described[38].

For human embryos donated in Belgium, the embryos were stained and imaged as previously described[39], with slight modifications. In brief, the blocking solution was performed with 10% FBS, 1 h at room temperature. Moreover, the primary and secondary antibodies were diluted in 2% FBS and the washing time was extended to 10 min. Nuclei were stained with Hoechst-33342 (ThermoFisher).

Quantification of YAP1 and GATA3 expression

The measurement of YAP1 and GATA3 expression was performed using the 3D ImageJ Suite in the Fiji software. In brief, the nuclei of immunofluorescently stained embryos were segmented automatically on the basis of the DAPI or Hoechst-33342 signal, and YAP1, GATA3, SOX2 and TEAD4 fluorescence intensities were measured within the nuclei. We manually verified that the 3D segmentation did not count the same nucleus twice or fuse together two adjacent nuclei, and in these cases measurements were excluded from the quantification.

Quantification of aPKC and AMOT expression at the apical membrane

The fluorescence intensity profile of aPKC and AMOT at the apical domain in Fig. 2d and Extended Data Fig. 6b, e of outer blastomeres were determined using the plot profile tool in Fiji. For cortical signal intensity of aPKC shown in Extended Data Figs. 10c, 11d, we drew an approximately 1-µm-thick line on the apical domain and measured the mean fluorescence intensity of aPKC. This value was normalized by dividing it by the mean fluorescence intensity measured in the nucleus. We picked confocal slices that cut the cell in the middle, to cover the whole apical domain, as previously performed[27,46,47].

Confocal imaging

For embryos imaged in the UK, confocal immunofluorescence images were taken with Leica SP5 and SP8 confocal microscopes and 1–2-µm-thick optical sections were collected. For embryos imaged in France, images were taken with a Nikon confocal microscope. For embryos imaged in Belgium, images were taken with a LSM800 (Zeiss) confocal microscope.

Re-analysis of single cell RNA-sequencing datasets

We integrated single-cell RNA-sequencing (scRNA-seq) time-course data from two studies[5,7], which investigated the transcriptome at different stages of development of the human pre-implantation embryo. The dataset from ref. [7] (ENA study identifier: PRJNA153427) measures gene expression of individual cells from the oocyte to the late blastocyst stage, and the dataset from ref. [5] (ENA study identifier: PRJEB11202) performed this analysis from the 8-cell to the late blastocyst stage. We also considered the scRNA-seq data from ref. [4] (ENA study identifier: PRJNA277181), which assayed the transcriptome of the three cell lineages of the human blastocyst (epiblast, primitive endoderm and TE).

In brief, reads were mapped to the hg19 reference genome using STAR v.2.7.0e with default settings[48]. Read summarization at the gene level was performed using featureCounts v.1.6.2 of Subread[49]. Then, we applied a previously described quality control and normalization pipeline[50] (using packages readr v.1.3.1, dplyr v.0.8.0.1, SummarizedExperiment v.1.12.0, SingleCellExperiment v.1.4.1, scran v.1.10.2, scater v.1.10.1, cowplot v.0.9.4, ggplot2 v.3.1.1 and ggrepel v.0.8.0) and performed imputation of drop-out events with DrImpute v.1.1[51].

We focused on TE cells with *GATA3* expression above the median in the dataset from ref. [4]. Next, we used package MGFR v.1.8.1[52] to detect genes with high expression in all the selected TE samples compared to the epiblast and primitive endoderm samples. Then, we looked at the expression of the resulting list of genes at the morula stage using the

integrated scRNA-seq data from refs. [5,7] (the datasets were integrated using Seurat 3.1.1). Finally, genes with average expression above 0 at the morula stage were selected to study the Pearson and Spearman correlation of their expression profile against the profile of *GATA3* at the same developmental stage. Functional enrichment analysis of the selected genes was performed with enrichR v.1.0[53].

To identify genes with the opposite trend, we centred our attention on the 10% of the morula samples with the lowest gene expression levels of *GATA3* (red cells in Fig. 1b). Genes with an average expression of at least twofold higher than that of *GATA3* in these samples were shortlisted for correlation analysis against the expression profile of *GATA3* at the morula stage as described above. We identified 1,722 genes that were highly expressed in these cells, of which 166 showed a negative correlation (Pearson's $r < -0.25$) with *GATA3* in all of the human morula cells we analysed.

The data processing and analysis pipelines are publicly available at https://github.com/galanisl/TE_differentiation.

Dimensionality reduction with griph

griph is a tool for automatically identifying clusters of cell types even in the presence of confounding factors for which it is difficult to control (for example, cell-cycle stage or batch effects)[54]. griph uses an ensemble of dissimilarity measures between samples to construct a graph in which only very similar nodes are connected with each other. Then, it applies a community detection algorithm to unsupervisedly spot groups of highly connected nodes that are linked to other groups by only a few edges. These results can be visualized in low-dimensional space using a graph layout algorithm. We applied griph v.0.1.1 to the gene expression matrix of raw counts from the morula stage (integrated dataset from refs. [5,7]) and coloured cells by size-factor-normalized log expression of *GATA3* at this stage. We used the default parameters of the method, except for parameter BatchAssignment, which we set to a character vector indicating the dataset of origin of each cell (that is, from ref. [5] or ref. [7]).

Re-analysis of low-input chromatin accessibility dataset

We re-analysed publicly available ATAC-seq chromatin accessibility datasets at the 8-cell and morula stages of human pre-implantation development (https://www.ebi.ac.uk/ena/browser/view/PRJNA494280) produced by low-input chromatin accessibility and transcriptome sequencing (LiCAT-seq)[8]. For this, we used version 1.1.0 of the nf-core[55] bioinformatics pipeline for ATAC-seq datasets. In brief, adapters and low-quality reads were removed with Trim Galore! v.0.6.5, trimmed FastQ files were mapped to the GRCh38 human reference genome with BWA v.0.7.17 and narrow peaks were called with MACS2 v.2.2.7.1. Tool versions and full details of the pipeline are available at https://nf-co.re/atacseq. For the identification of transcription factor footprints in ATAC-seq peaks, we used the HINT tool in version 0.13.0 of the Regulatory Genomics Toolbox[56]. Then, we used the MotifAnalysis tool from the same toolbox to identify transcription factor motifs reported in JASPAR[57]. Finally, we used HINT again to perform differential transcription-factor activity analysis between the binding sites predicted at the morula versus the 8-cell stage. Transcription factor activity scores measure the strength of binding in a particular biological context and are a function of the difference between the Tn5 transposase cleavage events in the transcription factor footprint and flanking regions, as well as the cleavage events around the predicted binding site[56].

Generation of scRNA-seq consensus coverage tracks

To generate the consensus coverage tracks for the morula scRNA-seq data (Fig. 1i, Extended Data Fig. 3d), we merged the BAM files associated with the high-*GATA3*-expressing group of morula cells (top-right cluster in Fig. 1b) using the merge functionality in SAMtools v.1.3.1[58]. After indexing the merged BAM file, we scaled it with BAMscale v.0.0.5-1[59]. The same operations were applied to the low-*GATA3*-expressing group of morula cells (red cells in Fig. 1b). The gene expression values of *KRT18* and *GRHL2* shown in the figures correspond to the average TPM expression of these genes in each group of morula cells.

Inhibitor treatment

The aPKC inhibitor CRT0276121 (Cancer Research Technology) was dissolved in DMSO to 10 mM stock concentration and diluted at the required concentrations in pre-equilibrated embryo culture medium. The tested concentrations of CRT0276121 are reported in Supplementary Table 6 for mouse embryos, in Supplementary Table 7 for cow embryos and in Supplementary Table 8 for human embryos. The following concentrations were used for the characterization experiments: 7–8 µM for mouse embryos, 1 µM for cow embryos and 1.5 µM for human embryos. Control embryos were developed in pre-equilibrated medium to which the same amount of DMSO was added. Schematics of the treatment timeline for each species are shown in Extended Data Fig. 7a, c, e.

Trim-Away experiment

pSMPP-mCherry-hTRIM21 vector was purchased from Addgene (https://www.addgene.org/104972/). mCherry-hTRIM21 was cloned into pCMV6-XL5 expression vector (https://www.addgene.org/vector-database/5211/) using EcoRI and XmaI restriction sites. To make in vitro-transcribed *mCherry-TRIM21* mRNA, the template DNA was linearized by AgeI-HF digestion. mRNA was generated by in vitro transcription using mMESSAGE mMACHINE T7 Transcription kit (no. AM1344; Thermo Fisher Scientific) and purified using RNeasy MinElute Cleanup Kit (no. 74204; Qiagen). mRNA was aliquoted at a concentration of 5 mg/ml and stored at −80 °C. Mouse anti-aPKC primary antibody (sc-17781, Santa Cruz) used for electroporation was passed through Amicon Ultra-0.5 100-kDa filter devices (Millipore) to concentrate it and to remove traces of azide and replace the buffer with Opti-MEM I medium (no. 31985062; Thermo Fisher Scientific), as previously described[32]. Concentrated antibody was stored in aliquots at −80 °C. Before electroporation, the *mCherry-TRIM21* mRNA and mouse anti-aPKC primary antibody complex was prepared at various concentration of *mCherry-TRIM21* mRNA (200–700 ng/µl) and of mouse anti-aPKC primary antibody (1,320–4,620 ng/µl) diluted in Opti-MEM I medium.

We electroporated embryos using the mouse anti-aPKC primary antibody for Trim-Away aPKC downregulation experiments. We used an anti-mouse IgG secondary antibody to recognize the mouse anti-aPKC primary antibody that was electroporated for the depletion experiment. To specifically detect the endogenous expression of aPKC, we used a rabbit anti-aPKC primary antibody together with an anti-rabbit IgG secondary antibody (schematic in Extended Data Fig. 9i).

Electroporation experiment

Mouse, human or cow embryos were placed in drops of FHM (no. MR-024D; Millipore) or oocyte and embryo wash medium (no. 51002; Egg Technologies), respectively, in a Petri dish overlaid with mineral oil (no. ART-4008-5P; Origio); the dish was placed on a heated microscope stage (Olympus IX70) and a single hole was created in the zona pellucida using a Saturn 5 laser (Research Instruments). The embryos were then washed through Opti-MEM I medium (no. 31985062, Thermo Fisher Scientific). Then, 5–8µl of the *mCherry-TRIM21* mRNA and mouse anti-aPKC primary antibody complex was added between the electrodes of the plate (CUY501P1-1.5, NEPA GENE); the impedance of the solution was checked and adjusted to between 0.19 and 0.21 Ω (typically 0.20 Ω). Electroporation of the embryos was performed using the NEPA21 system (NEPA GENE). For mouse and human embryos, the following electroporation parameters were used: voltage of 20 V, pulse length of 25 ms, pulse interval of 50 ms and the number of pulses was 2. For mouse embryos, various concentrations were tested (Supplementary Table 9), and the concentration used for phenotyping

was 600 ng/µl of *mCherry-TRIM21* mRNA and 3,960 ng/µl of mouse anti-aPKC primary antibody. For human embryos, the concentration used for phenotyping was 700 ng/µl of *mCherry-TRIM21* mRNA and 4,620 ng/µl of mouse anti-aPKC primary antibody. For cow embryos, the following electroporation parameters were used: voltage of 20 V, pulse length of 5 ms or 7.5 ms, pulse interval of 50 ms and the number of pulses was 2. For cow embryos, various concentrations were tested (Supplementary Table 10), and conditions for phenotyping were 400 ng/µl of *mCherry-TRIM21* mRNA and 2,640 ng/µl of mouse anti-aPKC primary antibody at a pulse length of 7.5 ms. Immediately after electroporation, the mouse and human embryos were removed from the plate, washed through FHM and cultured in Global medium (no. LGGG-20; LifeGlobal) supplemented with 5 mg/ml protein supplement (no. GHSA125: LifeGlobal) and overlaid with mineral oil until the morula stage, when they were fixed for phenotyping. In the case of the cow experiments, the embryos were removed from the plate, washed in oocyte and embryo wash medium and cultured in BO-IVC medium (no. 71005; Egg Technologies) until the morula stage, upon which the zona pellucida was removed followed by fixation and phenotyping by immunofluorescence. In the mouse experiments, the control embryos are embryos that were electroporated with *mCherry-TRIM21* mRNA only, whereas in cow and human experiments control embryos are absolute control embryos.

Statistics and reproducibility

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

All statistical analyses in this study were performed using GraphPad Prism 6.0 or using the R environment for statistical computing. The number of cells or embryos analysed (n), statistical tests and P values are all stated in each figure or figure legend. The tests performed in this paper are unpaired two-tailed Student's t-test and two-tailed Mann–Whitney U test. For multiple comparison, FDRs were controlled with the Benjamini–Hochberg method. Data are represented as mean ± s.d. Unless otherwise noted, each experiment was performed at least three times.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The datasets analysed in this study have previously been published, and are available at the Gene Expression Omnibus (GEO) in repositories GSE36552 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE36552) and GSE66507 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66507); at EMBL-EBI ArrayExpress accession number E-MTAB-3929 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-3929/); and at EMBL-EBI European Nucleotide Archive (ENA) accession code PRJNA494280 (https://www.ebi.ac.uk/ena/browser/view/PRJNA494280). Source data are provided with this paper.

Code availability

The data processing and analysis pipelines are publicly available at https://github.com/galanisl/TE_differentiation.

37. Kimmelman, J. et al. New ISSCR guidelines: clinical translation of stem cell research. *Lancet* **387**, 1979–1981 (2016).
38. Meistermann, D. et al. Spatio-temporal analysis of human preimplantation development reveals dynamics of epiblast and trophectoderm. *Cell Stem Cell* https://doi.org/10.2139/ssrn.3441907 (2019).
39. De Paepe, C. et al. BMP4 plays a role in apoptosis during human preimplantation development. *Mol. Reprod. Dev.* **86**, 53–62 (2019).
40. Fouladi Nashta, A. A., Waddington, D. & Campbell, K. H. S. Maintenance of bovine oocytes in meiotic arrest and subsequent development in vitro: a comparative evaluation of antral follicle culture with other methods. *Biol. Reprod.* **59**, 255–262 (1998).
41. Holm, P., Booth, P. J., Schmidt, M. H., Greve, T. & Callesen, H. High bovine blastocyst development in a static in vitro production system using SOFaa medium supplemented with sodium citrate and myo-inositol with or without serum-proteins. *Theriogenology* **52**, 683–700 (1999).
42. Fouladi-Nashta, A. A. et al. Differential staining combined with TUNEL labelling to detect apoptosis in preimplantation bovine embryos. *Reprod. Biomed. Online* **10**, 497–502 (2005).
43. Thouas, G. A., Korfiatis, N. A., French, A. J., Jones, G. M. & Trounson, A. O. Simplified technique for differential staining of inner cell mass and trophectoderm cells of mouse and bovine blastocysts. *Reprod. Biomed. Online* **3**, 25–29 (2001).
44. Goissis, M. D. & Cibelli, J. B. Functional characterization of CDX2 during bovine preimplantation development in vitro. *Mol. Reprod. Dev.* **81**, 962–970 (2014).
45. Koo, D.-B. et al. Aberrant allocations of inner cell mass and trophectoderm cells in bovine nuclear transfer blastocysts. *Biol. Reprod.* **67**, 487–492 (2002).
46. Zhu, M., Leung, C. Y., Shahbazi, M. N. & Zernicka-Goetz, M. Actomyosin polarisation through PLC–PKC triggers symmetry breaking of the mouse embryo. *Nat. Commun.* **8**, 921 (2017).
47. Maître, J.-L., Niwayama, R., Turlier, H., Nédélec, F. & Hiiragi, T. Pulsatile cell-autonomous contractility drives compaction in the mouse embryo. *Nat. Cell Biol.* **17**, 849–855 (2015).
48. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
49. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
50. Theunissen, T. W. et al. Molecular criteria for defining the naive human pluripotent state. *Cell Stem Cell* **19**, 502–515 (2016).
51. Gong, W., Kwak, I. Y., Pota, P., Koyano-Nakagawa, N. & Garry, D. J. DrImpute: imputing dropout events in single cell RNA sequencing data. *BMC Bioinformatics* **19**, 220 (2018).
52. El Amrani, K., Alanis-Lobato, G., Mah, N., Kurtz, A. & Andrade-Navarro, M. A. Detection of condition-specific marker genes from RNA-seq data with MGFR. *PeerJ* **7**, e6970 (2019).
53. Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).
54. Serra, D. et al. Self-organization and symmetry breaking in intestinal organoid development. *Nature* **569**, 66–72 (2019).
55. Ewels, P. A. et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.* **38**, 276–278 (2020).
56. Li, Z. et al. Identification of transcription factor binding sites using ATAC-seq. *Genome Biol.* **20**, 45 (2019).
57. Fornes, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
58. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
59. Pongor, L. S. et al. BAMscale: quantification of next-generation sequencing peaks and generation of scaled coverage tracks. *Epigenetics Chromatin* **13**, 21 (2020).

Acknowledgements We thank the donors whose contributions have enabled this research; P. Patel, A. Srikantharajah, M. Summers, A. Handyside, K. Ahuja, S. Lavery, A. Rattos and M. Jansa Perez for the coordination and donation of embryos to our research project; members of the laboratories of K.K.N., J. M. A. Turner and R. Lovell-Badge, as well as M. Marass, T. Rayon Alonso, B. Thompson and N. Goehring for discussion, advice and feedback on the manuscript; the laboratories of B. Thompson, N. Goehring and J. Briscoe for sharing reagents and advice; the Advanced Light Microscopy and Biological Research Facilities (Francis Crick Institute); and A. Brodie and K. Bacon for assisting with cow embryo culture. The AMOT antibody (Amot-C no. 10061-1) used in this paper for mouse embryos was provided by the laboratory of H. Sasaki. CRT0276121 was provided by Cancer Research Technology. Work in the laboratory of L.D. was funded by a donation from MSD to 'Fondation de l'Université de Nantes'. Work in the laboratory of H.V.d.V. was funded by the Fonds Wetenschappelijk Onderzoek Flanders (FWOAL722) and the Wetenschappelijk Fonds Willy Gepts (WFWG, UZ Brussel, G142). Work in the laboratory of A.A.F.-N. was supported by Comparative Biomedical Sciences Departmental fund from the Royal Veterinary College. Work in the laboratory of K.K.N. was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001120), the UK Medical Research Council (FC001120) and the Wellcome Trust (FC001120), and by the Rosa Beddington Fund.

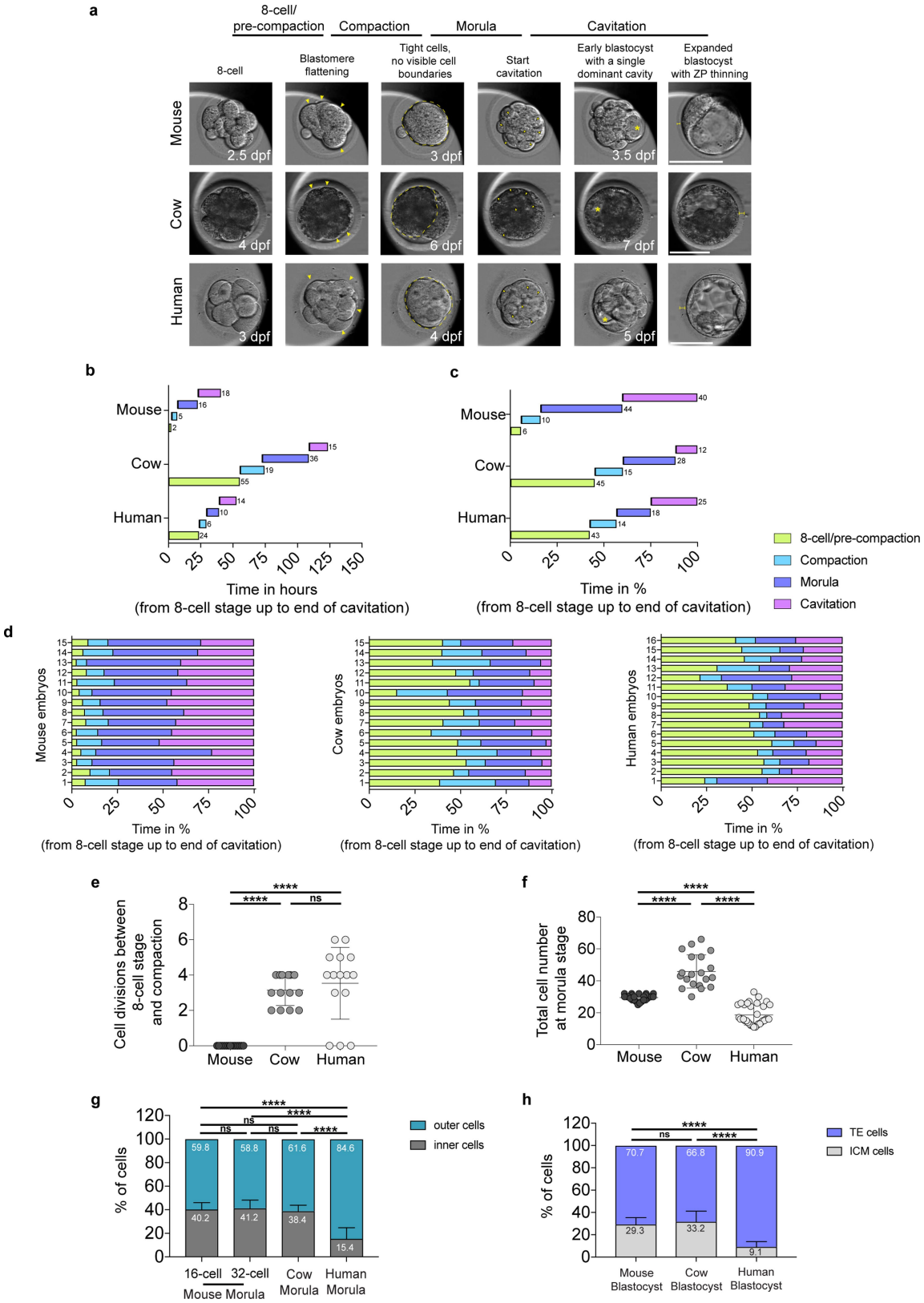
Author contributions C.G. and K.K.N. conceived the study; K.K.N. supervised the project; C.G., K.K.N. and A.M. designed the experiments; C.G., K.K.N., A.M. and N.M.E.F. performed experiments; G.A.-L. performed the bioinformatic analysis of scRNA-seq and ATAC-seq datasets; C.G., K.K.N., A.M. and G.A.-L. analysed data; S.L. managed human embryos donated to research in Nantes; C.G. and A.D. performed experiments on human embryos in Brussels; A.B. and S.L. performed experiments on human embryos in Nantes; K.E., P.S. and L.C. coordinated the donation of embryos to the research project in London; L.D. supervised experiments on human embryos in Nantes; H.V.d.V. supervised experiments on human embryos in Brussels; A.A.F.-N. and D.H. provided cow ovaries; A.A.F.-N. provided techniques for cow embryo generation, helped with the conceptualization and design of experiments on cow embryos, and hosted C.G. in his laboratory; C.G. and K.K.N. wrote the manuscript with help from all of the authors.

Competing interests The authors declare no competing interests.

Additional information

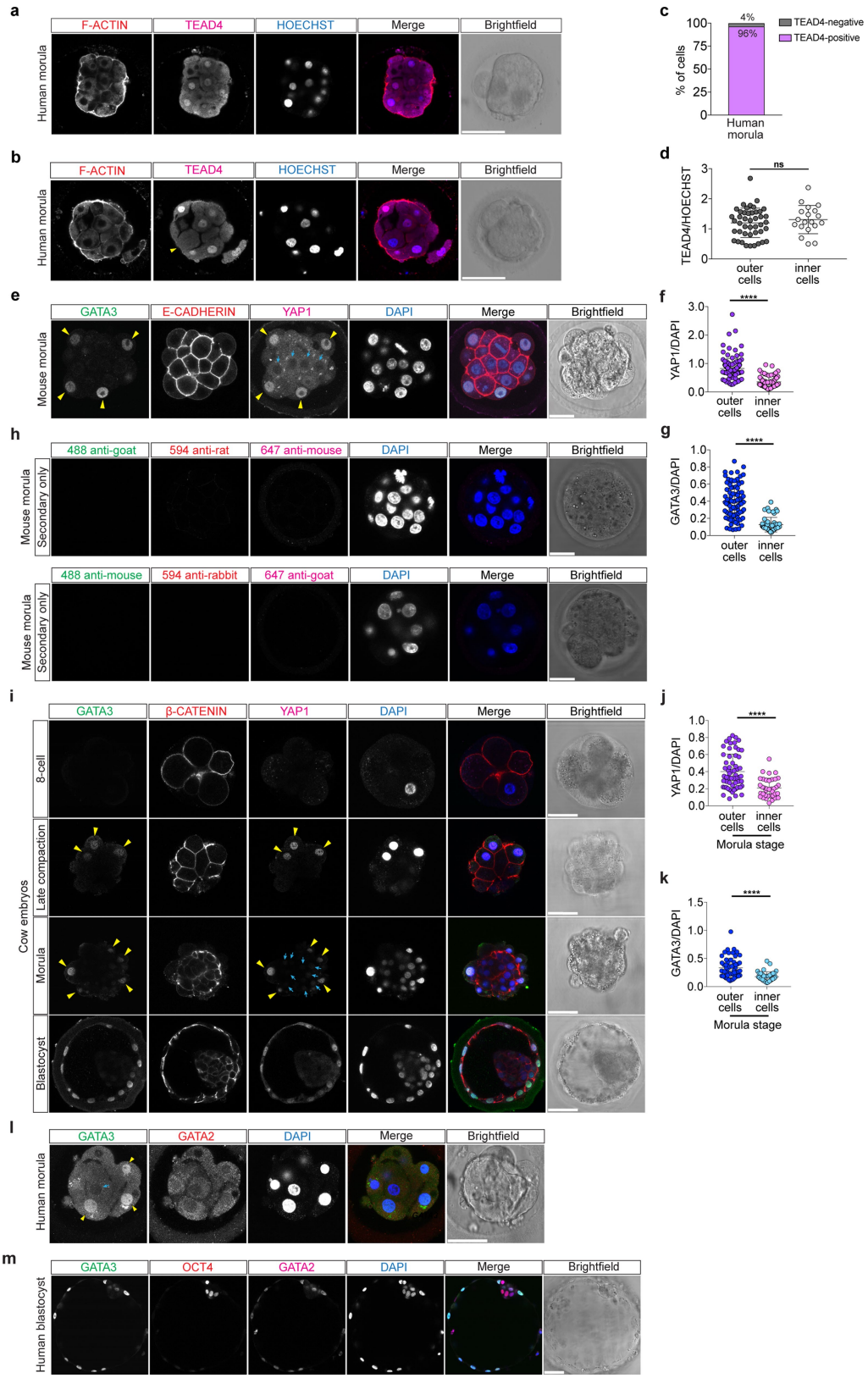
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2759-x.

Correspondence and requests for materials should be addressed to K.K.N.



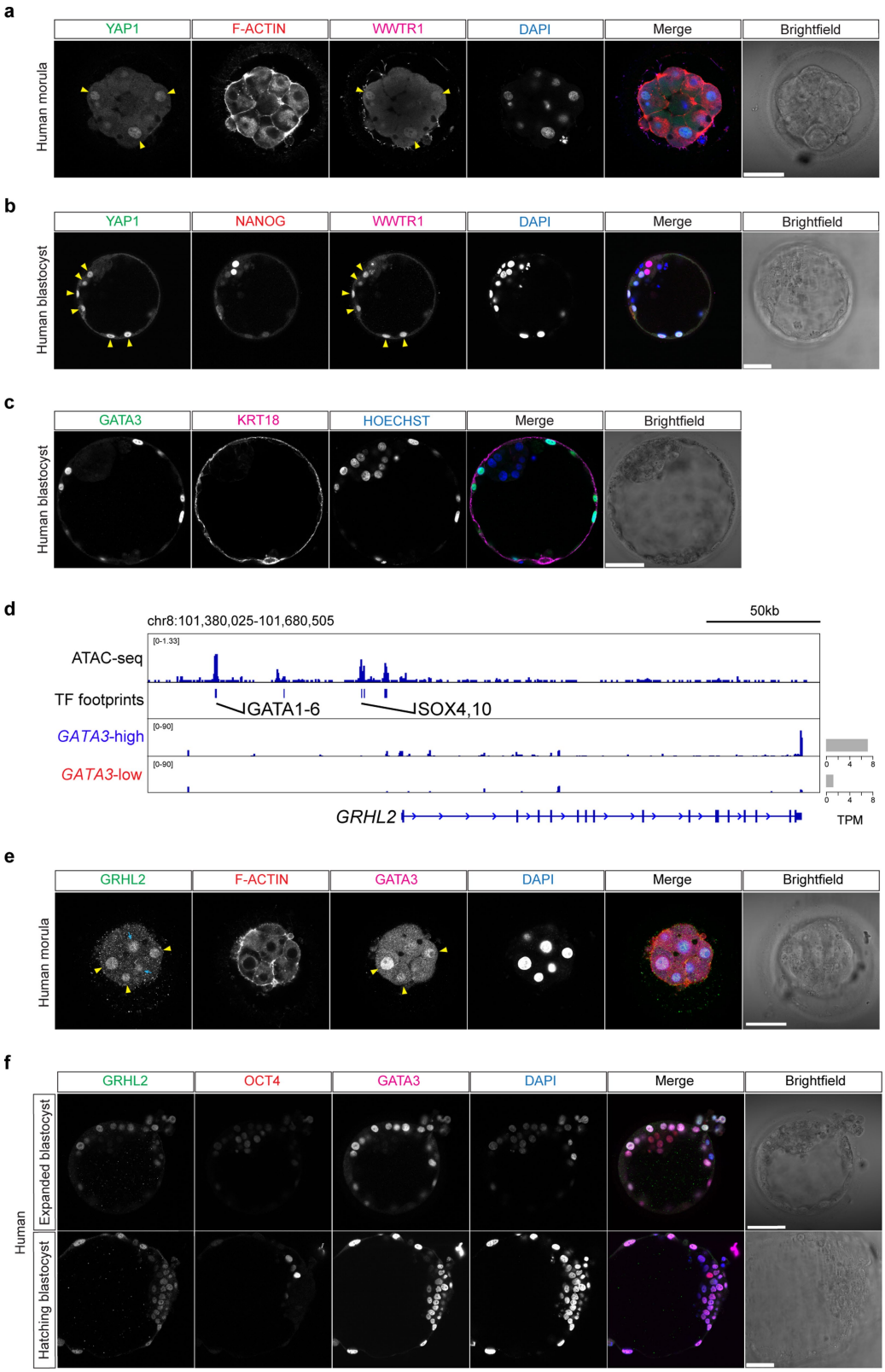
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Morphokinetic analysis to benchmark key morphological events in mouse, cow and human pre-implantation development. a, Representative images of mouse, cow and human embryos at each developmental stage analysed. Yellow arrows point to blastomere flattening. Yellow outline denotes the compacted morula. Yellow arrowheads mark microlumens. Yellow asterisks show single dominant cavity. Yellow bars show zona pellucida (ZP) thinning. dpf, days post-fertilization. Scale bars, 100 µm. **b**, **c**, Morphokinetic analysis of mouse, cow and human pre-implantation development, showing relative time in hours (**b**) and in percentage (**c**) (from 8-cell stage to the end of cavitation). $n = 15$ embryos for mouse and cow, and $n = 16$ for human. **d**, Morphokinetic analysis for each of the mouse, cow and human pre-implantation embryos used in the analysis, showing relative time in percentage (from the 8-cell stage up to the end of cavitation). $n = 15$ embryos for mouse and cow, and $n = 16$ for human. **e**, Quantification of the number of individual cells that divided between 8-cell stage and start of compaction in mouse, cow and human embryos. $n = 15$ embryos for mouse and human, and $n = 14$ for cow. Data are presented as mean ± s.d. Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$, ns, not significant. **f**, Quantification of the total cell number in mouse (32-cell stage), cow and human morula embryos. $n = 25$ for mouse, $n = 20$ for cow and $n = 34$ for human. Data are presented as mean ± s.d. Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. **g**, Quantification of the number of inner and outer cells in mouse, cow and human morula embryos. $n = 25$ for mouse, $n = 20$ for cow and $n = 34$ for human. Data are presented as mean ± s.d. Two-tailed Student's *t*-test, $^{****}P < 0.0001$, ns, not significant. **h**, Quantification of the number of ICM and TE cells in mouse, cow and human blastocysts. $n = 14$ for mouse, $n = 10$ for cow and $n = 12$ for human. Data are presented as mean ± s.d. Two-tailed Student's *t*-test, $^{****}P < 0.0001$, ns, not significant. More details about statistics and reproducibility are provided in Methods.



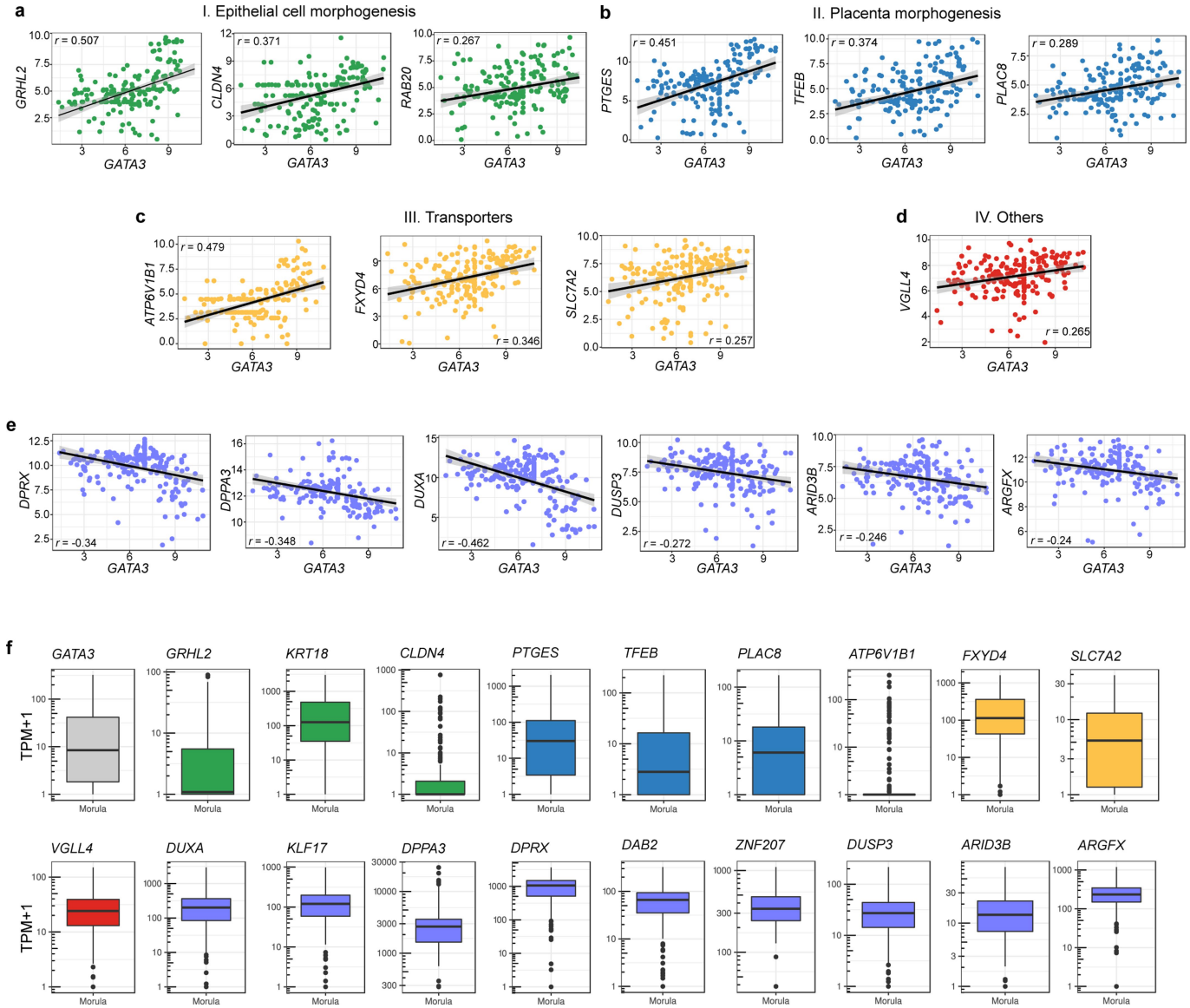
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Protein expression of TE-associated markers in mouse, cow and human morula embryos. a, **b**, Immunofluorescence analysis of F-actin, TEAD4 and Hoechst-33342 nuclear staining in human morula embryo. Yellow arrowheads point to a cell without TEAD4 expression. **c**, Percentage of TEAD4-positive and -negative cells in human morula embryos ($n = 112$ cells from 5 embryos). **d**, Quantification of TEAD4 normalized fluorescence intensity in human morula embryos ($n = 63$ cells from 5 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, ns, not significant. **e**, Immunofluorescence analysis of GATA3, E-cadherin, YAP1 and DAPI nuclear staining in mouse morula embryos. **f**, **g**, Quantification of YAP1 (**f**) and GATA3 (**g**) normalized fluorescence intensity in mouse embryos shown in **e** ($n = 182$ cells for YAP1 from 13 embryos and $n = 139$ cells for GATA3 from 9 embryos). Data are presented as mean ± s.d. Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. **h**, Immunofluorescence analysis using various secondary antibodies and DAPI nuclear staining (blue) in mouse morula embryos ($n = 3$). **i**, Time-course immunofluorescence analysis of GATA3, β-catenin, YAP1 and DAPI nuclear staining in cow embryos at different developmental stages: 8-cell ($n = 5$), late compaction ($n = 5$), morula ($n = 10$) and expanded blastocyst ($n = 5$). **j**, **k**, Quantification of YAP1 (**j**) and GATA3 (**k**) normalized fluorescence intensity in cow morula embryos shown in **i** ($n = 97$ cells from 10 embryos). Data are presented as mean ± s.d. Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. **l**, Immunofluorescence analysis of GATA3, GATA2 and DAPI nuclear staining in a human morula embryo ($n = 3$). **m**, Immunofluorescence analysis of GATA3, OCT4, GATA2 and DAPI nuclear staining in a human expanded blastocyst embryo ($n = 3$). Yellow arrowheads point to outer cells expressing GATA3, and cyan arrow points to an inner cell devoid of GATA3 expression. Scale bars, 25 µm (mouse embryos), 50 µm (cow and human embryos). More details about statistics and reproducibility are provided in Methods.

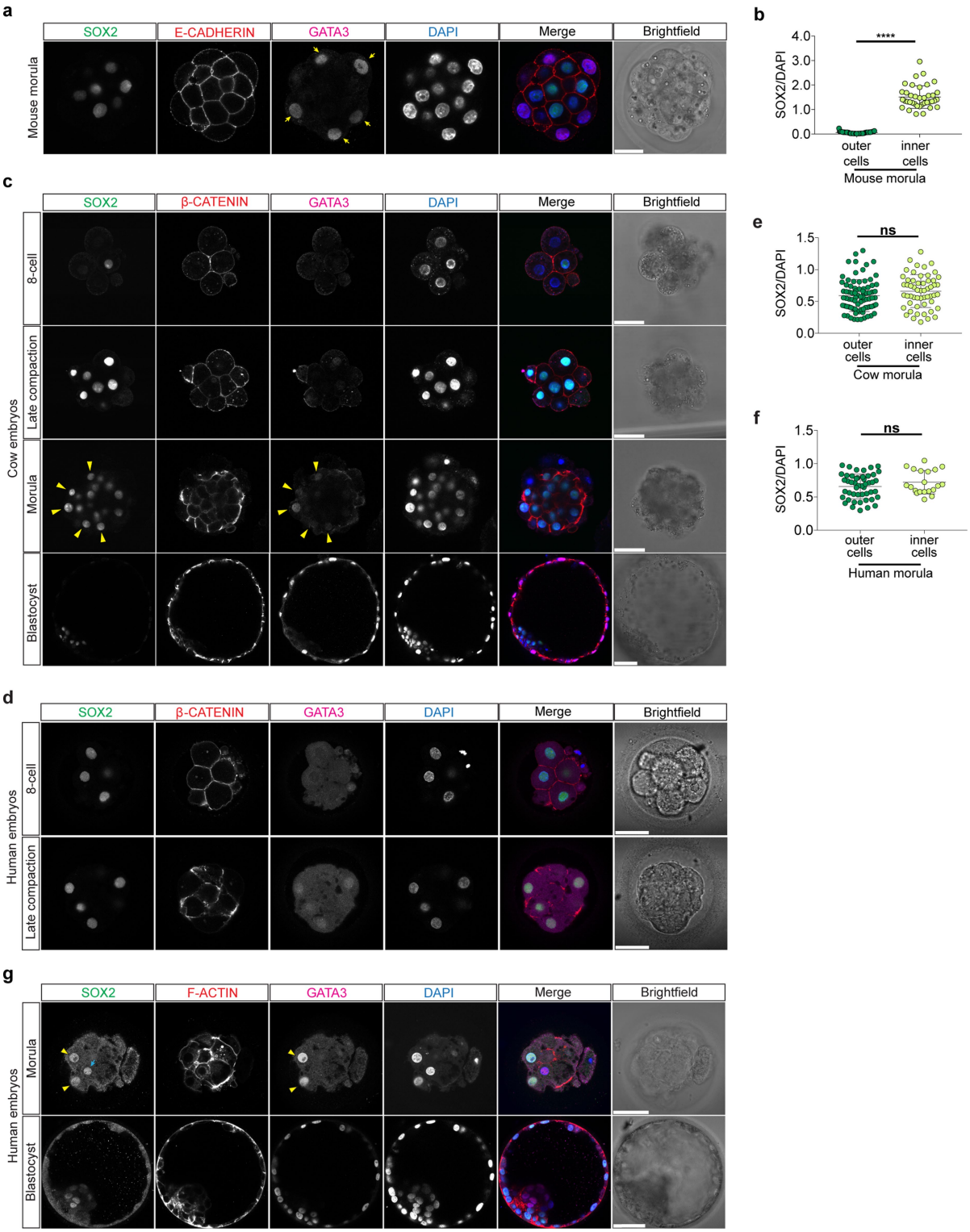


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | WWTR1, KRT18 and GRHL2 expression in human morula embryos. a, Immunofluorescence analysis of YAP1, F-actin, WWTR1 and DAPI nuclear staining in a human morula embryo ($n = 3$). **b**, Immunofluorescence analysis of YAP1, Nanog, WWTR1 and DAPI nuclear staining in a human expanded blastocyst embryo ($n = 3$). Yellow arrowheads point to outer and TE cells co-expressing YAP1 and WWTR1. **c**, Immunofluorescence analysis of GATA3, KRT18 and Hoechst-33342 nuclear staining in a human expanded blastocyst embryo ($n = 3$). **d**, Genome browser view of the ATAC-seq signal at the *GRHL2* locus. High-confidence peaks (FDR-corrected $P < 0.001$; MACS2 dynamic Poisson model) were used to identify transcription factor (TF) motifs. Representative binding motifs associated with the footprints are highlighted. The average expression of *GRHL2* in *GATA3*-high- and *GATA3*-low-expressing cells at the morula is shown and the TPM units are indicated. **e**, Immunofluorescence analysis of GRHL2, F-actin, GATA3 and DAPI nuclear staining in human morula embryos ($n = 3$). Yellow arrowheads point to outer cells co-expressing GRHL2 and GATA3, and cyan arrows point to inner cells expressing GRHL2 only. **f**, Immunofluorescence analysis of GRHL2, OCT4, GATA3 and DAPI nuclear staining in human expanded and hatching blastocyst embryo ($n = 3$ each stage). Scale bars, 50 µm. More details about statistics and reproducibility are provided in Methods.

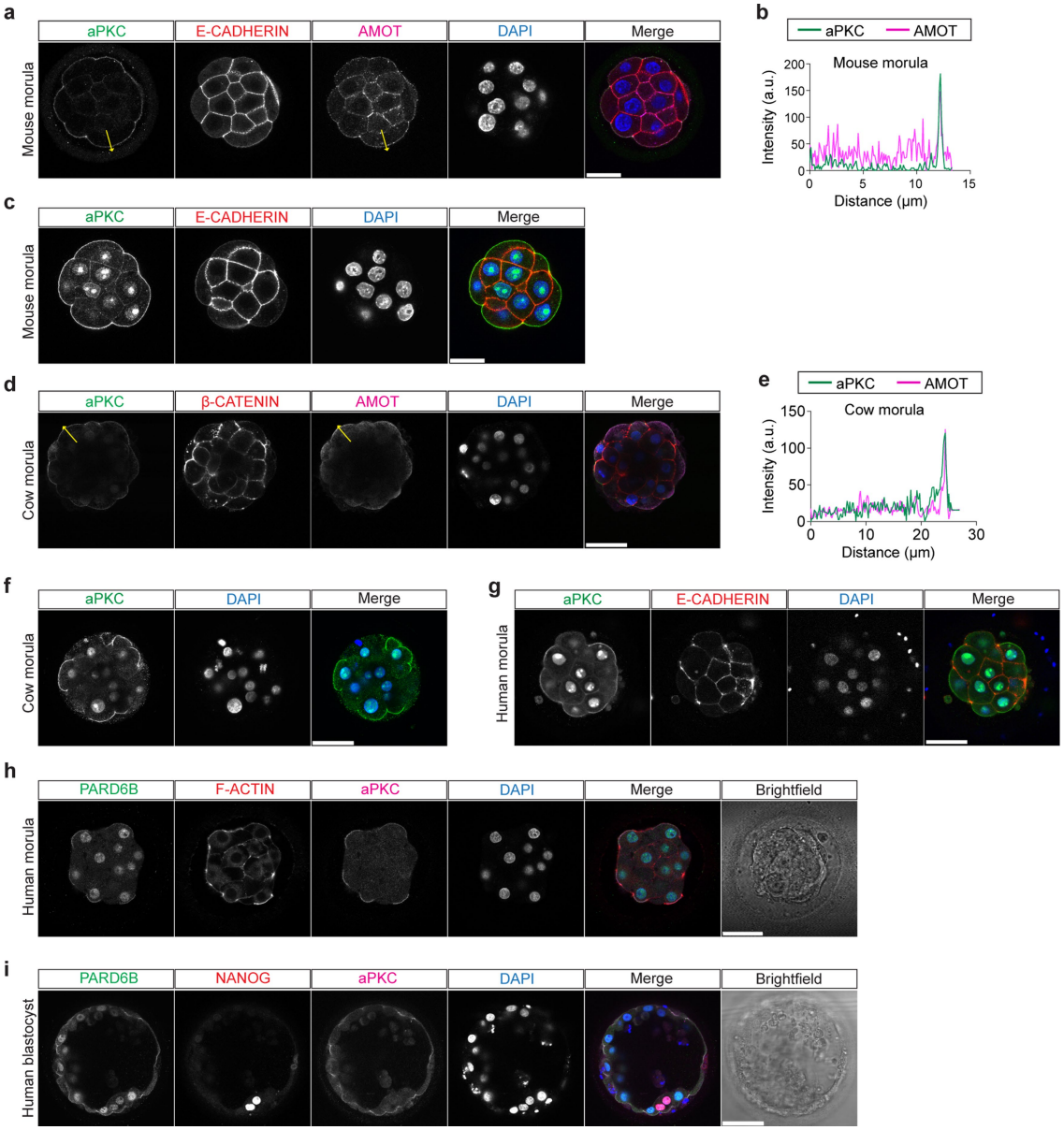


Extended Data Fig. 4 | Correlation analysis of *GATA3* expression in human morula cells. a–**d**, Scatter plots showing positive correlation of *GATA3* expression profile with *GRHL2*, *CLDN4*, *RAB20* (**a**), *PTGES*, *TFEB*, *PLAC8* (**b**), *ATP6V1B1*, *FXYD4*, *SLC7A2* (**c**), and *VGLL4* (**d**) expression profiles in human morula cells. $n = 197$ cells. r, Pearson correlation coefficient. Values are displayed as log-transformed size-factor-normalized counts. The black line corresponds to a linear regression model fitted to the data with 95% confidence bands. **e**, Scatter plots of selected genes implicated in embryonic stem cell pluripotency and/or enriched in precursors of the epiblast and ICM in human morula cells that were identified as negatively correlated with *GATA3* expression. $n = 197$ cells. r, Pearson correlation coefficient. Values are displayed as log-transformed size-factor-normalized counts. The black line corresponds to a linear regression model fitted to the data with 95% confidence bands. **f**, Box plots of selected genes showing positive or negative correlation with *GATA3* expression in human morula cells. $n = 197$ cells. Data are shown as TPM + 1. Boxes correspond to the first and third quartiles, horizontal lines to the median, whiskers extend to 1.5× the interquartile range and dots are outliers. More details about statistics and reproducibility are provided in Methods.

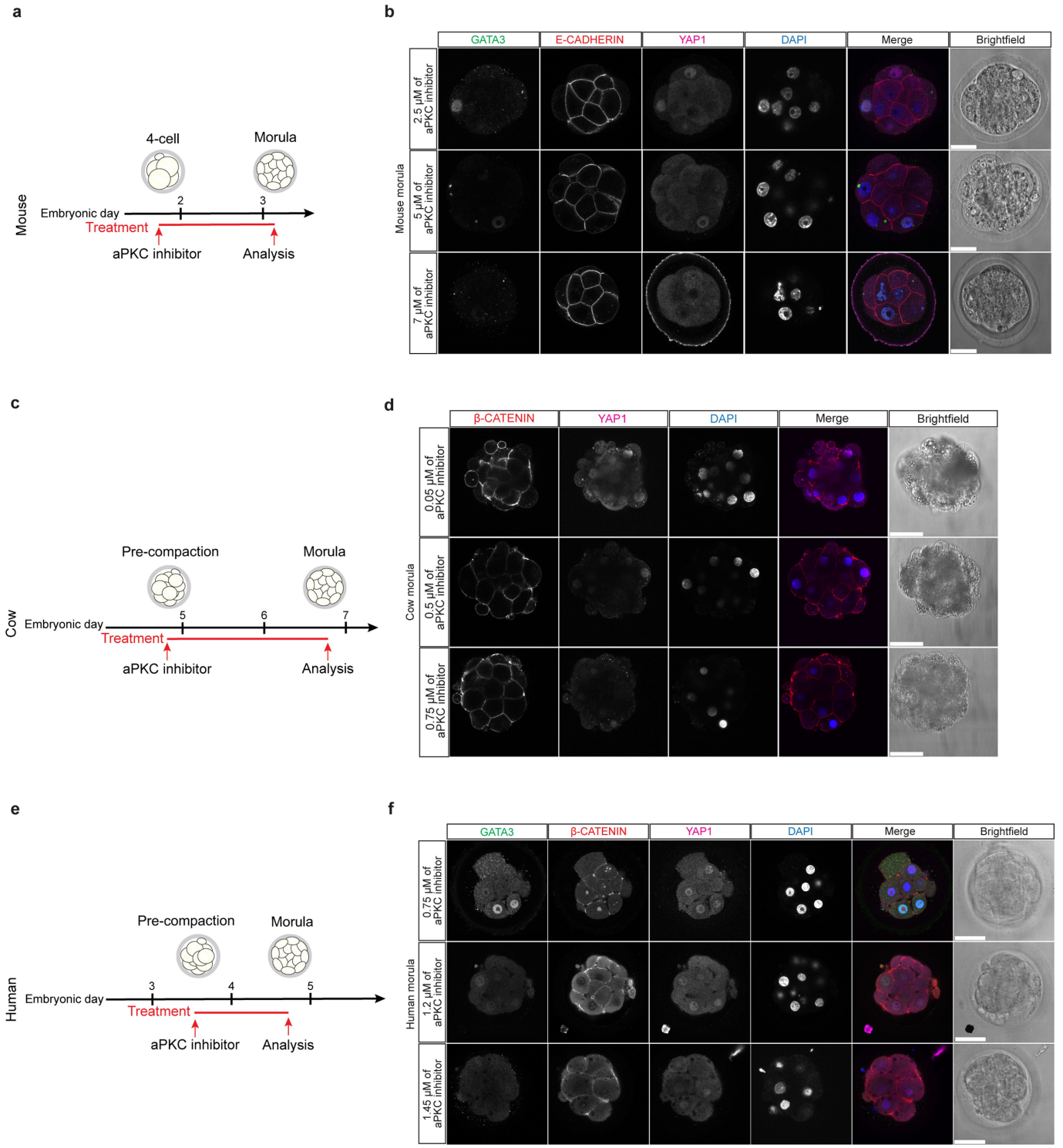


Extended Data Fig. 5 | See next page for caption.

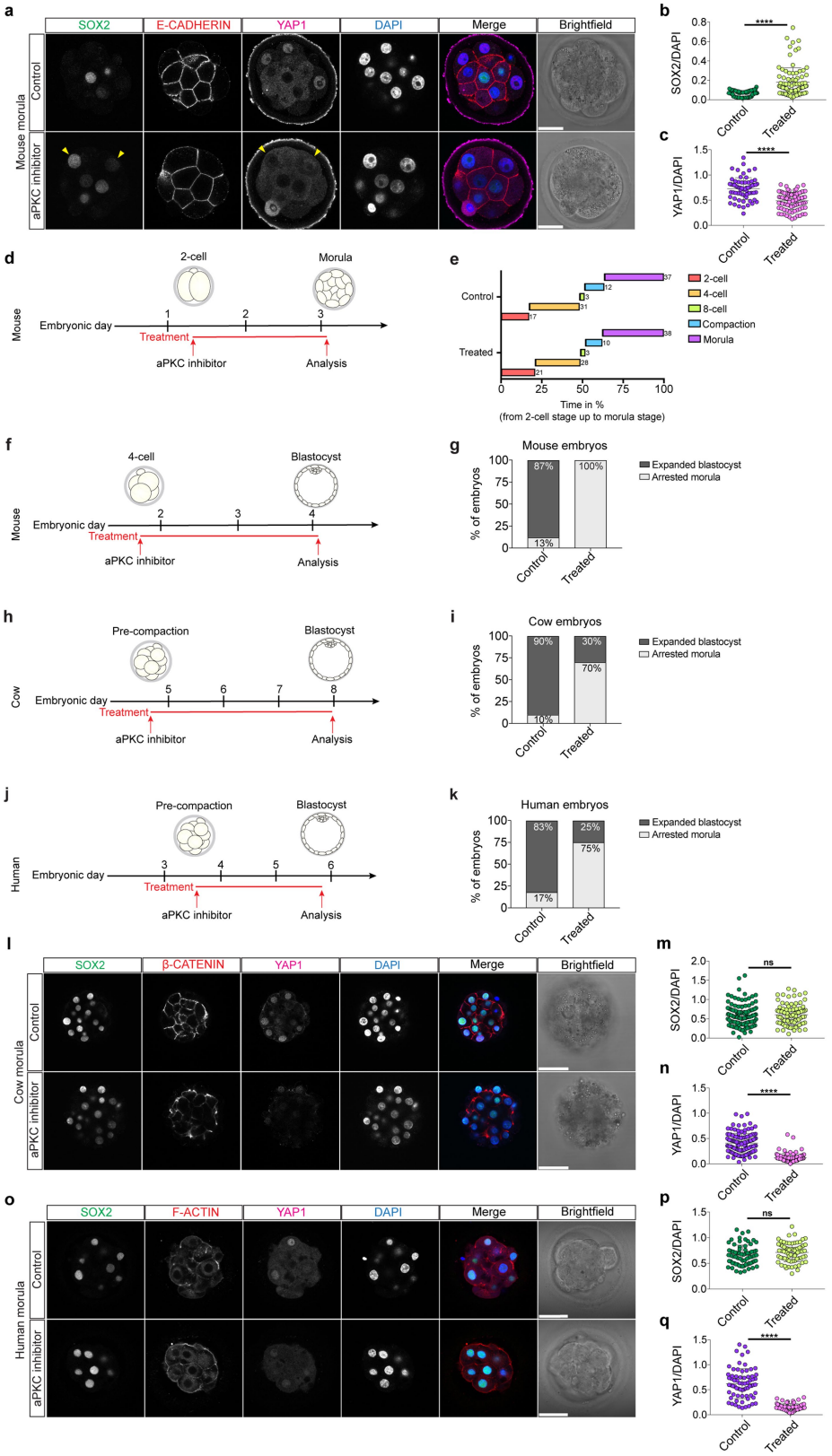
Extended Data Fig. 5 | SOX2 is an inner-cell-specific marker in mouse, but it is broadly expressed in cow and human morula embryos. a, Immunofluorescence analysis of SOX2, E-cadherin, GATA3 and DAPI nuclear staining in mouse morula embryos. **b**, Quantification of SOX2 normalized fluorescence intensity in mouse embryos shown in **a** ($n = 93$ cells from 7 embryos). Data are presented as mean ± s.d. Two-tailed Mann–Whitney *U* test, $^{****}P < 0.0001$. Yellow arrows mark outer cells expressing only GATA3. **c**, Time-course immunofluorescence analysis of SOX2, β-catenin, GATA3 and DAPI nuclear staining in cow embryos at 8-cell ($n = 5$), late compaction ($n = 5$), morula ($n = 10$) and expanded blastocyst ($n = 5$) stages. **d**, Immunofluorescence analysis of SOX2, β-catenin, GATA3 and DAPI nuclear staining in human 8-cell ($n = 5$) and late-compaction ($n = 5$) stage embryos. **e**, Quantification of SOX2 normalized fluorescence intensity in cow morula embryos shown in **c** ($n = 136$ cells from 10 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, ns, not significant. **f**, Quantification of SOX2 normalized fluorescence intensity in human morula embryos shown in Fig. 2b ($n = 68$ cells from 6 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, ns, not significant. **g**, Immunofluorescence analysis of SOX2, F-actin, GATA3 and DAPI nuclear staining in human morula and expanded blastocyst embryo ($n = 3$ each stage). The SOX2 antibody used in **g** in this figure is MAB2018 (R & D), whereas the one used in Fig. 2 is 14-9811-82 (Ebioscience); both show a consistent signal for SOX2. Yellow arrowheads point to outer cells co-expressing GATA3 and SOX2, and cyan arrow points to an inner cell showing SOX2 expression only. Scale bars, 25 µm (mouse embryos), 50 µm (cow and human embryos). More details about statistics and reproducibility are provided in Methods.



Extended Data Fig. 6 | Apical PAR complex and AMOT expression in mouse, cow and human morula embryos. a, Immunofluorescence analysis of aPKC, E-cadherin, AMOT and DAPI nuclear staining in mouse morula embryos ($n = 10$). **b**, Fluorescence intensity profile of aPKC and AMOT shown along the yellow arrows in the mouse morula embryo shown in **a**. a.u., arbitrary units. **c**, Immunofluorescence analysis of aPKC, E-cadherin and DAPI nuclear staining in mouse morula embryo ($n = 10$). **d**, Immunofluorescence analysis of aPKC, β-catenin, AMOT and DAPI nuclear staining in cow morula embryos ($n = 10$). **e**, Fluorescence intensity profile of aPKC and AMOT shown along the yellow arrows in the cow morula embryo shown in **d**. a.u., arbitrary units. **f**, Immunofluorescence analysis of aPKC and DAPI nuclear staining in cow morula embryo ($n = 3$). **g**, Immunofluorescence analysis of aPKC, E-cadherin and DAPI nuclear staining in human morula embryo ($n = 3$). The aPKC antibody used in **c**, **f**, **g** is LC-C354069 (LSBio), whereas the aPKC antibody used in Fig. 2c and in **a**, **d**, **h**, **i** is sc-17781 (Santa Cruz). Both antibodies show consistent and strong aPKC apical expression. **h**, Immunofluorescence analysis of PARD6B, F-actin, aPKC and DAPI nuclear staining in human morula embryo ($n = 3$). **i**, Immunofluorescence analysis of PARD6B, Nanog, aPKC and DAPI nuclear staining in human expanded blastocyst embryo ($n = 3$). Scale bars, 25 μm (mouse embryos), 50 μm (cow and human embryos). More details about statistics and reproducibility are provided in Methods.

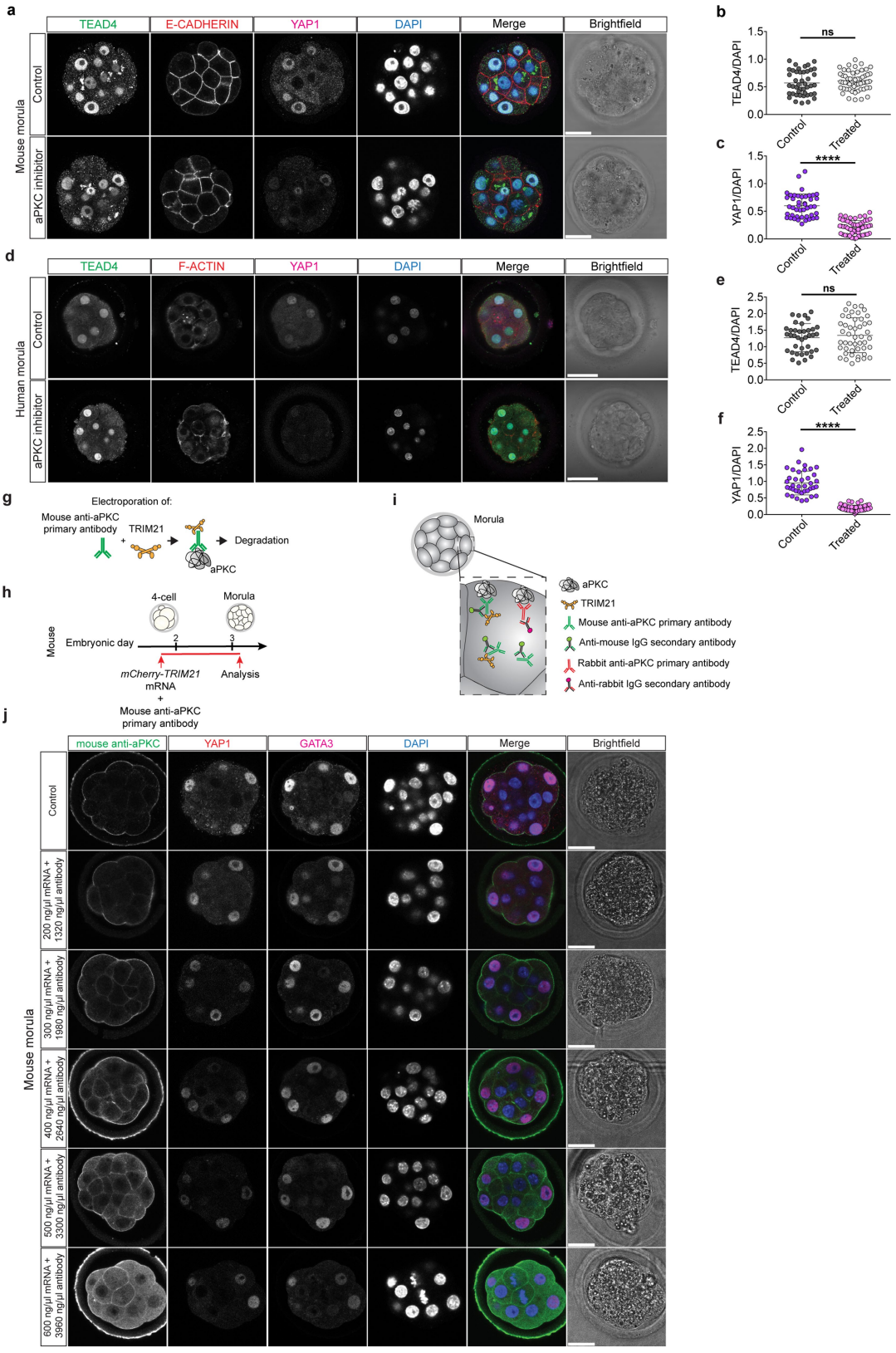


Extended Data Fig. 7 | aPKC inhibitor dose–response experiments in mouse, cow and human embryos. a, Schematic of aPKC inhibitor treatment in mouse embryos. **b**, Immunofluorescence analysis of GATA3, E-cadherin, YAP1 and DAPI nuclear staining in mouse morula embryos treated with different concentrations of aPKC inhibitor (*n* = reported in Supplementary Table 6 for each condition). **c**, Schematic of aPKC inhibitor treatment in cow embryos. **d**, Immunofluorescence analysis of β-catenin, YAP1 and DAPI nuclear staining in cow morula embryos treated with different concentrations of aPKC inhibitor (*n* = reported in Supplementary Table 7 for each condition). **e**, Schematic of aPKC inhibitor treatment in human embryos. **f**, Immunofluorescence analysis of GATA3, β-catenin, YAP1 and DAPI nuclear staining in human morula embryos treated with different concentrations of aPKC inhibitor (*n* = reported in Supplementary Table 8 for each condition). Scale bars, 25 µm (mouse embryos), 50 µm (cow and human embryos). More details about statistics and reproducibility are provided in Methods.



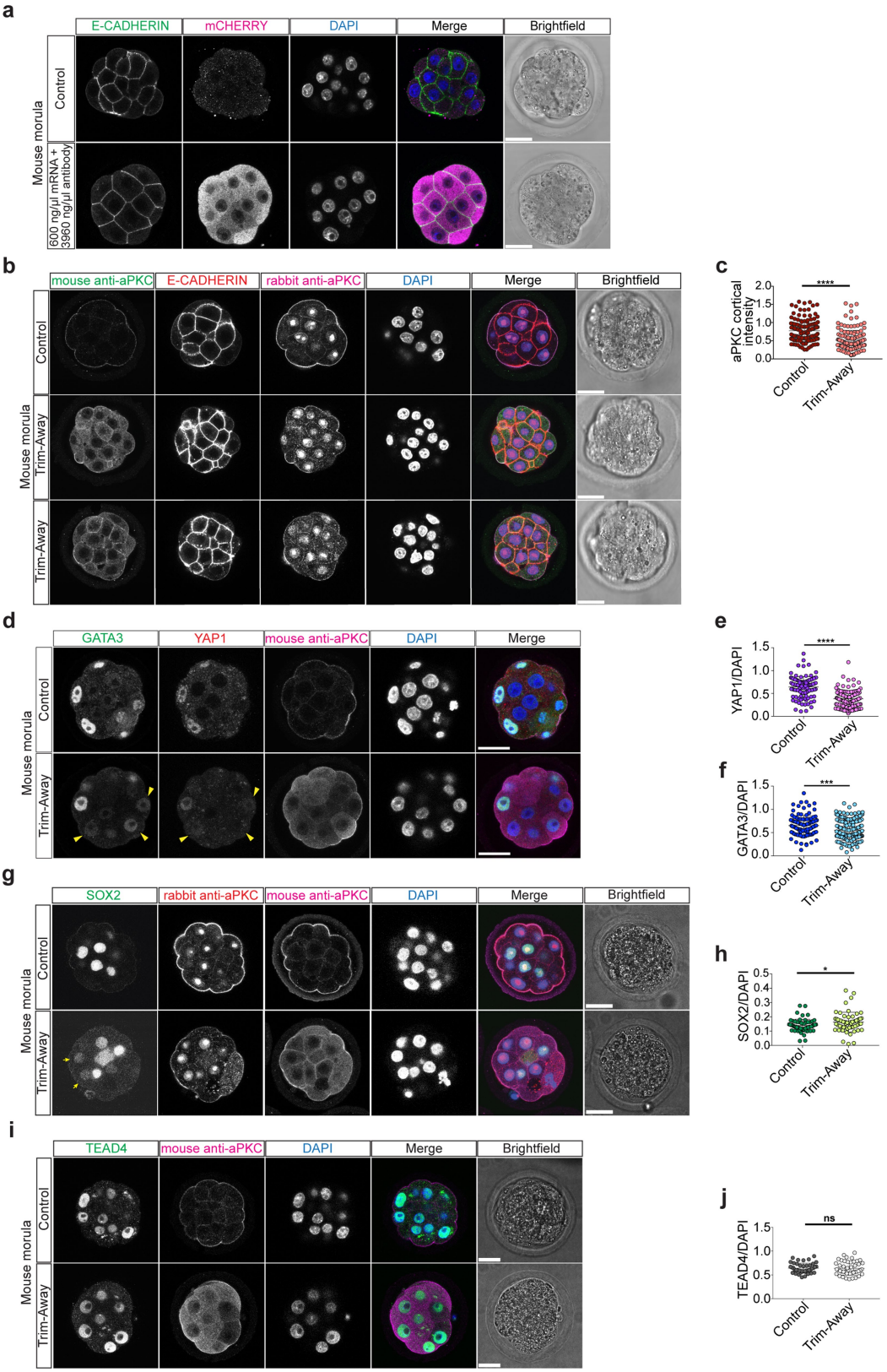
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Characterization of the effects of aPKC inhibition in mouse, cow and human embryos. a, Immunofluorescence analysis of SOX2, E-cadherin, YAP1 and DAPI nuclear staining in control and aPKC-inhibitor-treated mouse morula embryos. Yellow arrowheads point to outer cells expressing SOX2 in mouse morula embryos. **b**, **c**, Quantification of SOX2 (**b**) and YAP1 (**c**) normalized fluorescence intensity in outer cells of mouse embryos shown in **a** ($n = 155$ cells from 19 embryos). Data are presented as mean ± s.d. Two-tailed Mann–Whitney *U* test for SOX2, two-tailed Student's *t*-test for YAP1, $^{****}P < 0.0001$. **d**, Schematic of aPKC inhibitor treatment in mouse embryos. **e**, Morphokinetic analysis of control and aPKC-inhibitor-treated mouse embryos showing relative time in percentage (from 2-cell to morula stage) ($n = 88$). **f**, Schematic of aPKC inhibitor treatment in mouse embryos. **g**, Quantification of percentage of mouse embryos either developing to form an expanded blastocyst or arrested morula in control ($n = 30$) and aPKC inhibitor ($n = 30$) conditions. **h**, Schematic of aPKC inhibitor treatment in cow embryos. **i**, Quantification of percentage of cow embryos either developing to form an expanded blastocyst or arrested morula in control ($n = 10$) and aPKC inhibitor ($n = 10$) conditions. **j**, Schematic of aPKC inhibitor treatment in human embryos. **k**, Quantification of percentage of human embryos either developing to form an expanded blastocyst or arrested morula in control ($n = 12$) and aPKC inhibitor ($n = 12$) conditions. **l**, Immunofluorescence analysis of SOX2, β-catenin, YAP1 and DAPI nuclear staining in control and aPKC-inhibitor-treated cow morula embryos. **m**, **n**, Quantification of SOX2 (**m**) and YAP1 (**n**) normalized fluorescence intensity in outer cells of cow embryos shown in **l** ($n = 218$ cells from 15 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test for SOX2, ns, not significant. Two-tailed Mann–Whitney *U* test for YAP1, $^{****}P < 0.0001$. **o**, Immunofluorescence analysis of SOX2, F-actin, YAP1 and DAPI nuclear staining in control and aPKC-inhibitor-treated human morula embryos. **p**, **q**, Quantification of SOX2 (**p**) and YAP1 (**q**) normalized fluorescence intensity in outer cells of human embryos shown in **o** ($n = 140$ cells from 10 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test for SOX2, ns, not significant. Two-tailed Mann–Whitney *U* test for YAP1, $^{****}P < 0.0001$. Scale bars, 25 µm (mouse embryos), 50 µm (cow and human embryos). More details about statistics and reproducibility are provided in Methods.



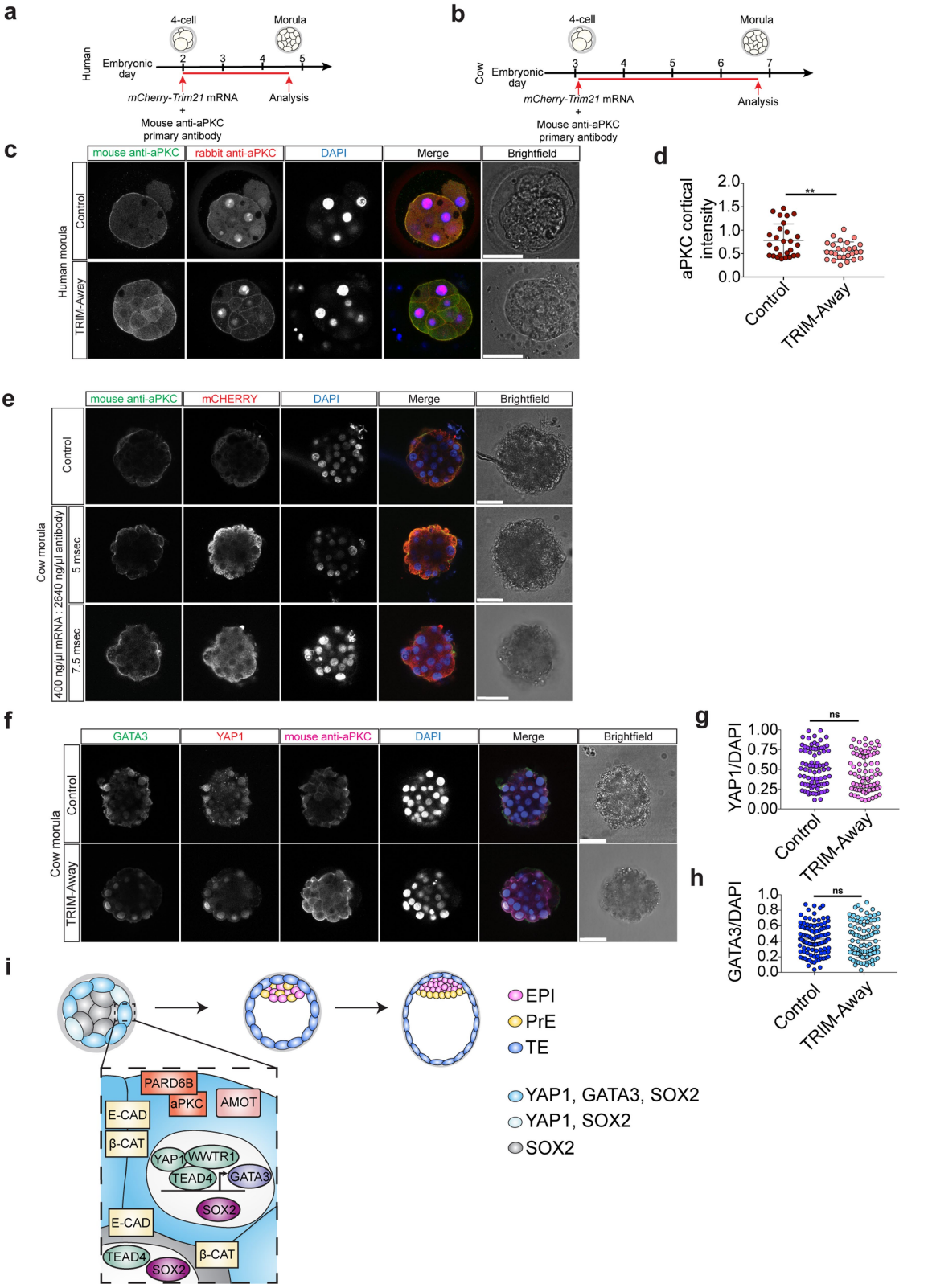
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Trim-Away optimization in mouse embryos. **a**, Immunofluorescence analysis of TEAD4, E-cadherin, YAP1 and DAPI nuclear staining in control and aPKC-inhibitor-treated mouse morula embryos. **b**, **c**, Quantification of TEAD4 (**b**) and YAP1 (**c**) normalized fluorescence intensity in outer cells of mouse embryos shown in **a** ($n = 101$ cells from 10 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test for TEAD4, ns, not significant. Two-tailed Mann–Whitney *U* test for YAP1, $^{****}P < 0.0001$. **d**, Immunofluorescence analysis of TEAD4, F-actin, YAP1 and DAPI nuclear staining in control and aPKC-inhibitor-treated human morula embryos. **e**, **f**, Quantification of TEAD4 (**e**) and YAP1 (**f**) normalized fluorescence intensity in outer cells of human embryos shown in **d** ($n = 85$ cells from 6 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test for TEAD4, ns, not significant. Two-tailed Mann–Whitney *U* test for YAP1, $^{****}P < 0.0001$. **g**, Schematic of the Trim-Away approach. **h**, Schematic of the Trim-Away experiment in mouse. **i**, Schematic of the two anti-aPKC primary antibodies used in the Trim-Away experiment. **j**, Immunofluorescence analysis of anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody, YAP1, GATA3 and DAPI nuclear staining in mouse morula embryos in control and Trim-Away conditions at different concentrations (n = reported in Supplementary Table 9 for each condition). Scale bars, 25 µm (mouse embryos), 50 µm (human embryos). More details about statistics and reproducibility are provided in Methods.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Characterization of the effects of aPKC Trim-Away in mouse embryos. a, Immunofluorescence analysis of E-cadherin, mCherry and DAPI nuclear staining in mouse morula embryos in control and Trim-Away conditions ($n = 10$). **b**, Immunofluorescence analysis of anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody, E-cadherin, rabbit anti-aPKC primary antibody to detect the aPKC protein and DAPI nuclear staining in mouse morula embryos in control and Trim-Away conditions. **c**, Quantification of aPKC cortical intensity (based on rabbit anti-aPKC signal) in outer cells at the morula stage of mouse embryos shown in **b** ($n = 270$ cells from 30 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, $^{****}P < 0.0001$. **d**, Immunofluorescence analysis of GATA3, YAP1, anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody and DAPI nuclear staining in mouse morula embryos in control and Trim-Away conditions. Yellow arrowheads point to decreased YAP1 and GATA3 expression. **e**, **f**, Quantification of YAP1 (**e**) and GATA3 (**f**) normalized fluorescence intensity in outer cells of mouse embryos shown in **d** ($n = 281$ cells for YAP1 from 32 embryos, $n = 263$ cells for GATA3 from 31 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, $^{***}P = 0.00078$ for GATA3 distribution, $^{****}P < 0.0001$ for YAP1 distribution. **g**, Immunofluorescence analysis of SOX2, rabbit anti-aPKC primary antibody to detect the aPKC protein, anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody and DAPI nuclear staining in mouse morula embryos in control and Trim-Away conditions. Yellow arrows point to outer cells expressing SOX2 in a mouse morula embryo. $n = 2$ independent experiments. **h**, Quantification of SOX2 normalized fluorescence intensity in outer cells of mouse embryos shown in **g** ($n = 120$ cells 18 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, $^{*}P = 0.026$. **i**, Immunofluorescence analysis of TEAD4, anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody and DAPI nuclear staining in mouse morula embryos in control and Trim-Away conditions. $n = 2$ independent experiments. **j**, Quantification of TEAD4 normalized fluorescence intensity in outer cells of mouse embryos shown in **i** ($n = 90$ cells from 10 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, ns, not significant. Scale bars, 25 µm. More details about statistics and reproducibility are provided in Methods.



Extended Data Fig. 11 | See next page for caption.

Extended Data Fig. 11 | Trim-Away experiments in cow and human embryos. **a**, **b**, Schematic of the Trim-Away experiments. **c**, Immunofluorescence analysis of anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody, rabbit anti-aPKC primary antibody to detect the aPKC protein, and DAPI nuclear staining in human morula embryos in control and Trim-Away conditions. **d**, Quantification of aPKC cortical intensity (based on rabbit anti-aPKC signal) in outer cells of human embryos shown in **c** ($n = 54$ cells from 6 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, $^{**}P = 0.006$. **e**, Immunofluorescence analysis of anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody, mCherry, and DAPI nuclear staining in cow morula embryos in control and Trim-Away conditions at a pulse length of 5 or 7.5 ms (n = reported in Supplementary Table 10 for each condition). **f**, Immunofluorescence analysis of GATA3, YAP1, anti-mouse IgG secondary antibody to detect the electroporated mouse anti-aPKC primary antibody and DAPI nuclear staining in cow morula embryos in control and Trim-Away conditions. **g**, **h**, Quantification of YAP1 (**g**) and GATA3 (**h**) normalized fluorescence intensity in outer cells of cow embryos shown in **f** ($n = 154$ cells for YAP1 from 16 embryos, and $n = 207$ cells for GATA3 from 20 embryos). Data are presented as mean ± s.d. Two-tailed Student's *t*-test, ns, not significant. Scale bars, 50 µm. **i**, Proposed model for human early lineage specification. EPI, epiblast; PrE, primitive endoderm; E-CAD, E-cadherin; β-CAT, β-catenin. More details about statistics and reproducibility are provided in Methods.

nature research

Corresponding author(s): Kathy Niakan

Last updated by author(s): August 9, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	scRNA-seq data (FASTQ files) were downloaded from the ENA (study IDs PRJNA153427, PRJEB11202 and PRJNA277181), mapped to the hg19 reference genome with STAR v2.7.0e and summarised to the gene level with Subread's featureCounts v1.6.2. Pre-processing and normalisation was carried out with the following packages under version 3.5.2 of the R environment for statistical computing and graphics: readr v1.3.1, dplyr v0.8.0.1, SummarizedExperiment v1.12.0, SingleCellExperiment v1.4.1, scran v1.10.2, scater v1.10.1 and DrImpute v1.1. For more details see https://github.com/galanisl/TE_differentiation. Chromatin accessibility profiles were downloaded from the ENA (study ID PRJNA494280) and pre-processed with version 1.1.0 of the nf-core pipeline for ATAC-seq data (https://nf-co.re/atacseq). Briefly, adapters and low quality reads were removed with Trim Galore! v0.6.5, trimmed FASTQ files were mapped to the GRCh38 human genome with BWA v0.7.17 and narrow peaks were called with MACS2 v2.2.7.1.
Data analysis	Analysis of the pre-processed and normalised scRNA-seq data was performed with the following packages under version 3.5.2 of R: SummarizedExperiment v1.12.0, SingleCellExperiment v1.4.1, scran v1.10.2, scater v1.10.1, MGFR v1.8.1, dplyr v0.8.0.1, readr v1.3.1, cowplot v0.9.4, ggplot2 v3.1.1, ggrepel v0.8.0, DrImpute v1.1, Seurat 3.1.1, griph v0.1.1 and enrichR v1.0. For more details see https://github.com/galanisl/TE_differentiation. Analysis of the pre-processed chromatin accessibility data was performed with the HINT and MotifAnalysis tools in version 0.13.0 of the Regulatory Genomics Toolbox. Genome browser tracks for scRNA-seq data were generated with SAMtools v1.3.1 and BAMscale v0.0.5-1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets analyzed in this study were previously published and are available at the GEO repositories GSE36552 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE36552) and GSE66507 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66507), at EMBL-EBI ArrayExpress: E-MTAB-3929 (https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-3929/) and at EMBL-EBI ENA: PRJNA494280 (https://www.ebi.ac.uk/ena/browser/view/PRJNA494280). Source data behind Figures 1-3 and Extended Data Figures 1, 2, 5, 6, 8-11 are available within the manuscript files.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determined based on our previous experience and experience from other groups' work. For human embryos: Fogarty et al., Nature, 2017 Blakeley et al., Development, 2015 Niakan et al., Developmental Biology, 2013 For mouse embryos: Hirate et al., Current Biology, 2013 Frum et al., Elife, 2018 For cow embryos: Berg et al., Developmental Cell, 2011 Fouladi-Nashta et al., Reproductive Biomedicine Online, 2005
Data exclusions	No data were excluded from the study design.
Replication	Experiment were replicated at least three times and the data were reproducible in the different attempts of replication. GATA3 and YAP1 immunofluorescence analysis was independently performed in three different laboratories: Niakan, David and Van de Velde Labs.
Randomization	Samples were allocated randomly into experimental groups
Blinding	The investigators were not blinded to group allocation during experiments and outcome assessment due to the experimental design in which embryos had to undergo specific treatments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

List of antibodies used available below and in Extended Data Table 11.
Anti-GATA3 antibody, catalogue number: AF2605, supplier: R&D, dilution: 1:200
Anti-GATA3 antibody, catalogue number: AB199428, supplier: Abcam, dilution: 2 µg/ml
Anti-E-CADHERIN antibody, catalogue number: 13-1900, supplier: ThermoFischer, dilution: 1:400 (mouse), 1:50 (human, cow)
Anti-YAP1 antibody, catalogue number: H00010413-M01, supplier: Abnova, dilution: 1:50
Anti-YAP1 antibody, catalogue number: 14074, supplier: Cell Signaling, dilution: 1:50
Anti-aPKC antibody, catalogue number: sc-17781, supplier: Santa Cruz, dilution: 1:50
Anti-aPKC antibody, catalogue number: LC-C354069, supplier: LSBio, dilution: 1:50
Anti-SOX2 antibody, catalogue number: 14-9811-82, supplier: eBioscience, dilution: 1:100
Anti-SOX2 antibody, catalogue number: MAB2018, supplier: R&D, dilution: 1:100
Anti-β-CATENIN antibody, catalogue number: 8814, supplier Cell Signaling, dilution: 1:100
Anti-AMOT antibody, catalogue number: sc-82493, supplier: Santa Cruz, dilution: 1:50 (cow, human)
Anti-AMOT antibody, specific for mouse, AMOT, gently provided by Hiroshi Sasaki's Lab, dilution: 1:200 (mouse)
Anti-KRT18 antibody, catalogue number: ab668, supplier: Abcam, dilution: 1:100
Anti-GATA2 antibody, catalogue number: sc-9008, supplier: Santa Cruz, dilution: 1:100
Anti-PARD6B antibody, catalogue number: BP1-87337, supplier: Novus Biologicals, dilution: 1:50
Anti-WWTR1, catalogue number: HPA007415, supplier: Sigma, dilution: 1:100
Anti-NANOG antibody, catalogue number: AF1997, supplier: R&D, dilution: 1:100
Anti-OCT3/4 antibody, catalogue number: sc-5279, supplier: Santa Cruz, dilution: 1:250
Anti-TEAD4 antibody, catalogue number: ab58310, supplier: Abcam, dilution: 1 µg/ml
Anti-GRHL2 antibody, catalogue number: HPA004820, supplier: Sigma, dilution: 1:200
Anti-mCHERRY antibody, catalogue number: ab167453, supplier: Abcam, dilution: 1:200
Alexa Fluor 594 Phalloidin, catalogue number A12381, supplier: ThermoFischer, dilution 1:250
Alexa Fluor 546 Phalloidin, catalogue number: A22283, supplier: ThermoFischer, dilution: 1:250
Also, Alexa Fluor secondary antibodies from ThermoFisher, dilution: 1:300

Validation

Here we report the list of antibodies along with manufacturer's website link where validation was reported:
Anti-GATA3 antibody, AF2605, R&D = https://www.rndsystems.com/products/human-gata-3-antibody_af2605
Anti-GATA3 antibody, AB199428, Abcam = https://www.abcam.com/gata3-antibody-epr16651-chip-grade-ab199428.html
Anti-E-CADHERIN antibody, 13-1900, Life Technologies = https://www.thermofisher.com/antibody/product/13-1900.html?CID=AFLCA-13-1900
Anti-YAP1 antibody, H00010413-M01, Abnova = http://www.abnova.com/products/products_detail.asp?catalog_id=H00010413-M01
Anti-YAP1 antibody, 14074, Cell Signaling = https://www.cellsignal.com/products/primary-antibodies/yap-d8h1x-xp-rabbit-mab/14074
Anti-aPKC antibody, sc-17781, Santa Cruz = https://www.scbt.com/scbt/product/pkc-zeta-antibody-h-1
Anti-aPKC, LC-C354069, LSBio = https://www.lsbio.com/antibodies/prkcz-antibody-pkc-zeta-antibody-c-terminus-icc-if-immunofluorescence-ihc-ip-wb-western-ls-c354069/365190
Anti-SOX2 antibody, 14-9811-82, Ebioscience = https://www.thermofisher.com/antibody/product/SOX2-Antibody-clone-Btjce-Monoclonal/14-9811-82
Anti-SOX2 antibody, MAB2018, R&D = https://www.rndsystems.com/products/human-mouse-rat-sox2-antibody-245610_mab2018
Anti-β-CATENIN antibody, 8814, Cell Signalling = https://www.cellsignal.com/products/primary-antibodies/non-phospho-active-b-catenin-ser33-37-thr41-d13a1-rabbit-mab/8814
Anti-AMOT antibody, provided by Hiroshi Sasaki's Lab = https://pubmed.ncbi.nlm.nih.gov/23791731/
Anti-AMOT antibody, sc-82493, Santa Cruz = http://datasheets.scbt.com/sc-82493.pdf
Anti-KRT18 antibody, ab668, Abcam = https://www.abcam.com/cytokeratin-18-antibody-c-04-ab668.html
Anti-GATA2 antibody, sc-9008, Santa Cruz = https://www.scbt.com/scbt/product/gata-2-antibody-h-116
Anti-PARD6B antibody, NBP1-87337, Novus Biologicals = https://www.novusbio.com/products/pard6b-antibody_nbp1-87337
Anti-WWTR1 antibody, HPA007415, Sigma = https://www.sigmaaldrich.com/catalog/product/sigma/hpa007415?lang=en&region=GB
Anti-NANOG antibody, AF1997, R&D = https://www.rndsystems.com/products/human-nanog-antibody_af1997
Anti-OCT3/4 antibody, sc-5279, Santa Cruz = https://www.scbt.com/scbt/product/oct-3-4-antibody-c-10
Anti-TEAD4 antibody, ab58310, Abcam = https://www.abcam.com/tead4-antibody-chip-grade-ab58310.html
Anti-GRHL2 antibody, HPA004820, Sigma-Aldrich = https://www.sigmaaldrich.com/catalog/product/sigma/hpa004820?lang=en&region=GB
Anti-mCherry antibody, ab167453, Abcam = https://www.abcam.com/mcherry-antibody-ab167453.html
Alexa Fluor 594 Phalloidin, A12381, ThermoFischer = https://www.thermofisher.com/order/catalog/product/A12381?SID=srch-srp-A12381
Alexa Fluor 546 Phalloidin, A22283, ThermoFischer = https://www.thermofisher.com/order/catalog/product/A22283?SID=srch-srp-A22283

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Four- to eight-week-old (C57BL6×CBA) F1 female mice were super-ovulated using injection of 5 IU of pregnant mare serum gonadotrophin (PMSG; Sigma-Aldrich). Forty-eight hours after PMSG injection, 5 IU of human chorionic gonadotrophin (HCG; Sigma-Aldrich) was administered. Superovulated females were set up for mating with eight-week-old or older (C57BL6×CBA) F1 males. Mice were maintained on a 12h light–dark cycle, ambient temperature 19/22°C, and humidity 45/65%.

Wild animals	No wild animals were used in this study
Field-collected samples	Cow ovaries were obtained from the abattoir and transported to laboratory in PBS kept at 38°C.
Ethics oversight	This study was approved by the Animal Ethics Committee and by the UK Home Office Licence Number 70/8560 and followed all relevant institutional and national guidelines and regulations. The cow work was approved by the Ethics Committee at the Royal Veterinary College. All experiments followed all relevant institutional and national guidelines and regulations

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	This is not applicable as we used donated embryos surplus to IVF treatment.
Recruitment	Informed consent was obtained from all couples that donated spare embryos following IVF treatment.
Ethics oversight	All human embryo experiments followed all relevant institutional and national guidelines and regulations. For the work conducted in UK: This study was approved by the UK Human Fertilisation and Embryology Authority (HFEA): research licence number R0162, and the Health Research Authority's Research Ethics Committee (Cambridge Central reference number 19/EE/0297). For the work conducted in France: 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. For the work conducted in Belgium: The use of human embryos donated to research was allowed by the Local Ethical Committee of UZ Brussel (BUN 143201526417) and the Belgian Federal Committee for research on human embryos (AdV057).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

1.2.3 Discussion

Embryologie comparée : les morphocinétiques

Une observation importante a été l'étude comparative des morphocinétiques de souris, d'humain, et de bovin.

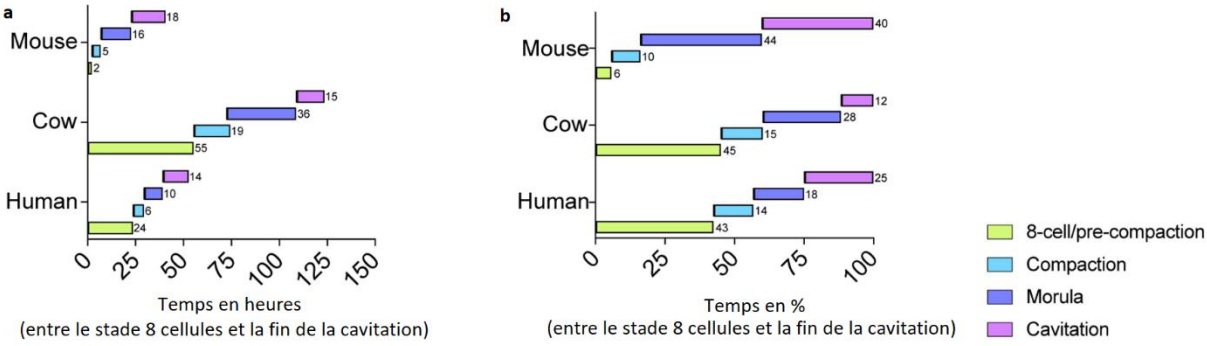


Figure 17 : morphocinétiques comparées du développement embryonnaire murin, humain, et bovin

(a-b) Mesure du temps en heures (a) ou en % (b) entre le stade 8 cellules et la fin de la cavitation. Figure adaptée de Gerri et al., 2020.

Chez la souris, la durée passée au stade morula est particulièrement longue (Figure 17a). Entre le stade 8-cellules et la fin de la cavitation, le stade morula occupe quasiment la moitié du temps de développement (44%), contre 18% chez l'Homme. En valeur absolue, le stade morula dure en moyenne 6h de plus chez la souris que chez l'Homme (16h contre 10h), alors que le développement préimplantatoire de la souris est plus court de 2,5 jours (4,5 jours chez la souris contre 7 pour l'Homme) (Shahbazi, 2020). Entre sa compaction au stade 8-16C et le début la cavitation à 32C, il y a donc environ deux cycles cellulaires chez la souris contre un seul chez l'homme. Ces informations peuvent expliquer pourquoi nous détectons la première spécification avant la blastulation chez la souris, mais au stade blastocyste B2-B3 chez l'homme.

Nous émettons l'hypothèse que l'embryon de souris a « plus de temps » pour spécifier ses cellules pendant sa longue phase en morula, là ou chez l'homme le stade morula est court. Aussi, il apparait que le temps en pourcentage du développement passé à chaque stade est relativement proche entre l'embryon humain et bovin comparé à la souris (Figure 17b). De même, des données morphocinétiques récentes montrent que le rat (Gerri et al., 2022) et le cochon d'inde (Petropoulos, données non publiées) ont des temps de développement proches de ceux de l'Homme et du bovin. Sur des critères morphocinétiques, il semble donc que l'embryon de souris se développe différemment des autres mammifères.

La souris, le meilleur ennemi de l'Homme ?

Nous avons montré que les morulas de souris, d'homme, et de bovin partagent certaines caractéristiques comme la polarisation des cellules externes et l'expression du marqueur GATA3 dans ces mêmes cellules polarisées. La présence de TEAD4 dans le noyau de toutes les cellules de la morula, et la restriction nucléaire de YAP1 dans les cellules externes suite à la polarisation montrent que la voie HIPPO est conservée entre ces espèces.

L'inhibition de la polarité a montré une diminution de l'expression nucléaire de YAP1 et GATA3 dans les cellules externes pour les trois espèces. Aussi, dans ces conditions, la majorité des embryons des trois espèces ne cavitent pas. Ces résultats montrent que la polarité est une étape critique du développement, nécessaire à l'établissement d'un TE *bona fide*, et ce dans plusieurs espèces de mammifères.

Cependant, nous notons une différence notable dans la localisation du facteur de transcription SOX2. Chez la souris, SOX2 est localisée uniquement dans les cellules internes apolaires de la morula (Gerri et al., 2020). En inhibant la polarisation, les

cellules externes adoptent une identité de cellules internes et SOX2 se retrouve exprimé dans les cellules externes. Chez l'Homme, le bovin, et le rat, SOX2 est exprimée dans toutes les cellules de la morula puis est progressivement restreint à l'épiblaste au stade blastocyste expansé (Cauffman et al., 2009; Goissis and Cibelli, 2014). En bloquant la polarisation chez l'homme et le bovin, l'expression et la localisation de SOX2 n'est pas modifiée.

Les embryons de souris *Tead4-/-* n'expriment pas ou peu CDX2 et GATA3, et stoppent leur développement au stade morula (Nishioka et al., 2008; Yagi et al., 2007). Chez l'homme, une récente étude a montré des résultats différents. Les embryons humains *TEAD4-/-* n'expriment pas CDX2, mais blastulent aussi souvent que les embryons contrôles (Stamatiadis et al., 2022). Aussi, l'expression de GATA3 dans le TE n'est pas perturbée.

En revanche, l'intégrité du TE s'en retrouve grandement dégradé avec la plupart des embryons ayant un TE au grade « C », suggérant que ces embryons ont peu de chances de s'implanter. Cette expérience suggère que :

- L'expression de *CDX2* est dépendante de TEAD4 chez l'homme
- L'expression de *GATA3* chez l'Homme est indépendante de celle de *TEAD4* contrairement à la souris (Ralston et al., 2010).
- L'intégrité du trophectoderme humain est dépendante d'une co-expression *GATA3-CDX2.*

Compte tenu des résultats de Gerri et Stamatiadis, il semble qu'il existe chez l'homme un mécanisme d'expression de *GATA3*, dépendant de la polarité mais indépendant de TEAD4. Des résultats préliminaires non publiés d'un KO-*GATA3* dans l'embryon humain suggèrent même que GATA3 ne serait pas strictement indispensable à la

formation du blastocyste (Cosemans et al., 2022). L'embryon humain *GATA3-/-* n'exprime qu'une à deux cellules CDX2 dans le TE. Chose étonnante, « quelques » cellules du TE expriment NANOG, et la fraction ICM/TE est augmentée. Ceci suggère que GATA3 est un inhibiteur du programme de pluripotence dans l'embryon humain.

Si ces observations venaient à être confirmées, il serait intéressant d'envisager *GATA3* comme l'un des premiers marqueurs du destin TE chez l'Homme, mais pas son initiateur. Nous avons réfléchi à d'autres acteurs qui pourraient être impliqués dans un mécanisme qui lierait polarité et destin cellulaire.

KLF5 est un facteur de transcription exprimé dans toutes les cellules de la morula de souris et progressivement restreint au TE du blastocyste ainsi qu'aux cellules du PrE (Lin et al., 2010; Parisi and Russo, 2011). Les embryons de souris *Klf5-/-* n'expriment pas *Cdx2* et stoppent leur développement au stade morula.

Une étude récente a montré que la protéine KLF5 est un marqueur du TE dans le blastocyste humain (Fogarty et al., 2021). Curieusement, nos données transcriptomiques montrent pourtant que le gène *KLF5* est exprimé dès le stade 8 cellules, avant même *GATA3,* et reste fortement exprimé dans les lignées EPI, TE, PrE jusqu'au 7ème jour. Dans les contextes du cancer du sein et du colon, il a été montré que KLF5 forme un complexe avec YAP1 pour être stabilisé et éviter sa dégradation par le protéasome (Wei et al., 2017; Zhi et al., 2012). Il serait intéressant d'envisager KLF5 comme un acteur de la voie HIPPO. YAP1 serait capable de stabiliser KLF5 dans le noyau des cellules externes polarisées de la morula et du blastocyste, et ce indépendamment de TEAD4. Nous pourrions envisager l'axe YAP1-KLF5 comme une voie précoce de l'établissement du destin TE chez l'Homme.

1.3 Manuscrit secondaire #2 : Human blastoids model blastocyst development and implantation.

1.3.1 Contexte de l'étude

Depuis 2019, nous collaborons avec Nicolas Rivron qui était intéressé par nos travaux sur les cellules souches trophoblastiques (Castel et al., 2020). Il prévoyait d'utiliser ces cellules comme analogue du TE, et tenter de former des blastoïdes humains comme il l'avait fait chez la souris un an plus tôt (Rivron et al., 2018).

En 2020, il a finalement pu générer des blastoïdes humains avec seulement des cellules souches pluripotentes naïves. Nous avons dès lors intensifié notre collaboration en l'aidant à valider son modèle de blastoïde.

Notre expertise couvrait le développement péri-implantatoire humain ainsi que les modèles cellulaires qui y sont associés. Dans cette collaboration, nous avions deux objectifs :

- Montrer que les blastoïdes ressemblent aux embryons humains.
- Proposer des hypothèses issues de nos travaux sur l'embryon humain à tester sur le modèle blastoïde.

Nous avons montré que la formation et le développement des blastoïdes ressemblent à ceux des embryons humains. Nous avons également pu tester et valider des hypothèses que nous avions chez l'Homme, notamment le rôle de l'IL6 dans le dialogue moléculaire EPI/TE.

1.3.2 Manuscrit Kagawa, Javali, Khoei et al. 2021

Kagawa H, Javali A, Khoei HH, Sommer TM, Sestini G, Novatchkova M, Scholte op Reimer Y, Castel G, **Bruneau A**, Maenhoudt N, Lammers J, Loubersac S, Freour T, Vankelecom H, David L, Rivron N. 2021. Human blastoids model blastocyst development and implantation. Nature 1–6. doi:10.1038/s41586-021-04267-8

Article

Human blastoids model blastocyst development and implantation

https://doi.org/10.1038/s41586-021-04267-8

Received: 12 February 2021

Accepted: 18 November 2021

Published online: 02 December 2021

Open access

Check for updates

Harunobu Kagawa[1,7], Alok Javali[1,7], Heidar Heidari Khoei[1,7], Theresa Maria Sommer[1], Giovanni Sestini[1], Maria Novatchkova[1,2], Yvonne Scholte op Reimer[1], Gaël Castel[3], Alexandre Bruneau[3], Nina Maenhoudt[4], Jenna Lammers[3,5], Sophie Loubersac[3,5], Thomas Freour[3,5], Hugo Vankelecom[4], Laurent David[3,6] & Nicolas Rivron[1]✉

One week after fertilization, human embryos implant into the uterus. This event requires the embryo to form a blastocyst consisting of a sphere encircling a cavity lodging the embryo proper. Stem cells can form a blastocyst model that we called a blastoid[1]. Here we show that naive human pluripotent stem cells cultured in PXGL medium[2] and triply inhibited for the Hippo, TGF-β and ERK pathways efficiently (with more than 70% efficiency) form blastoids generating blastocyst-stage analogues of the three founding lineages (more than 97% trophectoderm, epiblast and primitive endoderm) according to the sequence and timing of blastocyst development. Blastoids spontaneously form the first axis, and we observe that the epiblast induces the local maturation of the polar trophectoderm, thereby endowing blastoids with the capacity to directionally attach to hormonally stimulated endometrial cells, as during implantation. Thus, we propose that such a human blastoid is a faithful, scalable and ethical model for investigating human implantation and development[3,4].

A model of the human blastocyst would support scientific and medical progress. Its ability to predict human development will, however, depend on its ability to reproduce the sequences of blastocyst cellular determination and morphogenesis effectively, faithfully, and according to the developmental sequence and pace. Such modelling would ensure the formation of cells that reflect the blastocyst stage as a starting point to recapitulate aspects of subsequent developmental steps, including implantation. During this year, diverse ways of forming models of the human blastocyst have been proposed[5–9]. However, the cells generated often do not match those of the blastocyst[5,7–9] (at 5–7 days post fertilization (dpf)) and have been proposed to rather reflect later developmental stages, including gastrulation (E14) and germ layers (mesoderm and endoderm) stages[10]. Here we form a model of the human blastocyst that specifically generates and spatially patterns cellular analogues of the blastocyst stage with similar developmental sequence and pace, which enables the model to mimic aspects of implantation.

Inhibition of Hippo, ERK and TGFβ pathways

At 4 dpf, the conceptus forms a morula that initiates cavitation to make a blastocyst. Blastocyst development (at 5–7 dpf) supports the generation of the three founding lineages[11]: the epiblast (EPI), which is embryonic; trophectoderm (TE), which is extraembryonic; and primitive endoderm (PrE), which is extraembryonic (Fig. 1a). Peripheral cells become TE through inhibition of the Hippo pathway[12,13]. Naive human pluripotent stem cells (PSCs) cultured in PXGL[2] efficiently form TE analogues upon inhibition of TGFβ and ERK pathways[14–16]. We aggregated naive PSCs in non-adherent hydrogel microwells and inhibited these three pathways (Fig. 1b, Extended Data Fig. 1ac). Upon exposure to lysophosphatidic acid (LPA) (a Hippo pathway inhibitor), A83-01 (an inhibitor of TGFβ family receptors) and PD0325901 (an ERK inhibitor) in a chemically defined medium containing the STAT activator leukaemia inhibitory factor (LIF) and Y-27632 (a ROCK inhibitor), blastocyst-like structures formed efficiently (Fig. 1c–e, Supplementary Videos 1, 2; more than 70% efficiency, diameters 150–250 µm; full morphometric criteria are presented in Methods) and consistently (Extended Data Fig. 1d, more than 20 passages). LPA was essential for this high efficiency (Extended Data Fig. 1b–d). Within 4 days, the cell number (47 ± 9 to 129 ± 27) and overall size (65–200 µm) had increased (Extended Data Fig. 1e, f) to ranges similar to those for 5–7 dpf blastocysts[17] (stages B3 to B6). TE cell analogues[11] (identified as $GATA2^+GATA3^+CDX2^+TROP2^+$) formed, proliferated (Fig. 1f–h, Extended Data Fig. 1g–l), and established adherens junctions (marked by epithelial cadherin (CDH1)), apical–basal polarity (indicated by atypical PKC (aPKC) localization) and tight junctions (marked by ZO-1; Fig. 1i, Extended Data Fig. 1m) while undergoing cycles of inflations and deflations[18] (Extended Data Fig. 1n, Supplementary Video 2). Of note, all blastocyst-like structures set apart a unique inner cell cluster reflecting the EPI ($OCT4^+$; 27 ± 13 cells; 26% of total cells) and PrE ($GATA4^+SOX17^+PDGFRa^+$; 7 ± 5 cells; 7% of total cells) (Fig. 1f–h, Extended Data Fig. 1i, j, l). Multiple lines of naive human embryonic stem (ES) cells (Shef6, H9 and HNES1) and human naive induced PSCs (niPSC 16.2.b and cR-NCRM2) formed similar structures with comparably high efficiency (Fig. 1e, Extended Data Fig. 1o), whereas primed PSCs that reflect the post-implantation EPI did not (Extended Data Fig. 1p).

[1]Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna BioCenter (VBC), Vienna, Austria. [2]Institute of Molecular Pathology (IMP), Vienna Biocenter, Vienna, Austria. [3]Université de Nantes, CHU Nantes, INSERM, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ITUN, Nantes, France. [4]Unit of Stem Cell Research, Cluster of Stem Cell and Developmental Biology, Department of Development and Regeneration, KU Leuven, (University of Leuven), Leuven, Belgium. [5]CHU Nantes, Service de Biologie de la Reproduction, Nantes, France. [6]Université de Nantes, CHU Nantes, INSERM, CNRS, SFR Santé, FED 4203, INSERM UMS 016, CNRS UMS 3556, Nantes, France. [7]These authors contributed equally: Harunobu Kagawa, Alok Javali, Heidar Heidari Khoei. ✉e-mail: nicolas.rivron@imba.oeaw.ac.at

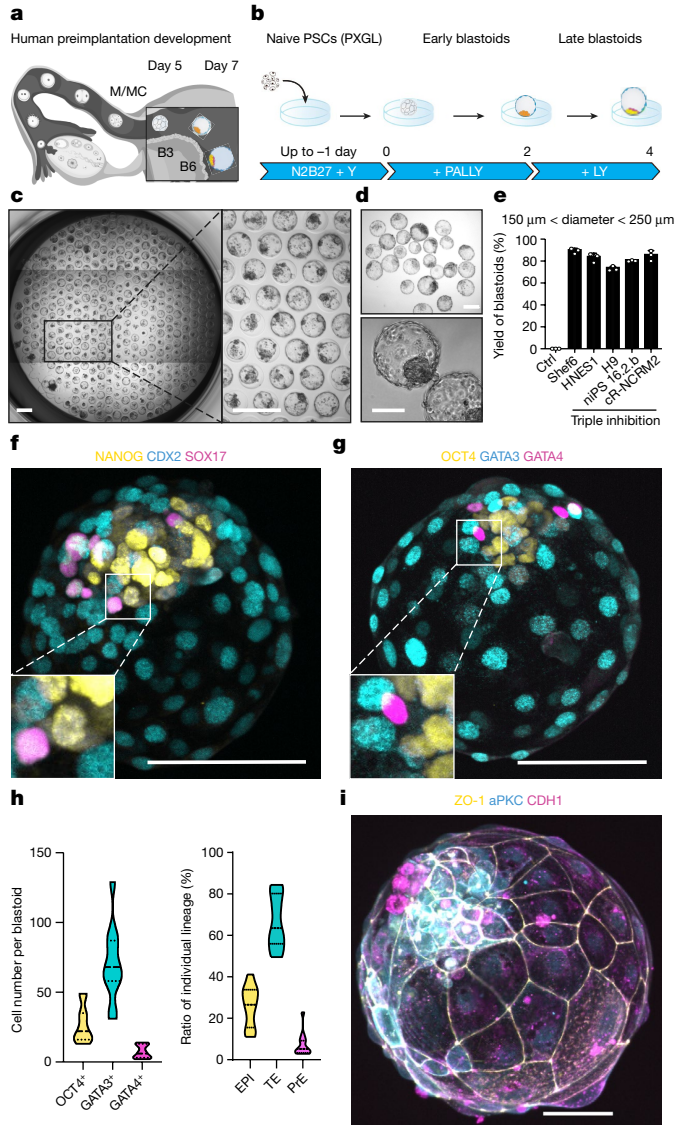


Fig. 1 | Triply inhibited naive PSCs efficiently form human blastocyst-like structures comprising analogues of the three founding lineages. **a**, A schematic of the time window of human peri-implantation development modelled by blastoids (days 5–7). M/MC, morula/morula compacted; B, blastocyst. **b**, One-step protocol of human blastocyst-like structure formation. N2B27, serum-free medium; PALLY, PD0325901 + A83-01 + LPA + hLIF + Y-27632. **c**, Phase-contrast image of human blastocyst-like structures formed on a non-adherent hydrogel microwell array after 96 h. Each microwell is 200 µm in diameter. Scale bars, 400 µm. **d**, Phase-contrast images of representative human blastocyst-like structures harvested from microwells. Scale bars, 200 µm (top) and 100 µm (bottom). **e**, Percentages of microwells including a human blastocyst-like structure for different naive PSC lines cultured in the PALLY condition with optimized LPA concentration compared with a H9 control (Ctrl) deprived of the three inhibitors. The morphometric definition of blastocyst-like structures is provided in Methods. $n = 3$ microwell arrays; mean ± s.d. **f**, **g**, Immunofluorescence of the epiblast (EPI) markers NANOG (yellow) (**f**) and OCT4 (yellow) (**g**), the TE markers CDX2 (cyan) (**f**) and GATA3 (cyan) (**g**), and the PrE markers SOX17 (magenta) (**f**) and GATA4 (magenta) (**g**) in human blastocyst-like structures. Scale bars, 100 µm. **h**, Absolute number of cells positive for OCT4, GATA3 and GATA4 (left) and ratios of cells belonging to individual lineages represented as percentage of total number of cells (right) in blastocyst-like structures (96 h) based on immunofluorescence. **i**, Representative immunofluorescence of the tight junction molecule ZO-1 (yellow), the adherence junction molecule CDH1 (magenta) and the apical domain molecule aPKC (cyan) in a representative human blastocyst-like structure. Scale bar, 50 µm.

Formation of blastocyst-stage analogues

Single-cell transcriptomics analysis showed that blastocyst-like structures formed only three distinct transcriptomic states (Fig. 2a, b, Extended Data Fig. 2a) marked by genes specific to the three founding lineages, including *GATA2* and *GATA3* (TE), *POU5F1* and *KLF17* (EPI), and *GATA4* and *SOX17* (PrE) (Fig. 2c, d, Extended Data Fig. 2b). Comparison with cells from blastocysts, in vitro cultured blastocysts and a gastrulation-stage embryo indicated that the cells in the blastocyst-like structures were transcriptionally similar to the blastocyst stage and distinct from post-implantation stages (Fig. 2e, f, Extended Data Fig. 2c–g). A higher-resolution clustering analysis (from resolution 0.02 to resolution 1) isolated one cluster of non-blastocyst-like cells with a gene-expression pattern reminiscent of post-implantation tissues[15] (*GABRP*, *ISL1*, *APLNR* and *CRABP2*) (Extended Data Fig. 3a–c) that also appeared transcriptionally similar to amnion (annotated as non-neural ectoderm) and extra-embryonic mesoderm (Extended Data Fig. 3d–j). This sub-population constituted less than 3% of all sequenced cells (Extended Data Fig. 3i). Of note, we found that naive PSC culture also contained 5–6% similarly differentiated cells[19] (Extended Data Fig. 3i). Bulk RNA-sequencing (RNA-seq) analysis showed that isolated trophoblast analogues (TROP2$^+$) had an intermediate transcriptome between those of naive PSCs and post-implantation-like trophoblasts[20] (TSCs) (Extended Data Fig. 4a). Furthermore, trophoblasts were enriched in blastocyst-stage TE transcripts[11] (*ESRRB*, *GRHL1*, *OVOL1*, *GATA2*, *GATA3*, *TBX3*, *KRT19*, *CGA*, *CGB5* and *CGB7*) but not in some post-implantation trophoblast markers[11] (*SIGLEC6* and *DPP4*) (Extended Data Fig. 4b, c). The transcriptome of isolated EPI analogues (TROP2$^-$PDGFRa$^-$) resembled that of naive PSCs (Extended Data Fig. 4a), was enriched in markers specific for blastocyst-stage EPI[21] (*KLF17*, *ATG2A*, *SUSD2*, *TFCP2L1*, *DPPA2* and *PRDM14*), and differed from the transcriptome of primed PSCs (Extended Data Fig. 4a, d). Finally, isolated PrE analogues (PDGFRa$^+$) had an intermediate transcriptome between those of naive PSCs and extraembryonic endoderm cell lines[22] (nEND cells) (Extended Data Fig. 4a). PrE analogues were enriched in blastocyst-stage PrE markers (early blastocyst: *GATA6*, *MSX2* and *HNF4A*; late blastocyst: *PDGFRA*, *GATA4*, *SOX17*, *HNF1B* and *FOXA2*) and downregulated in EPI genes (*ARGFX*, *PRDM14*, *SOX2*, *NANOG*, *DPPA2* and *POU5F1*), similar to during blastocyst development[21] (Extended Data Fig. 4e). Blastocysts have the ability to establish stem cell lines[2]; similarly, blastocyst-like structures enabled de novo derivation of naive PSCs[2] (NANOG$^+$SOX2$^+$OCT4$^+$KLF17$^+$) (Extended Data Fig. 5a) that could form second-generation blastocyst-like structures (Extended Data Fig. 5b, c) and of TSCs[20] (CDX2$^-$GATA3$^+$CK7$^+$) (Extended Data Fig. 5d) endowed with the capacity for rapid differentiation into syncytio trophoblasts (SCT) and extravillous trophoblasts (EVT) (over 3–6 days; Extended Data Fig. 5e–j). Of note, derivation of PrE cell lines from human blastocysts has not been reported. Thus, blastocyst-like structures formed blastocyst-stage cellular analogues (accounting for more than 97% of the cells sequenced).

Hippo inhibition is essential

Knowledge about human blastocyst lineage segregation is limited (Fig. 3a). However, inhibition of the Hippo pathway is known to occur in peripheral cells upon acquisition of an apical domain, and is required to initiate TE specification[12] (Extended Data Fig. 6a). We tested whether blastocyst-like structures co-opted this mechanism. Of note, aPKC and F-actin expression domains appeared co-aligned in outer cells that also accumulated the Hippo downstream effector YAP1 in nuclei (Extended Data Fig. 6b, c). YAP1 nuclear location correlated with GATA2 and GATA3 expression, contrasted with NANOG expression, and became restricted to TE analogues[12] (Fig. 3b, Extended Data Fig. 6d, e). An aPKC inhibitor (CRT0103390)[12] largely prevented YAP1 nuclear accumulation, decreased the number of GATA3$^+$ cells

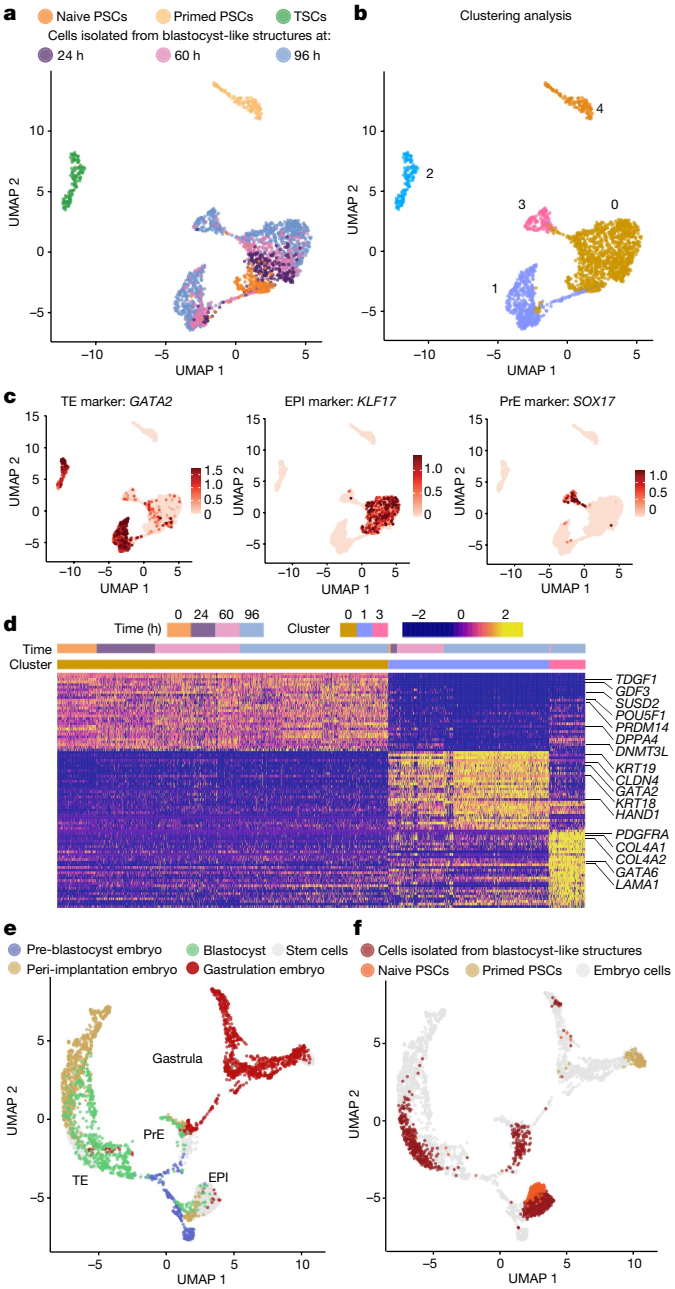


Fig. 2 | Human blastocyst-like structures form analogues of the three pre-implantation lineages. a, **b**, Uniform manifold approximation and projection (UMAP) of the transcriptome of single cells originating from blastocyst-like structures (at 24, 60 and 96 h), naive PSCs, primed PSCs and TSCs (representing post-implantation cytotrophoblasts); individual cells are coloured on the basis of their origin (**a**) or their unsupervised cluster affiliation (**b**). **c**, Expression level of markers of each blastocyst lineage. **d**, Unsupervised distance map generated using the top 30 genes that are enriched in clusters 0, 1 and 3 (defined in the UMAP in **b**). Note that this list includes epiblast markers specific to the blastocyst stage (for example, SUSD2, KLF17 and PRDM14). **e**, **f**, UMAP of single-cell transcriptome of cells from blastocyst-like structures, naive PSCs and primed PSCs integrated with published datasets from human embryos at pre-implantation, peri-implantation (in vitro cultured blastocysts) and gastrulation (Carnegie stage 7, that is, between embryonic days 16 and 19) stages. Individual cells are coloured on the basis of their origin in human embryos (**e**) or blastocyst-like structures or stem cells (**f**).

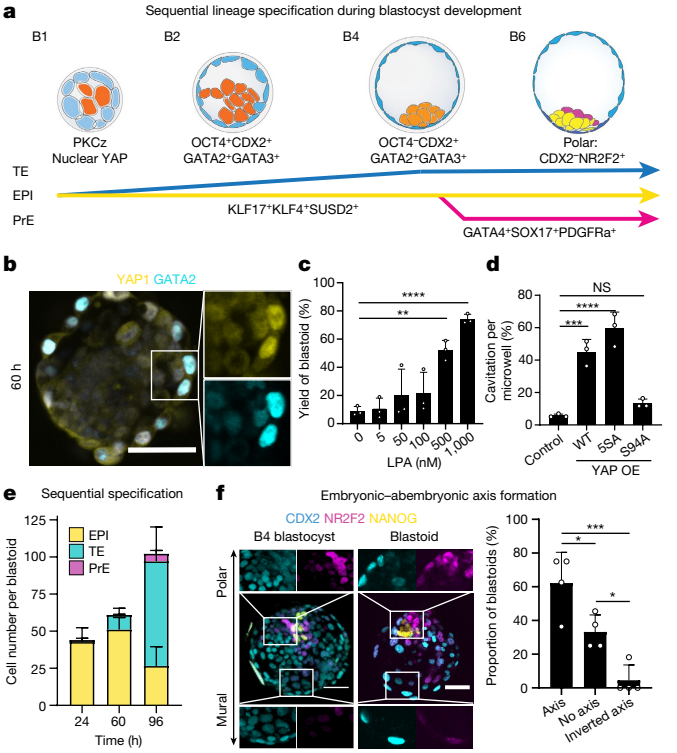


Fig. 3 | The three lineages form according to the sequence and time of blastocyst development. a, Schematic depicting the sequential lineage specification of human blastocysts. **b**, Immunofluorescence of YAP1 (yellow) and GATA2 (cyan) in aggregates of naive PSCs cultured in PALLY medium for 60 h. Scale bar, 50 µm. **c**, Dose-dependent effect of LPA on the yield of blastocyst-like structures. $n = 3$ independent microwell arrays; mean ± s.d.; one-way analysis of variance (ANOVA) and Dunnett's multiple comparisons test. $^{**}P = 0.0016$, $^{****}P < 0.0001$. **d**, Effect of the overexpression of different variants of YAP1 on cavitation events in early blastocyst-like structures. $n = 3$ experiments; mean ± s.d.; one-way ANOVA and Tukey's multiple comparisons test. NS, not significant; $^{***}P = 0.0004$, $^{****}P = 0.00004$. **e**, Total cell numbers per lineage developing blastocyst-like structures at three time points of development (24, 60 and 96 h). Mean ± s.d. EPI: $n = 11$ blastocyst-like structures at 24, 68 and 96 h; TE: $n = 8$ (24 h), $n = 14$ (48 h) and $n = 15$ (96 h) blastocyst-like structures; PrE: $n = 9$ (24 h), $n = 37$ (48 h) and $n = 9$ (96 h) blastocyst-like structures. **f**, Immunofluorescence of CDX2 (cyan), NR2F2 (magenta) and NANOG (yellow) in representative B4-stage human blastocyst (left) and blastocyst-like structures (middle). Quantification of the proportion of blastocyst-like structures with a preferentially polar NR2F2 expression pattern (axis) compared with a preferentially mural NR2F2 expression pattern (inverted axis) (right). $n = 4$ independent experiments with 4–12 blastocyst-like structures in each experiment; mean ± s.d.; one-way ANOVA and Tukey's multiple comparisons test. $^{*}P < 0.05$, $^{***}P < 0.001$. Scale bar, 50 µm.

and prevented the formation of blastocyst-like structures (Extended Data Fig. 6f–h). Conversely, ligands of LPA receptors (LPA and NAEPA) that can inhibit the Hippo pathway enhanced the formation of blastocyst-like structures (Fig. 3c, Extended Data Fig. 6i). Because Hippo pathway inhibition frees YAP1 to enter the nucleus, we tested whether genetically engineered levels and functions of YAP1 could affect morphogenesis. Overexpression of wild-type or constitutively active forms of YAP1 (5SA) accelerated cavitation (Fig. 3d). The interaction between YAP1 and TEAD transcription factors is necessary for downstream gene regulation. Accordingly, over-expression of YAP1 with a mutation in the TEAD binding site (S94A) did not affect cavitation (Fig. 3d, Extended Data Fig. 6j), and verteporfin—which disrupts the YAP1–TEAD interaction—prevented the formation of blastocyst-like structures (Extended Data Fig. 6k). Cavity morphogenesis occurred through the apparent coalescence of multiple fluid-filled cavities[23] (Extended Data Fig. 6l). Aquaporin 3 (AQP3), the water transporter most highly expressed in human blastocysts[11], was initially visible in all cells (36 h) and was then restricted to TE analogues (96 h) (Extended

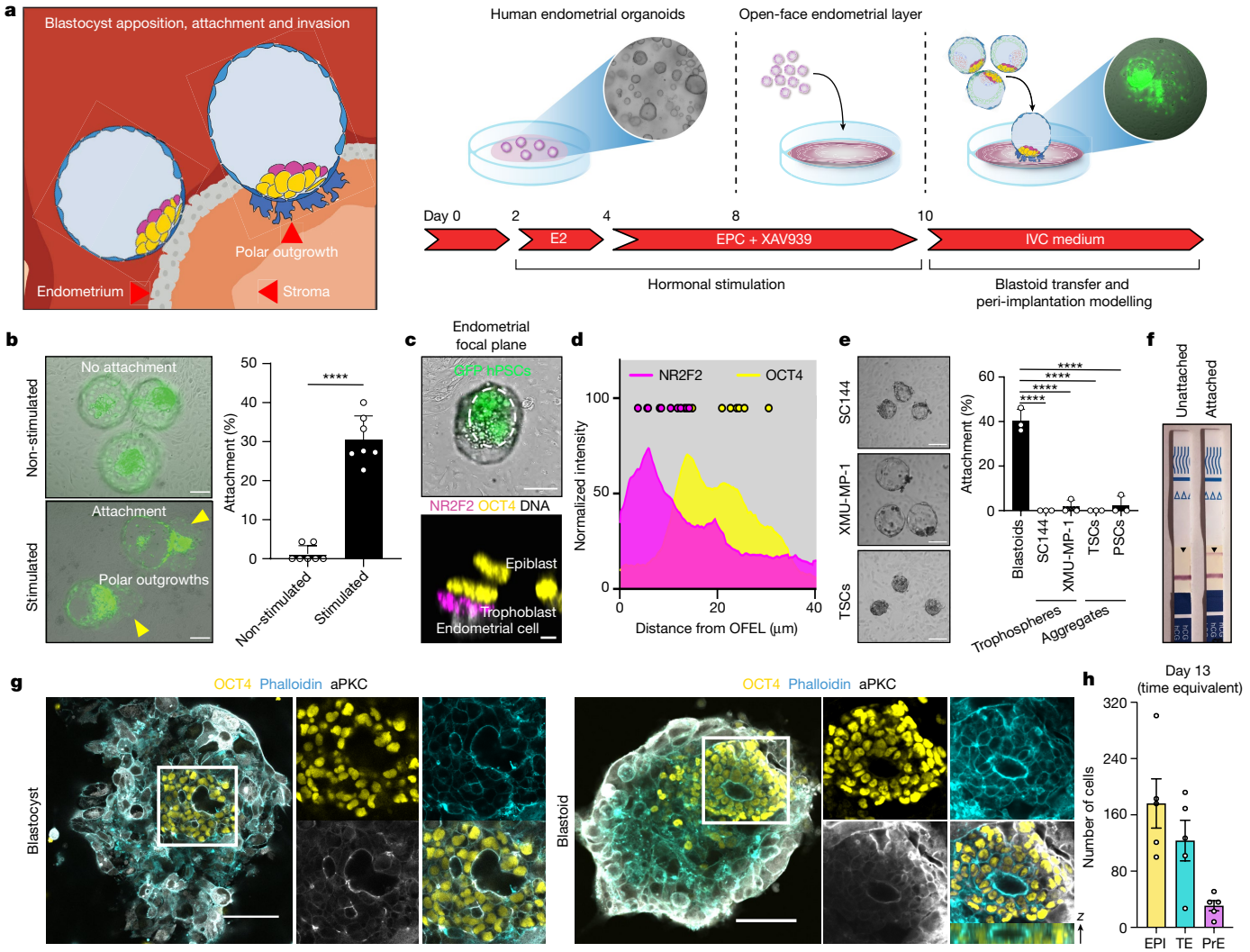


Fig. 4 | Human blastoids recapitulate aspects of implantation. a, Left, schematic of the modelled implantation process. Right, OFEL priming assay using EPC/XAV939. E2, β-oestradiol; EPC, E2 + progesterone + cAMP. **b**, Representative phase-contrast images of blastoids (GFP$^+$) 24 h after deposition onto non-stimulated (top left) or stimulated (bottom left) OFELs. Scale bar, 100 μm. Attachment efficiency of human blastoids (right). $n = 7$ independent experiments from 3 different donors; mean ± s.d.; unpaired two-tailed t-test. ****$P = 4.5 \times 10^{-8}$. **c**, Representative images of recently attached human blastoids (12 ± 4 h). Top, the dashed delineates the inner cluster of blastoids formed from GFP$^+$ naive PSCs (also see Supplementary Video 3). Scale bar, 100 μm. Bottom, x–z plane of NR2F2 (magenta) and OCT4 (yellow) immunofluorescence in blastoids immediately after attachment. Scale bar, 5 μm. **d**, Intensity profile of immunofluorescence of NR2F2 and OCT4 in blastoids immediately after attachment. $n = 10$. **e**, Left, representative phase-contrast images of trophospheres formed using 3 μM SC144 (top) or 2 μM XMU-MP-1 (middle), and aggregates of TSCs (bottom) deposited onto stimulated OFELs. Scale bar, 100 μm. Right, attachment efficiency. $n = 3$ independent experiments; mean ± s.d.; one-way ANOVA and Dunnett's multiple comparisons test. ****$P < 0.0001$. **f**, Pregnancy test strips detecting secretion of CGβ into the medium of unstimulated OFELs with unattached blastoids and stimulated OFELs with attached blastoids (48 h on OFEL; see ELISA assay in Extended Data Fig. 10b). **g**, Immunofluorescence of OCT4 (yellow) and aPKC (grey) in human blastocysts (left) or blastoids (right) grown in post-implantation culture condition for 4 days, counterstained with phalloidin marking F-actin (cyan). Scale bars, 100 μm. **h**, Number of cells positive for OCT4, GATA3 and GATA4 in blastoids grown in post-implantation culture for 6 days (time equivalent, day 13). $n = 5$. mean ± s.d.

Data Fig. 6m). Thus, similar to human blastocysts[12], TE specification and morphogenesis within these structures depends on aPKC, inhibition of the Hippo pathway, nuclear translocation of YAP1 and the ability of YAP1 to bind TEAD transcription factors.

Adequate developmental sequence

In blastocysts, TE (GATA2$^+$DAB2$^+$) and EPI (KLF17$^+$NANOG$^+$) cells appear first[11,21] (5–6 dpf) and PrE cells (GATA6$^+$ADM$^+$) and polar TE cells (pTE) (CDX2$^-$NR2F2$^+$) appear last[21] (6–7 dpf). This sequence is recapitulated in the blastocyst-like structures. Trophoblasts (*DAB2*$^+$, *CDX2*$^+$, *GATA2*$^+$, *GATA3*$^+$) formed first (within 24 h) (Fig. 3e, Extended Data Fig. 7a), and changed the levels of transcripts related to PKC and Hippo signalling (*AKAP12*, *CAPZB*, *ULK4*, *MOB1A*, *AMOT*, *AMOTL2*, *LATS2* and *TEAD1*) (Supplementary Table 1). At protein level, early TE-like cells were first YAP1^{nuclear}GATA2$^+$ (at 24 h) and then CDX2$^+$GATA3$^+$, while maintaining expression of KLF17 and OCT4, but not NANOG (at 60 h) (Extended Data Fig. 7b–d). Subsequently, OCT4 became undetectable[11] (Fig. 1g, Extended Data Fig. 1i). Genes associated with SMAD, ERK, Notch and Wnt signaling pathways were regulated during this process (Extended Data Fig. 7e, f, Supplementary Table 1). Finally, pTE analogues matured as marked by expression of *OVOL1*, *GREM2*, *CCR7*, *SP6* and *NR2F2* (Extended Data Fig. 7g–j), upregulation of NR2F2 and CCR7[11] and downregulation of CDX2 (Fig. 3f, Extended Data Fig. 7h, j).

The transcriptome of EPI analogues maintained core blastocyst markers (*POU5F1*, *NANOG*, *KLF17*, *SUSD2*, *KLF4*, *ARGFX* and *GDF3*) (Fig. 3e, Extended Data Fig. 7k, l, Supplementary Table 1), while undergoing a progression characterized by regulation of Nodal (*NODAL*, *LEFTY1* and *LEFTY2*) and mTOR (*LAMTOR1*, *LAMTOR4*, *LAMTOR5*, *XBP1P1* (*XBP1*, also known as) *SEC13* and *MLST8*) signalling-related genes, and of the X chromosome activation-related gene *XACT* (Extended Data Fig. 7k–m, Supplementary Table 1, cluster 4 versus cluster 0). At the protein level, EPI analogues were marked by KLF17 and SUSD2, which are specifically highly expressed at the blastocyst stage (Extended Data Fig. 7l). PrE analogues appeared within 60 h and GATA4, OTX2 and SOX17 were detected[11] within 72 h (Fig. 3e, Extended Data Fig. 7n–p). Early PrE marker genes[21] (*GATA6*, *LBH*, *ADM* and *LAMA1*) were uniformly expressed among the PrE analogues, while some late PrE marker genes (*CTSE*, *APOA1*, *PITX2* and *SLCO2A1*) were expressed only in a subpopulation of cells, suggesting a progression toward the late blastocyst stage[11] (Extended Data Fig. 7q). By 96 h, mature PrE analogues had regulated SMAD (*NODAL*, *BMP2*, *BMP6*, *GDF3*, *ID1* and *IDI2*) and Wnt signalling-related transcripts (*WNT3*, *RSPO3* and *LBH*) and were enriched in transcripts controlling extracellular matrix organization (*LAMA1*, *LAMB1*, *LAMC1*, *COL4A1* and *COL4A2*), and endodermal and epithelial differentiation (Extended Data Fig. 7q, r, Supplementary Table 1, cluster 6 versus cluster 8). Because this model morphologically resembles the human blastocyst (see criteria in Methods), efficiently generates analogues of the three lineages with transcriptomes matching the blastocyst stage, and forms these analogues according to the sequence and approximate pace of blastocyst development, we refer to it as a human blastoid.

Distinct attachment to endometrial cells

At 7 dpf, the human blastocyst initiates implantation in utero through the attachment of its TE to a receptive endometrium (Fig. 4a, left). We tested whether blastoids could model this interaction by seeding endometrial organoids[24] in 2D to form an open-faced endometrial layer (OFEL) to facilitate the deposition of blastoids (Fig. 4a, right). Subpopulations of the OFEL were positive for acetylated α-tubulin, marking ciliated epithelial cells[24] (Extended Data Fig. 8a), and FOXA2, marking glandular epithelial cells (Extended Data Fig. 8b). The window of implantation is the period during which the uterus becomes receptive for blastocyst implantation. It opens upon exposure to oestrogen (E2) and progesterone (P4), and correlates with regulation of Wnt signalling[25]. Accordingly, OFELs responded to E2, P4, cAMP and XAV939 by upregulating the expression of genes that mark the mid-secretory-phase endometrium (Extended Data Fig. 8c–e) and decreasing proliferation, which are hallmarks of receptivity[25] (Extended Data Fig. 8e, f). Notably, blastoids deposited onto non-stimulated OFELs did not attach; however, they did attach to and repel the endometrial cells of stimulated OFELs, as occurrs in utero (Fig. 4b, Extended Data Fig. 8g, h). The contraceptive levonorgestrel impaired blastoid attachment (Extended Data Fig. 8i). We concluded that human blastoids are capable of interacting specifically with endometrial cells that have been made receptive.

Epiblast signals gatekeep trophectoderm attachment

Human blastocysts attach to the endometrium via the pTE, which is defined by its contact with the EPI. Similarly, blastoids reproducibly initiated attachment through this region (Fig 4c, d, Extended Data Fig. 9a–c, Supplementary Videos 3, 4). We tested the role of the pTE–EPI interface by forming trophospheres devoid of EPI. IL-6 is highly expressed in the pTE and transcripts for its receptor (*IL6R* and *IL6ST* (also known as *GP130*)) and effector (*STAT3*) are present at high levels in the EPI (Extended Data Fig. 9d). Consistent with a role for STAT signalling in the EPI, the efficiency of blastoid formation increased with LIF concentration (Extended Data Fig. 9e), whereas the addition of a GP130 inhibitor (SC144) yielded trophospheres (Fig. 4e, Extended Data Fig. 9f). The presence of a potent inhibitor of the Hippo kinases MST1 and MST2 (XMU-MP-1) also yielded trophospheres (Fig. 4e, Extended Data Fig. 9g). The transcriptomes of these trophospheres reflected early and late blastocyst-stage TE (Extended Data Fig. 9h, i). Neither type of trophosphere attached to OFELs (Fig. 4e), and nor did aggregates of TSCs[20] that reflect post-implantation cytotrophoblasts[26] ($CDX2^{-}CK7^{+}$) or aggregates of naive PSCs (Fig. 4e, Extended Data Fig. 9j, k). We thus conclude that signals from the EPI induce pTE maturation and endows it with the potential to interact with endometrial cells. This potential appears lost in TSCs reflecting a post-implantation stage. On the basis of transcriptome analysis and in utero data[25], we propose several pairs of molecules whose transcripts became more abundant upon endometrial cell stimulation and pTE analogue maturation (Extended Data Fig. 9l) s that might mediate the first contact between blastocyst and uterus. Overall, we conclude that a polar-like TE state, whose maturation depends on EPI inductions, gatekeeps the interaction of the blastocyst with the endometrium. This interaction and subsequent maturation create a window of opportunity for blastocyst implantation.

Modelling post-stages on day 13

The blastoid morphology was stable for two days in peri-implantation culture conditions[27,28] (Extended Data Fig. 10a). Clinical pregnancy is characterized by the detection of chorionic gonadotropin-β hormone (CGβ). Accordingly, upon attachment, blastoids formed trophoblasts expressing CGβ at levels detectable using standard pregnancy tests and ELISA (Fig. 4f, Extended Data Fig. 10b). $NR2F2^{+}$ pTE analogues proliferated and decreased CDX2 expression while upregulating the peri-implantation gene cytokeratin 7 (*KRT7* (a.k.a.*CK7*)) (Extended Data Fig. 10c, d). Some trophoblasts further differentiated into SCT and EVT expressing CGβ and HLA-G, respectively (Extended Data Fig. 10e, f). EPI analogues maintained expression of OCT4 and SOX2, upregulated the primed pluripotency marker CD24 (Fig. 4g, Extended Data Fig. 10g) and patterned cortical F-actin as during post-implantation EPI epithelization, and some blastoids cultured in vitro for 4 days past the equivalent of the blastocyst stage (day 7) formed pro-amniotic-like cavities enriched with F-actin, PODXL and aPKC (Fig. 4g, Extended Data Fig. 10h). A subpopulation in the periphery of the EPI analogue expressed CDX2 along with SOX2 or TFAP2C, suggestive of early amnion analogues (Extended Data Fig. 10i, j). Extra-embryonic endoderm analogues were characterized by restricted expression of OTX2[11] (Extended Data Figs. 7o, 10k). Upon prolonged culture (up to 6 days), the three lineages consistently expanded (Fig. 4h, Extended Data Fig. 10l, m) until a time equivalent of day 13, although, similar to blastocysts, their organization did not reflect that developmental stage.

Discussion

Human blastoids morphologically resemble the human blastocyst (criteria described in Methods), efficiently generate analogues of its three lineages with transcriptomes matching the human blastocyst stage, and form these analogues according to the sequence (TE and EPI, then pTE and PrE) and approximate pace (4 days) of blastocyst development. We therefore propose that this model is relevant for the study of human blastocyst development and implantation. Some initial parameters and end-point criteria that are useful to form and define these models[5–9] are summarized in Supplementary Table 2. Mimicking the interaction between the epiblast and trophectoderm revealed that the epiblast induces the local maturation of polar trophectoderm and subsequently endows it with the capacity to attach onto stimulated endometrial cells. In future, human blastoids may be used to help identify therapeutic targets and contribute to preclinical modelling (for example, in vitro fertilization medium complements such as LPA and NAEPA or contraceptives such as SC144 (ref.[3])). Considering the proportionality

(balancing the benefits and harms) and subsidiarity (pursuing goals using the morally least problematic means) of human embryology, blastoids represent an ethical opportunity to complement research using embryos[4].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-04267-8.

1. Rivron, N. C. et al. Blastocyst-like structures generated solely from stem cells. *Nature* **557**, 106–111 (2018).
2. Guo, G. et al. Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass. *Stem Cell Rep.* **6**, 437–446 (2016).
3. Rivron, N. et al. Debate ethics of embryo models from stem cells. *Nature* **564**, 183–185 (2018).
4. Clark, A. T. et al. Human embryo research, stem cell-derived embryo models and in vitro gametogenesis: Considerations leading to the revised ISSCR guidelines. *Stem Cell Rep.* **16**, 1416–1424 (2021).
5. Liu, X. et al. Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* **591**, 627–632 (2021).
6. Yanagida, A. et al. Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* **28**, 1016–1022 (2021).
7. Yu, L. et al. Blastocyst-like structures generated from human pluripotent stem cells. *Nature* **591**, 620–626 (2021).
8. Sozen, B. et al. Reconstructing aspects of human embryogenesis with pluripotent stem cells. *Nat. Commun.* **12**, 5550 (2021).
9. Fan, Y. et al. Generation of human blastocyst-like structures from pluripotent stem cells. *Cell Discov.* **7**, 81 (2021).
10. Zhao, C. et al. Reprogrammed iBlastoids contain amnion-like cells but not trophectoderm. Preprint at https://doi.org/10.1101/2021.05.07.442980 (2021).
11. Meistermann, D. et al. Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. *Cell Stem Cell* **28**, 1625–1640.e6 (2021).
12. Gerri, C. et al. Initiation of a conserved trophectoderm program in human, cow and mouse embryos. *Nature* **587**, 443–447 (2020).
13. Rossant, J. Making the mouse blastocyst: past, present, and future. *Curr. Top. Dev. Biol.* **117**, 275–288 (2016).
14. Amita, M. et al. Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. *Proc. Natl. Acad. Sci. USA* **110**, E1212–E1221 (2013).
15. Io, S. et al. Capturing human trophoblast development with naive pluripotent stem cells in vitro. *Cell Stem Cell* **28**, 1023–1039.e13 (2021).
16. Guo, G. et al. Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* **28**, 1040–1056.e6 (2021).
17. Hardy, K., Handyside, A. H. & Winston, R. M. The human blastocyst: cell number, death and allocation during late preimplantation development in vitro. *Development* **107**, 597–604 (1989).
18. Lewis, W. H. & Gregory, P. W. Cinematographs of living developing rabbit-eggs. *Science* **69**, 226–229 (1929).
19. Messmer, T. et al. Transcriptional heterogeneity in naive and primed human pluripotent stem cells at single-cell resolution. *Cell Rep.* **26**, 815–824.e4 (2019).
20. Okae, H. et al. Derivation of human trophoblast stem cells. *Cell Stem Cell* **22**, 50–63.e6 (2018).
21. Stirparo, G. G. et al. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* **145**, dev158501 (2018).
22. Linneberg-Agerholm, M. et al. Naïve human pluripotent stem cells respond to Wnt, Nodal, and LIF signalling to produce expandable naïve extra-embryonic endoderm. *Development* **146**, dev180620 (2019).
23. Dumortier, J. G. et al. Hydraulic fracturing and active coarsening position the lumen of the mouse blastocyst. *Science* **365**, 465–468 (2019).
24. Boretto, M. et al. Development of organoids from mouse and human endometrium showing endometrial epithelium physiology and long-term expandability. *Development* **144**, 1775–1786 (2017).
25. Wang, W. et al. Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. *Nat. Med.* **26**, 1644–1653 (2020).
26. Castel, G. et al. Induction of human trophoblast stem cells from somatic cells and pluripotent stem cells. *Cell Rep.* **33**, 108419 (2020).
27. Ma, H. et al. In vitro culture of cynomolgus monkey embryos beyond early gastrulation. *Science* **366**, eaax7890 (2019).
28. Xiang, L. et al. A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* **577**, 537–542 (2020).

Methods

Ethical approvals

The use of human embryos 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 following informed consent from the couples who donated embryos 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. Human endometrium samples were obtained from patients who signed an informed consent form and protocols approved by the Ethics Committee of Royan Institute (IR.ACECR.ROYAN.REC. 1397.93) and of Shahid Beheshti University of Medical Sciences (IR.SBMU.MSP.REC. 1396.25). The Wicell line H9 was used under the agreement 20-WO-341 for a research program entitled 'Modeling early human development: Establishing a stem cell based 3D in vitro model of human blastocyst (blastoids)'. Blastoid generation was approved by the Commission for Science Ethics of the Austrian Academy of Sciences. This work did not exceed a developmental stage normally associated with 14 consecutive days in culture after fertilization even though this is not forbidden by the ISSCR Guidelines as far as embryo models are concerned. All experiments complied with all relevant guidelines and regulations, including the 2021 ISSCR guidelines that forbid the transfer of human blastoids into any uterus[4].

Culture of human naive pluripotent stem cells

Experiments were done using the following PSC lines; human ES cell lines: H9, Shef6 and HNES1. Induced pluripotent stem cell (iPSC) lines: cR-NCRM2 and niPSC 16.2.b. The H9 and H9-GFP lines reset to the naive state were provided by the laboratory of Y. Takashima. Other naive human ES cells and iPSCs were provided by the laboratory of A. Smith. Naive PSCs were cultured on gelatin-coated plates including a feeder layer of gamma-irradiated mouse embryonic fibroblasts (MEFs) in PXGL medium, as previously reported[29]. PXGL medium is prepared using N2B27 basal medium supplemented with PD0325901(1 μM, MedChemExpress, HY-10254), XAV-939 (1 μM, MedChemExpress, HY-15147), Gö 6983 (2 μM, MedChemExpress, HY-13689) and human leukemia inhibitory factor (hLIF, 10 ng ml^{-1}, in-house made) as previously reported[29]. N2B27 basal medium contained DMEM/F12 (50%, in house made), neurobasal medium (50%, in-house made), N-2 supplement (Thermo Fisher Science, 17502048), B-27 supplement (Thermo Fisher Science, 17504044), GultaMAX supplement (Thermo Fisher Science, 35050-038), non-essential amino acid, 2-mercaptoethanol (100 μM, Thermo Fisher Science, 31350010), and bovine serum albumin solution (0.45%, Sigma-Aldrich, A7979-50ML). Cells were routinely cultured in hypoxic chambers (5% CO_2, 5% O_2) and passaged as single cells every three to four days. All cell lines had routinely tested negative for mycoplasma.

Culture of primed pluripotent embryonic stem cells

Primed H9 cells were cultured on Vitronectin XF (STEMCELL Technologies, 07180) coated plates (1.0 μg cm^{-2}) using Essential 8 medium (prepared in-house).

Microwell arrays

Microwell arrays comprising microwells of 200 μm diameter were imprinted into 96-well plates as previously described[30,31].

Induction of blastoids and trophospheres

Naive PSCs were treated with Accutase (Biozym, B423201) at 37 °C for 5 min, followed by gentle mechanical dissociation with a pipette. After centrifugation, the cell pellet was resuspended in PXGL medium, supplemented with Y-27632 (10 μM, MedChemExpress, HY-10583). To exclude MEFs, the cell suspension was transferred onto gelatin-coated plates and incubated at 37 °C for 70 min. After MEF exclusion, the cell number was determined using a Countess automated cell counter (Thermo Fisher Scientific) and trypan blue staining to assess cell viability. The cells were then resuspended in N2B27 medium containing 10 μM Y-27632 (aggregation medium) and 3.0×10^4 cells were seeded onto a microwell array included into a well of a 96-well plate and placed in a hypoxic chamber (5% CO_2, 5% O_2) for the whole period of blastoid or trophosphere formation. The cells were allowed to form aggregates inside the microwell for a period ranging from 0 to 24 h depending on the cell lines and based on their propensity for aggregation. Subsequently, the aggregation medium was replaced with PALLY medium (N2B27 supplemented with PD0325901 (1 μM), A 83-01 (1 μM, MedChemExpress, HY-10432), 1-oleoyl lysophosphatidic acid sodium salt (LPA)[32] (500 nM, Tocris, 3854), hLIF (10 ng ml^{-1}) and Y-27632 (10 μM)). The PALLY medium was refreshed every 24 h. After 48 h, the PALLY medium was replaced with N2B27 medium containing 500 nM LPA and 10 μM Y-27632. At 96 h, a blastoid is defined based (1) on morphological similarities to B6 staged human blastocyst, as a structure composed of a monolayered cyst with an overall diameter of 150–250 μm comprising one inner cell cluster, and (2) on similarities to the molecular dynamic of human development as a structure that forms analogues of the three blastocyst cell lineages in the sequential and timely manner of a blastocyst. For example, >90% of morphologically adequate structures generated from the lines analysed formed >97% of analogues of three blastocyst-stage lineages (see Fig. 1h and Extended Data Fig. 3i). An exception is the line Shef6, which efficiently formed morphologically correct structures but appeared less efficient at forming PrE analogues. See also Supplementary Table 2. Blastoids reproducibly formed at high efficiency and we did not observe differences based on the number of passages after resetting in PXGL culture conditions. The effect of LPA, NAEPA (Sigma-Aldrich, N0912) and Verteporfin (Selleck Chemicals Llc, S1786) on the yield of blastoid formation was assessed by culturing naive PSC aggregates in PALY medium (without LPA) complemented with molecules added every day from 0 to 96 h. The Verteporfin treatment was executed without exposure to the light. The effect of the aPKC inhibitor CRT0103390 (a gift from the laboratory of K. Niakan) was assessed by culturing naive PSC aggregates in PALLY medium complemented with 2 μM CRT0103390 every day from 0 to 96 h. The formation of trophospheres was induced by culturing naive PSC aggregates in PALLY medium complemented with 2 μM XMU-MP-1 (Med Chem Express, HY-100526) or 3 μM SC-144 (Axon, 2324) every day from 0 to 96 h. The BSA concentration was titrated within the range of 0–0.3% for individual cell lines used for the formation of the blastoids and trophospheres. A step-by-step protocol is available on Protocol Exchange (https://doi.org/10.21203/rs.3.pex-1639/v1).

Derivation of cell lines from human blastoids

Derivation experiments were performed with blastoids cultured for 96 h as described in the previous section. Blastoids were individually transferred on gelatin-coated 96-well plates with feeder layers of gamma-irradiated MEFs. Naive PSCs were derived in PXGL medium[2]. TSCs were derived in human TSC medium[20]. After 24 h of culture on feeders, blastoids attached and, within one week, colonies were formed. Derivation was considered successful after three passages after blastoid transfer. For immunofluorescence assays, naive PSCs were transferred onto Geltrex (0.5 μl cm^{-2})-coated coverslips, and TSCs were transferred onto fibronectin-coated coverslips (5 μg ml^{-1}, Sigma Aldrich, 08012).

Trophoblast organoid formation

Organoid formation was performed with blastoid-derived TSC lines. Organoids were cultured as previously described[33] with some modifications. Colonies of TSCs were dissociated into single cells using 1× trypsin at 37 °C for 5 min. After centrifugation, 200,000 cells were resuspended in 150 μl Matrigel (Corning, 356231). Droplets of 20 μl per well were

placed into a prewarmed 48-well cell culture plate and placed upside down into the incubator for 20 min. Organoids were cultured in 250 ml TOM medium (Advanced DMEM-F12, N2 supplement, B27 supplement minus vitamin a, PenStrep, *N*-acetyl-L-cysteine (1.25 mM), L-glutamine (2 mM), A83-01 (500 nM), CHIR99021 (1.5 uM), recombinant human EGF (50 ng ml^{-1}), 10% R-spondin 1 conditioned medium, recombinant human FGF2 (100 ng ml^{-1}), recombinant human HGF (50 ng ml^{-1}), PGE2 (2.5 µM). Medium was refreshed every other day. For SCT formation organoids were maintained in TOM medium until day 7.

2D trophoblast differentiations

The differentiation of blastoid derived TSCs was performed as described previously[20] with some modifications. TSC lines were adapted to Fibronectin coating (5 µg ml^{-1}, Sigma Aldrich, 08012) for at least three passages prior to the experiments. For EVT and SCT differentiation, cells were dissociated with TrypLE for 5 min at 37 °C and cells were seeded at a density of 55,000 cells per well onto 12-well plates. For SCT differentiation, the plates were precoated with 10 µg ml^{-1} fibronectin and cultured in SCT medium (DMEM/F12, supplemented with 0.1 mM 2-mercaptoethanol, 0.5% penicillin-streptomycin, 1% ITS-X supplement, 7.5 mM A83-01, 2.5 mM Y27632, 4% KnockOut Serum Replacement and 2 mM forskolin) for 3 days. For EVT differentiation, plates were precoated with Matrigel and cells were cultured in EVT medium (DMEM/F12, supplemented with 0.1 mM 2-mercaptoethanol, 0.5% penicillin-streptomycin, 1% ITS-X supplement, 2% Matrigel, 7.5 mM A83-01, 2.5 mM Y27632, 4% KnockOut Serum Replacement and 100 ng ml^{-1} NRG1). After three days, the medium was changed to EVT medium with 0.5% Matrigel and without NRG1. Cells were cultured until day 6.

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 the 8-cell stage were loaded into a 12-well dish (Vitrolife: Embryoslide Ibidi) with non-sequential culture medium (Vitrolife G2 plus) under mineral oil (Origio: Liquid Paraffin) at 37 °C in 5% O_2/6% CO_2.

Plasmid construction

The cDNA sequence of hYAP1, hYAP1(5SA) and hYAP1(5SA + S94A) were amplified from the pQCXIH-Myc-YAP, pQCXIH-Myc-YAP-5SA and pQCXIH-Myc-YAP-S94A plasmids, respectively. These YAP plasmids[34] were gifts from K. Guan (Addgene #33091, #33093 and #33094). The individual cDNA sequences were cloned into pDONR211, followed by cloning into PB-TAC-ERP2 using Gateway (invitrogen) cloning strategy. PB-TAC-ERP2[35] was a gift from K. Woltjen (Addgene plasmid #80478). Complete sequences of the resulting plasmids are available upon request.

Cell transfection in human naive PSCs

pCAG-PBase (5 µg) and PB-TAC-YAP1-ERP (5 µg) were transfected by NEPA21 electroporation (Nepa Gene) into 5×10^4 cells in single-cell suspension. Electroporated naive PSCs were plated on Geltrex (0.5 µl cm^{-2}, Thermo Fisher Science, A1413302)-coated 6-well plates with PXGL medium containing Y-27632 (10 µM). Puromycin (0.5 µg ml^{-1}, Sigma-aldrich, P7255) was added to PXGL medium from day 1 to day 3–4 to select transformed cells. pCAG-PBase was a gift from K. Woltjen.

YAP overexpression in naive PSC aggregates

The naive PSC aggregates were formed from naive H9 cell lines integrated with the doxycycline inducible cassette as described in the section above. The aggregates were cultured in PALLY medium with reduced LPA concentration (5 nM) from 0 h to 48 h along with 100 ng ml^{-1} doxycycline. Higher LPA concentrations masked the effects of the genetic overexpression of the YAP1 variants. The number of cavitated aggregates was counted at 72 h.

Single-cell RNA-seq library preparation and sequencing

To avoid over-representation of TE cells, blastoids were collected, dissociated and the cell suspension was stained using antibodies against TROP2 and PDGFRa that mark trophoblasts and primitive endoderm, respectively. For the 96 h time point, blastoids were selectively picked up from the microwell arrays before the dissociation, according to the morphological criteria described above. On the contrary, for the 24 and 65 h time points, all structures, including the ones that will not develop into a blastoid, were included. Accordingly, this non-selective picking correlated with the presence of more off-target cells. Cells were FACS-sorted into 384-well-plates containing the lysis buffer for Smart-seq2 and immediately frozen. The antibody staining was exploited in order to harvest specific numbers of TROP2$^+$, PDGFRa$^+$ and double-negative cells. The abutted FACS gates (DiVa 9.0.1) covered the whole spectrum and no blastoid cells were excluded. The H9 naive cells cultured on MEF were stained using an antibody against SUSD2, then FACS-sorted. Dead cells were excluded by DAPI staining. Smart-seq2 libraries were generated as described previously with minor optimization[36]. Maxima H Minus reverse transcriptase (3 U per reaction, Thermo Fisher Science, EP0751) was used for the cDNA synthesis. The prepared libraries were sequenced on the S1 or SP flow cell using an Illumina Novaseq instrument in 50-bp paired-end mode.

Single-cell RNA-seq data analysis

Smart-Seq transcriptome sequencing experiments were analysed using genome sequence and gene annotation from Ensembl GRCh38 release 103 as reference. For gene-expression quantification RNA-seq reads were first trimmed using trim-galore v0.6.6 and thereafter aligned to the human genome using hisat2 v2.2.1. Uniquely mapping reads in genes were quantified using htseq-count v0.13.5 with parameter -s no. TPM estimates were obtained using RSEM v1.3.3 with parameter–single-cell-prior. Further analysis was performed in R v4.0.3 with Seurat v4.0.1. Based on initial evaluation of per-cell quality control metrics and outlier identification using the median absolute deviation algorithm, cells with ≤2,000 detected genes or ≥12.5% mitochondrial gene percentage were filtered out. Only genes detected in at least five cells were retained. Count-data were log-normalized, top 3,000 highly variable were selected, and standardization of per-gene expression values across cells was performed using NormalizeData, FindVariableFeatures and ScaleData data functions in Seurat. Principal component analysis (PCA) based on the standardized highly variable features was used for linear dimension reduction, a shared nearest neighbor (SNN) graph was constructed on the dimensionally reduced data, and the graph was partitioned using a SNN modularity optimization-based clustering algorithm at a range of resolutions using RunPCA, FindNeighbors and FindClusters from Seurat with default settings. Cluster marker and marker genes between identity groups were determined with the Wilcox likelihood-ratio test (two-sided) using the FindAllMarker and FindMarkers functions with *P*-value adjustment using Bonferroni correction and followed by filtering at a adjusted *P* value cut-off of 0.05. UMAP was used for visualization.

For integration of Smart-seq experiments from multiple sources we followed the previously described procedure[10]. Published data from E-MTAB-3929 (human preimplantation embryos[37] ranging from embryonic day 3 to 7), GSE109555 (in vitro cultured blastocysts[38]) were downloaded, and data from Carnegie stage 7 embryo were kindly provided by the authors of the study[39]. All the data was preprocessed to obtain per gene read counts using the same protocol as described for blastoid cells, in the case of GSE109555 including adaptations to

accommodate UMI and CB information following the authors' instructions (https://github.com/WRui/Post_Implantation). For GSE109555 we used 1,000 cells randomly subsampled from the 3,184 high-quality single cells described in the original publication. Similar to ref. [10], we excluded cells belonging to haemogenic endothelial progenitors and erythroblasts. After evaluation of per-cell quality control metrics, and as in ref. [10], cells with >2,000 detected genes and <12.5% mitochondrial gene percentage were retained. Genes detected in at least five cells in any dataset were retained. log-normalization was performed using computeSumFactors in scran package v1.18.7, per-batch scaling normalization using multiBatchNorm in batchelor v1.6.3. Datasets were aligned using the fastMNN approach via SeuratWrappers v0.3.0 using the log-normalized batch-adjusted expression values. MNN low-dimensional coordinates were then used for clustering and visualization by UMAP. The data processing and analysis pipelines are publicly available at https://github.com/RivronLab/Human_Blastoid_Kagawa_et_al-.

Bulk RNA-seq library preparation and sequencing

Bulk RNA-seq libraries were prepared using Smart-Seq2 protocol as previously described[36]. For each sample, 50 cells were pooled together and prepared for sequencing. The libraries were then sequenced using an Illumina Novaseq 6000 with 50-bp paired end mode. For each sample, approximately 10 million reads were obtained.

Bulk RNA-seq data analysis

RNA-seq reads were first trimmed using trimgalore v0.5.0 and reads mapping to abundant sequences included in the iGenomes Ensembl GRCh38 bundle (rDNA, mitochondrial chromosome, phiX174 genome, adapter) were removed using bowtie2 v2.3.4.1 alignment. Remaining reads were analyzed using genome and gene annotation for the GRCh38/hg38 assembly obtained from Ensembl release 94. Reads were aligned to the genome using star v2.6.0c and reads in genes were counted with featureCounts (subread v1.6.2) and parameter -s 0. Differential gene-expression analysis on raw counts and variance-stabilized transformation of count data for heatmap visualization were performed using DESeq2 v1.18.1.

Culture of human trophoblast stem cells and aggregate formation

Experiments were performed using the human blastocyst-derived TSC line bTS5 provided by the laboratory of T. Arima. Cells were cultured on Laminin 511 (5 μg ml^{-1}, BioLamina, LN511) coated plates in TSC medium as previously described[20]. Aggregates of TSCs were formed as follows. Colonies were dissociated into single cells using Accutase at 37 °C for 5 min. The cells were resuspended into TSC medium containing 10 μM Y-27632, and 3.0×10^4 cells were seeded onto a microwell array imprinted into a well of a 96-well plate. The same medium[20] was refreshed daily. After 72 h, the aggregates were used for both characterization and implantation experiments.

Endometrial organoid culture

Cryopreserved human endometrial organoids were provided by the H. Baharvand laboratory (Royan Institute) within the framework of collaboration agreements. Human endometrial organoids were established from healthy human donors following the protocol described previously[24,40] with some modifications. In brief, organoids were cultured in human endometrial expansion medium composed of 10% R-spondin 1 conditioned medium (in-house made) and 10% noggin-Fc-conditioned medium[41] (in-house made), supplemented with 1× N2 supplement, 1× B27 supplement, 1× insulin-transferrin-selenium (in-house), Glutamax (1 μM), *N*-acetylcysteine (1.25 mM, Sigma-Aldrich, A7250), nicotinamide (2.5 mM, Sigma-Aldrich, 72340), EGF (50 ng ml^{-1}, Peprotech, 100-47), bFGF (2 ng ml^{-1}, Peprotech, 100-18B), HGF (10 ng ml^{-1}, Peprotech, 315-23), FGF10 (10 ng ml^{-1}, Peprotech, 100-26), A83-01 (500 nM) and SB202190 (10 μM, Tocris, 1264). Y-27632 (10 μM) was used in the first 2 days after passaging to prevent apoptosis. The medium was changed every 2 days and the organoids were passaged with TrypLE followed by mechanical dissociation every 7–9 days.

Hormonal stimulation of endometrial organoids and OFEL culture

Endometrial organoids were passaged as described in the previous section. The dissociated cells were resuspended in Matrigel supplemented with Y-27632 (10 μM), cell suspension was deposited in 48-well plates and were cultured in endometrial expansion medium for 2 days. The organoids were stimulated first with E2 (10 nM, Sigma-Aldrich, E2758) for 2 days, followed by the mixture of E2 (10 nM), P4 (1 μM, Sigma-Aldrich, P8783), and cAMP (250 μM, Biolog, B 007) with or without XAV939 (10 μM) (EPC or EPCX respectively) for 4 days. For OFEL culture, organoids were recovered from the matrigel droplets with ice-cold DMEM/F12 and mechanical pipetting. The organoids were dissociated using TrypLE and mechanically triturated to generate single cells and seeded at a density of $3–3.5 \times 10^4$ cells per well into a 96-well glass bottom plate (Cellvis, P96-1.5H-N) and cultured for 2–3 days with stimulation. For contraceptive treatment, levonorgestrel[42] (LNG) (10 μM, Sigma-Aldrich, PHR1850) was added every day to the medium after hormonal stimulation and continued until the end of the experiment.

In vitro implantation assay

Confluent OFELs were prepared for the implantation assay at least 2 h prior to the deposition of blastoids, trophospheres, naive PSCs or TSCs aggregates by washing the OFEL two times with DMEM/F12 and adding mIVC1 medium[28]. Structures were then transferred onto the OFELs using a mouth pipette under an inverted microscope. After 24–48 h, the medium was removed, the well was washed with PBS, fixed using 4% formaldehyde for 30 min at room temperature and subsequently processed for immunofluorescence staining. The percentage of attached structures was reported as the percentage of total transferred structures.

In vitro culture of human blastoids in post implantation conditions

Human blastoids were selected using a mouth pipette, washed with CMRL1066 medium and transferred into suspension culture plates or 96-well plates coated with Matrigel containing pre-equilibrated media adapted from monkey blastocyst culture[27] with minor modifications as followed. For the first day, the culture medium was CMRL1066 supplemented with 10% (v/v) FBS, 1 mM L-glutamine (Gibco), 1× N2 supplement, 1× B27 supplement, 1 mM sodium pyruvate (Sigma) and 10 μM Y27632. After 24 h, half of the medium was replaced with a new medium including 5% Matrigel. After 48h, 50% of medium was replaced with a new medium supplemented with 20% (v/v) FBS and 5% Matrigel. After 72 h, half of the medium was replaced with a new medium supplemented with 30% (v/v) KSR and 5% Matrigel. Then, half of the medium was replaced every day and blastoids were cultured for up to 6 days. Cultures were fixed for staining after 4 and 6 days of in vitro culture with 4% PFA as mentioned above.

Human pre-implantation embryos

The use of human embryos donated to research as surplus of IVF treatment was allowed by the French embryo research oversight committee: Agence de la Biomédecine, under approval numbers 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 is 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 had fulfilled their reproductive needs or the embryos had 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 the opportunity to receive counselling. No financial inducements were offered for donation. Molecular analysis of the embryos was performed in compliance with the guidelines of the embryo research oversight committee and The International Society for Stem Cell Research (ISSCR)[43].

RNA extraction, cDNA synthesis and RT–qPCR

RNA was extracted using the RNeasy mini kit (Qiagen, 74106) and cDNA synthesis was performed using the Superscript III (Invitrogen, 18080093) enzyme. qPCR reactions were performed using GoTaq qPCR Master Mix (Promega, A6001) on CFX384 Touch Real-Time PCR Detection System (Bio-rad). Quantification was performed using Microsoft Office Excel by applying the comparative cycle threshold (C_t) method. Relative expression levels were normalized to GAPDH. The primers used for the qPCR analysis are listed in Supplementary Table 3.

ELISA assay for CGβ detection

Medium from wells containing unattached or attached blastoids was collected and centrifuged to remove debris and stored at −80 °C until use. The supernatant was subject to CGβ ELISA (Abcam, ab178633), according to the manufacturer's instructions, alongside CGβ standards.

Ligand–receptor analysis

The Cellinker web-platform was used to predict putative receptor–ligand interactions between polar TE and endometrial epithelial cells. Enriched genes in polar TE along with upregulated genes in stimulated OFELs were used as the query to search ligands and receptors in the database.

Immunohistochemistry

The samples were fixed with 4% formaldehyde for 30 min at room temperature. Post fixation, formaldehyde solution was removed and the samples were washed at least three times with PBS. The samples were then permeabilized and blocked using 0.3% Triton X-100 and 10% normal donkey serum in PBS for at least 60 min. The samples were then incubated overnight at 4 °C with primary antibodies diluted in fresh blocking/permeabilization solution. The samples were washed with PBS containing 0.1% Triton X-100 (PBST) at least three times for 10 min each. The washing buffer was then replaced with Alexafluor tagged secondary antibodies (Abcam or Thermofisher scientific) along with a nuclear dye Hoechst-33342 (1:500 or 1:300 for 2D or 3D samples respectively, Life Technologies, H3570) diluted in PBST for 30 min in dark at room temperature. The samples were then washed with PBST three times for 10 min each. For human blastocysts, the samples were fixed at the B4 or B6 stage according to the grading system proposed by Gardner and Schoolcraft[44] or at B3 or B4 +72 h in vitro culture. Embryos were fixed with 4% paraformaldehyde for 10 min at room temperature and washed in PBS/BSA. Embryos were permeabilized and blocked in PBS containing 0.2% Triton-x100 and 10% FBS at room temperature for 60 min. Samples were incubated with primary antibodies overnight at 4 °C. Incubation with secondary antibodies was performed for 2 h at room temperature along with Hoechst counterstaining. The samples were mounted for imaging in PBS in the wells of glass bottom micro slides (Ibidi, 81507). The details of antibodies and their dilutions along with stainings previously performed on human blastocysts (other studies) are provided in the Supplementary Tables 4, 5. EdU staining was done using Click-iT EdU Alexa Fluor 647 Imaging Kit (Thermo Scientific, C10640) following the manufacturer's instructions.

Microscopy and image analysis

The phase-contrast images were acquired using Thermo Fisher scientific EVOS cell imaging system and inverted wide-field microscope Axio VertA1. The number of blastoids or cavitated structures were counted manually for each well. After 96 h, a blastoid is defined based on the morphological parameters as described in previous sections. The fluorescent images and time-lapse images were acquired using Olympus IX83 microscope with Yokogawa W1 spinning disk (Software: CellSense 2.3; camera: Hamamatsu Orca Flash 4.0) or Nikon Eclipse Ti E inverted microscope, equipped with a Yokogawa W1 spinning disc (Software: Visiview 4.5.0.7 ; camera: Andor Ixon Ultra 888 EMCCD). The confocal images were analysed and display images were exported using FIJI 1.53k or Bitplane Imaris 9.7.0 softwares. For cell counting, Bitplane Imaris software was used. Cell count parameters were set for size and fluorescence strength of voxels and then overall cell count data was obtained for each image using the Imaris spot function. Note that large cavities in blastoids increase the depth of the imaging field and cause poor signal from deeply located cells. Therefore, our counting data in Figs. 1h, 3g could be underrepresented values, particularly in the case of trophectoderm cells. The quantification of the percentage of blastoids forming the NR2F2 axis was done manually. To do so, blastoids stained to detect NR2F2 expression were imaged using a confocal spinning disk microscope. The images were projected using a 3D-project function in FIJI. The blastoid was classified to have an axis when NR2F2 expression was restricted to its polar half with no expression or lower level of expression in the mural half. The inverted pattern of NR2F2 expression was classified as an invert axis. The blastoids with NR2F2 expression on their both polar and mural halves were classified to have no axis. Confocal immunofluorescence images of human blastocysts were acquired with a Nikon confocal microscope and a 20× mim or 25× silicon objective. Optical sections of 1 µm-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.

Statistics and reproducibility

All the experiments were performed at least in three biological replicates unless specifically described in the Methods and the figure legends. Statistical analyses were performed using Graphpad prism 8.1.1 (330).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Single-cell RNA-seq and bulk RNA-seq data for human blastoids used in this study were deposited at the Gene Expression Omnibus under the accession number GSE177689. Source data are provided with this paper.

29. Guo, G. et al. Epigenetic resetting of human pluripotency. *Development* **144**, 2748–2763 (2017).
30. Rivron, N. C. et al. Tissue deformation spatially modulates VEGF signaling and angiogenesis. *Proc. Natl. Acad. Sci. USA* **109**, 6886–6891 (2012).
31. Vrij, E. J. et al. 3D high throughput screening and profiling of embryoid bodies in thermoformed microwell plates. *Lab Chip* **16**, 734–742 (2016).
32. Yu, F.-X. et al. Regulation of the Hippo–YAP pathway by G-protein-coupled receptor signaling. *Cell* **150**, 780–791 (2012).
33. Turco, M. Y. et al. Trophoblast organoids as a model for maternal-fetal interactions during human placentation. *Nature* **564**, 263–267 (2018).
34. Zhao, B. et al. Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. *Genes Dev.* **21**, 2747–2761 (2007).
35. Kim, S.-I. et al. Inducible transgene expression in human iPS cells using versatile all-in-one piggyBac transposons. *Methods Mol. Biol.* **1357**, 111–131 (2016).
36. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).

37. Petropoulos, S. et al. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* **165**, 1012–1026 (2016).
38. Zhou, F. et al. Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature* **572**, 660–664 (2019).
39. Tyser, R. C. V. et al. Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature* **600**, 285–289 (2021).
40. Turco, M. Y. et al. Long-term, hormone-responsive organoid cultures of human endometrium in a chemically defined medium. *Nat. Cell Biol.* **19**, 568–577 (2017).
41. Heijmans, J. et al. ER stress causes rapid loss of intestinal epithelial stemness through activation of the unfolded protein response. *Cell Rep.* **3**, 1128–1139 (2013).
42. Matsuo, M. et al. Levonorgestrel inhibits embryo attachment by eliminating uterine induction of leukemia inhibitory factor. *Endocrinology* **161**, bqz005 (2020).
43. Kimmelman, J. et al. New ISSCR guidelines: clinical translation of stem cell research. *Lancet* **387**, 1979–1981 (2016).
44. Gardner, D. K., Lane, M., Stevens, J., Schlenker, T. & Schoolcraft, W. B. Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. *Fertil. Steril.* **73**, 1155–1158 (2000).
45. Mischler, A. et al. Two distinct trophectoderm lineage stem cells from human pluripotent stem cells. *J. Biol. Chem.* **296**, 100386 (2021).
46. Jeschke, U. et al. The human endometrium expresses the glycoprotein mucin-1 and shows positive correlation for Thomsen–Friedenreich epitope expression and galectin-1 binding. *J. Histochem. Cytochem.* **57**, 871–881 (2009).
47. Zhang, Y. et al. Cellinker: a platform of ligand–receptor interactions for intercellular communication analysis. *Bioinformatics* **37**, 2025–2032 (2021).

Acknowledgements This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-Co grant agreement no.101002317 'BLASTOID: a discovery platform for early human embryogenesis'). H.H.K. is supported by the Austrian Science Fund (FWF), Lise Meitner Programme M3131-B. This project has also received funding from the ANR 'BOOSTIVF'. L.D. thanks the iPSCDTC and MicroPICell core facilities. We thank Y. Takashima for sharing the H9 and H9-GFP cell lines; A. Smith, P. Andrews and G. Guo for sharing the HNES1, Shef6, niPSC 16.2b and cR-NCRM2 cell lines; H. Baharvand for sharing the endometrial organoids; K. Woltjen for sharing the PB-TAC-ERP2 and pCAG-PBase plasmids; K. Guan for sharing pQCXIH-Myc-YAP, pQCXIH-Myc-YAP-5SA, pQCXIH-Myc-YAP-S94A plasmids; J. M. Brickman for sharing the RNA isolated from PrE differentiated cells and nEND cells; S. Srinivas for sharing the single-cell RNA-seq data of peri-gastrulation embryo; A. Bykov and L. Cochella for technical assistance for SMARTSeq2 library preparation; and the NGS, Biooptic and Stem Cell facility at IMBA for critical assistance.

Author contributions H.K., A.J., H.H.K. and N.R. conceived the study; N.R. supervised the project; H.K., A.J., H.H.K., T.M.S. and N.R. designed the blastoid experiments; H.K., A.J., H.H.K., T.M.S. and Y.S.o.R. performed blastoid experiments; G.S., G.C. and M.N. performed the bioinformatic analysis of single-cell RNA-seq datasets; J.L., S.L. and T.F. managed human embryos donated for research in Nantes; A.B., J.L., S.L. and G.C. performed experiments on human embryos in Nantes; L.D. supervised experiments on human embryos in Nantes; H.H.K., N.M., H.V. and N.R. designed the experiments with endometrial organoids; N.R. hosted N.M. and G.C. in his laboratory; H.K., A.J., H.H.K., T.M.S., Y.S.o.R., G.C., A.B., N.M. and N.R. analysed data; N.R. wrote the manuscript with help from all of the authors.

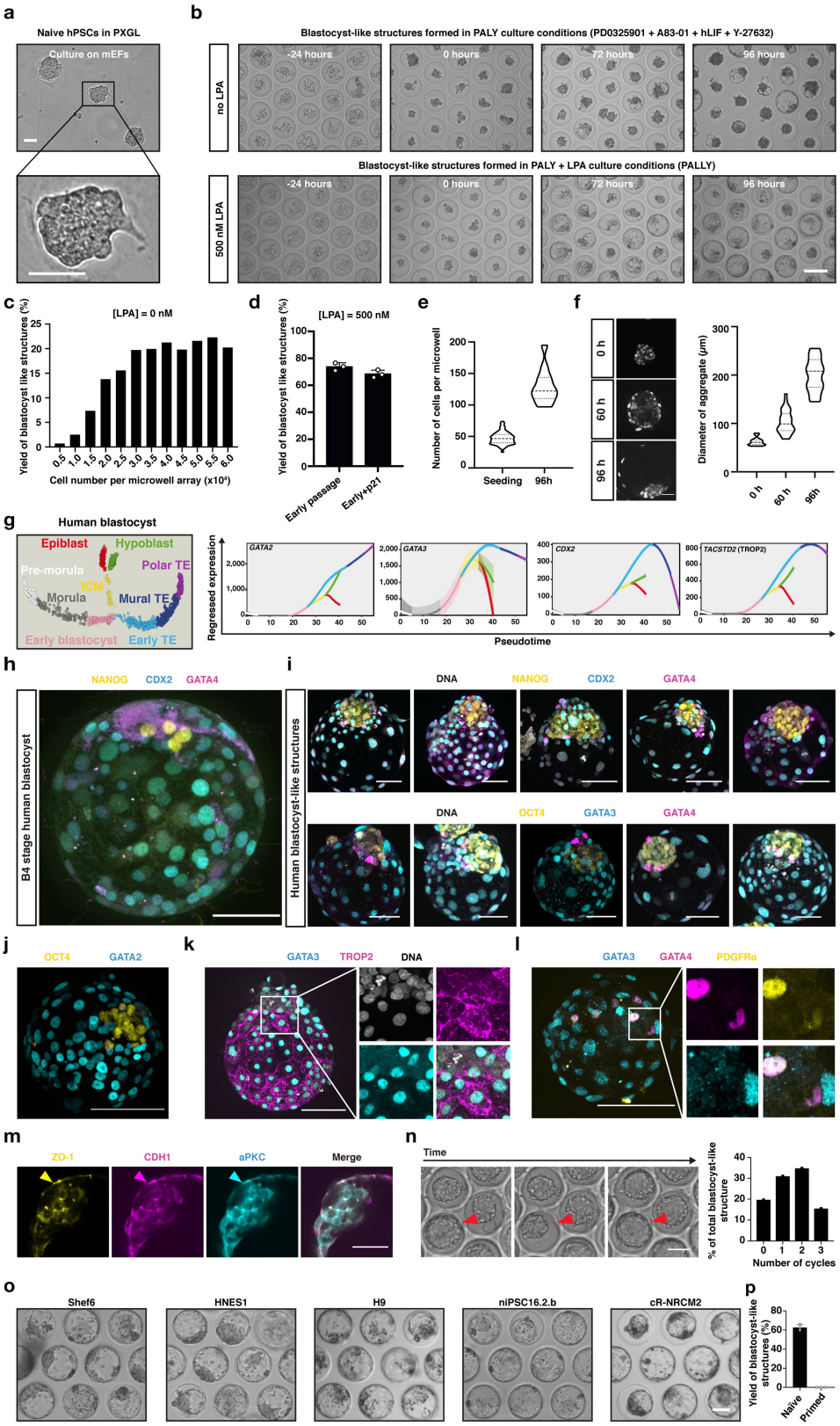
Competing interests The Institute for Molecular Biotechnology, Austrian Academy of Sciences has filed patent application EP21151455.9 describing the protocols for human blastoid formation and for the blastoid–endometrium interaction assay. H.K., A.J., H.H.K. and N.R. are the inventors on this patent. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-04267-8.

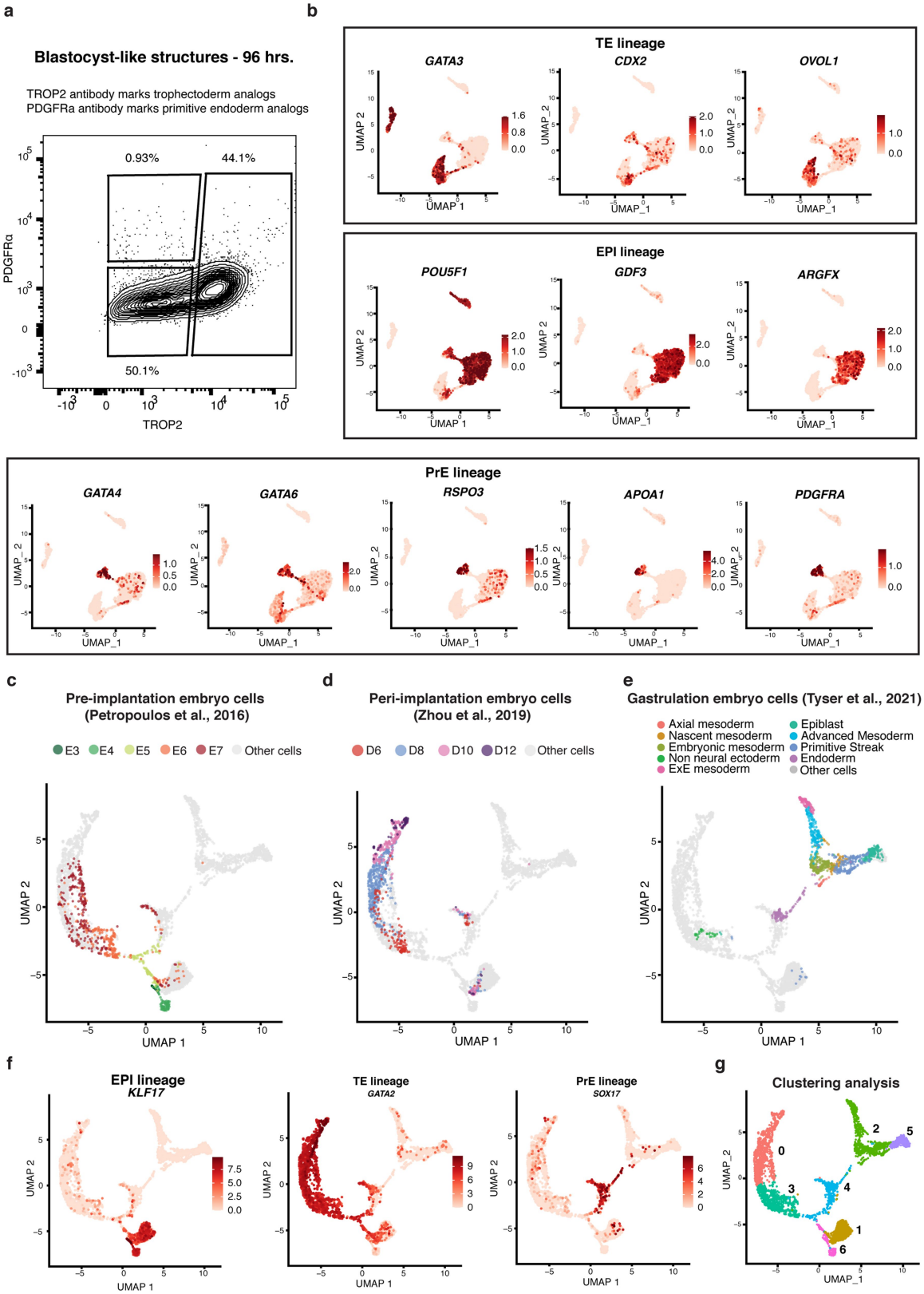
Correspondence and requests for materials should be addressed to Nicolas Rivron.

Peer review information *Nature* thanks Jan Brosens, Jianping Fu, Insoo Hyun and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



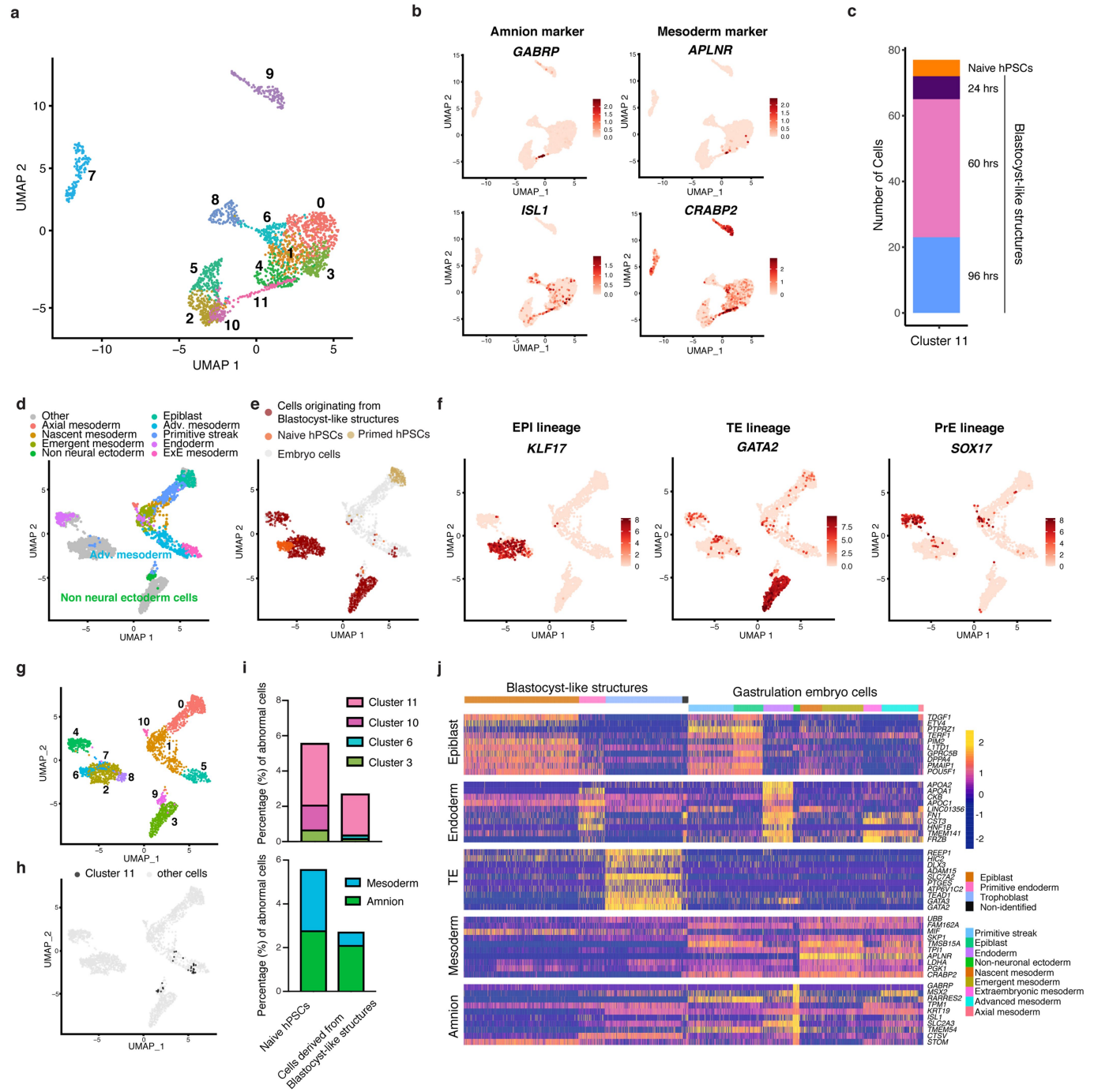
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Naive hPSCs form human blastocyst-like structures comprising analogs of the three founding lineages. a. Phase contrast images of naive hPSCs cultured in PXGL medium and on MEF feeder layers. Scale bar: 50 µm. **b**. Time course phase contrast images of naive hPSCs aggregates cultured within microwell arrays either without LPA (PALY medium, **top**) or with 500 nM LPA (PALLY medium, **bottom**). Scale bar: 200 µm. **c**. Quantification of the effect of the initial cell numbers per microwell array on the yield of blastocyst-like structures. n=1 microwell arrays. **d**. Quantification of the effect of serial passaging of naive hPSCs on the yield of blastocyst-like structures. n=3 microwell arrays. mean± S.D. **e**. Quantification of the cell numbers per microwell at the time of seeding and in blastocyst-like structures at 96 h when cells are seeded at 3.0×10^4 cells per microwell array. n=190 microwells (seeding) and n=12 blastocyst-like structures (96 hrs.). **f**. Fluorescence staining of DNA using Hoechst in representative naive hPSCs aggregates over the course of formation of blastocyst-like structures (96 h, **left**). Scale bar: 50 µm. Measurement of the distributed diameters of the structures over the course of formation of blastocyst-like structures (**right**). n=15, 31 and 11 for 0, 60 and 96 h, respectively. **g**. Pseudotime analysis of human pre-implantation development showing the expression of the TE markers *GATA2, GATA3, CDX2 and TACSTD2*. Gene expression analysis was performed by using the public data analysis tool (https://bird2cluster.univ-nantes.fr/demo/PseudoTimeUI/). **h**. Immunofluorescence stainings for EPI marker NANOG (Yellow), TE marker CDX2 (Cyan) and primitive endoderm marker GATA4 (Magenta) in a representative B4-stage human blastocyst. Scale bar: 50 µm. **i**. Immunofluorescence stainings for the EPI markers (Yellow) NANOG (**top**) and OCT4 (**bottom**); the TE markers (Cyan) CDX2 (**top**) and GATA3 (**bottom**); and the primitive endoderm marker (Magenta) GATA4 in five representative blastocyst-like structures. Counterstain with Hoechst (Grey) marking DNA. Scale bar: 50 µm. **j**. Immunofluorescence staining for EPI marker OCT4 (yellow) and TE marker GATA2 (Cyan) in blastocyst-like structures. Scale bar: 100 µm. **k**. Immunofluorescence staining for TE markers GATA3 (Cyan) and TROP2 (Magenta) in blastocyst-like structures. Scale bar: 100 µm. **l**. Immunofluorescence staining for TE markers GATA3 (Cyan) and PrE marker GATA4 (Magenta) and PDGFRa (Yellow) in blastocyst-like structures. Scale bar: 100 µm. **m**. Single optical section of immunofluorescence staining image for the tight junction molecule ZO-1 (Yellow), the adherence junction molecule CDH1 (Magenta), and the apical domain molecule aPKC (Cyan) in a representative human blastocyst-like structures. Scale bars: 50 µm. **n**. Representative time points from a timelapse image of naive cell aggregates, cavitating into blastocyst-like structures while showing cycles of cavity inflation and deflation (**left**) - quantification of blastocyst-like structures showing distinct frequencies of inflation and deflation (**right**). n=1 microwell arrays. Scale bar: 100 µm. **o**. Phase contrast images of representative areas of microwell arrays showing blastocyst-like structures formed from different naive hPSCs and hiPSCs lines. n>3. Scale bar: 100 µm. **p**. Quantification of the yield of blastocyst-like structures obtained from naive and primed H9 hPSCs. n=3 microwell arrays. mean± S.D.

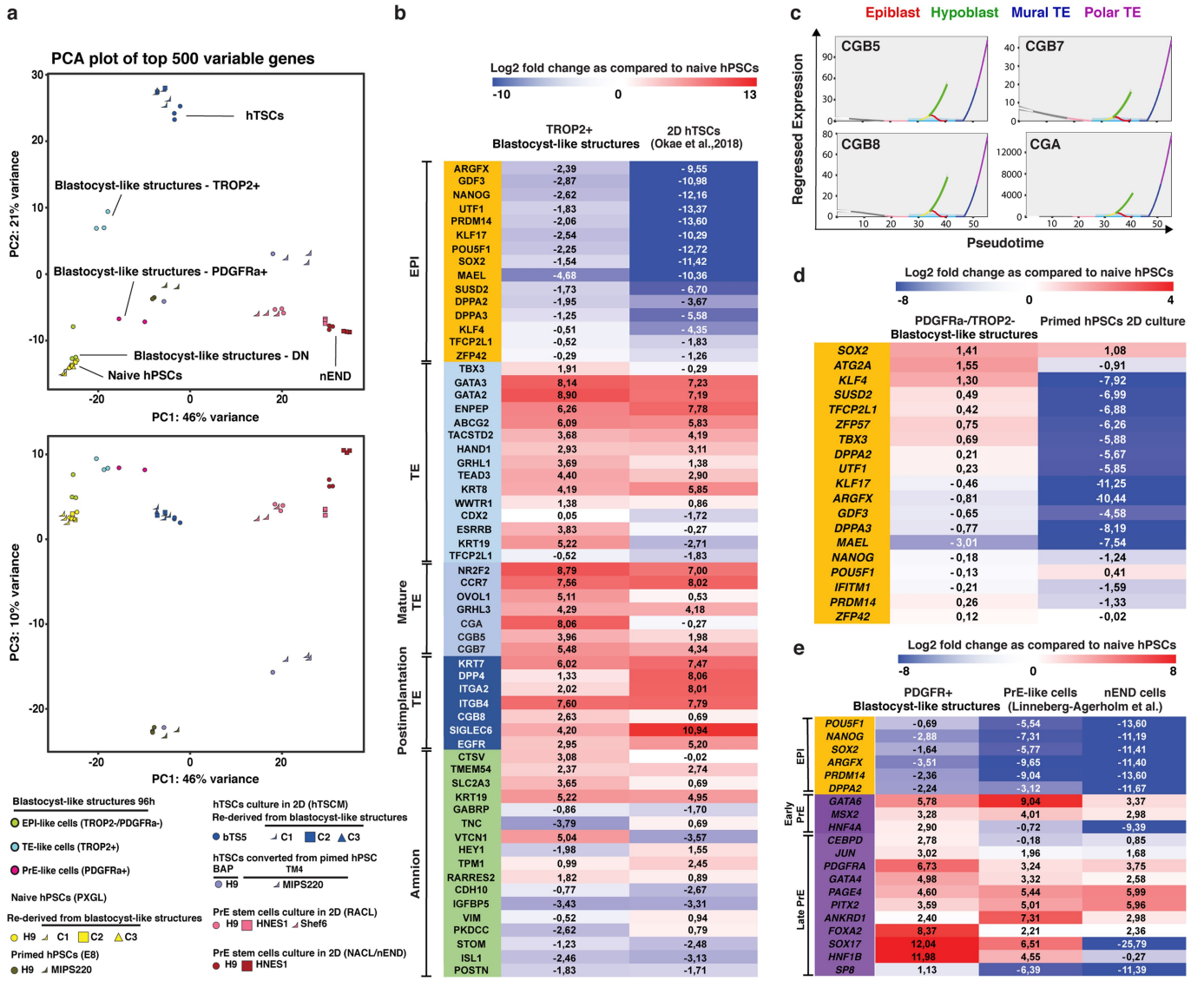


Extended Data Fig. 2 | See next page for caption.

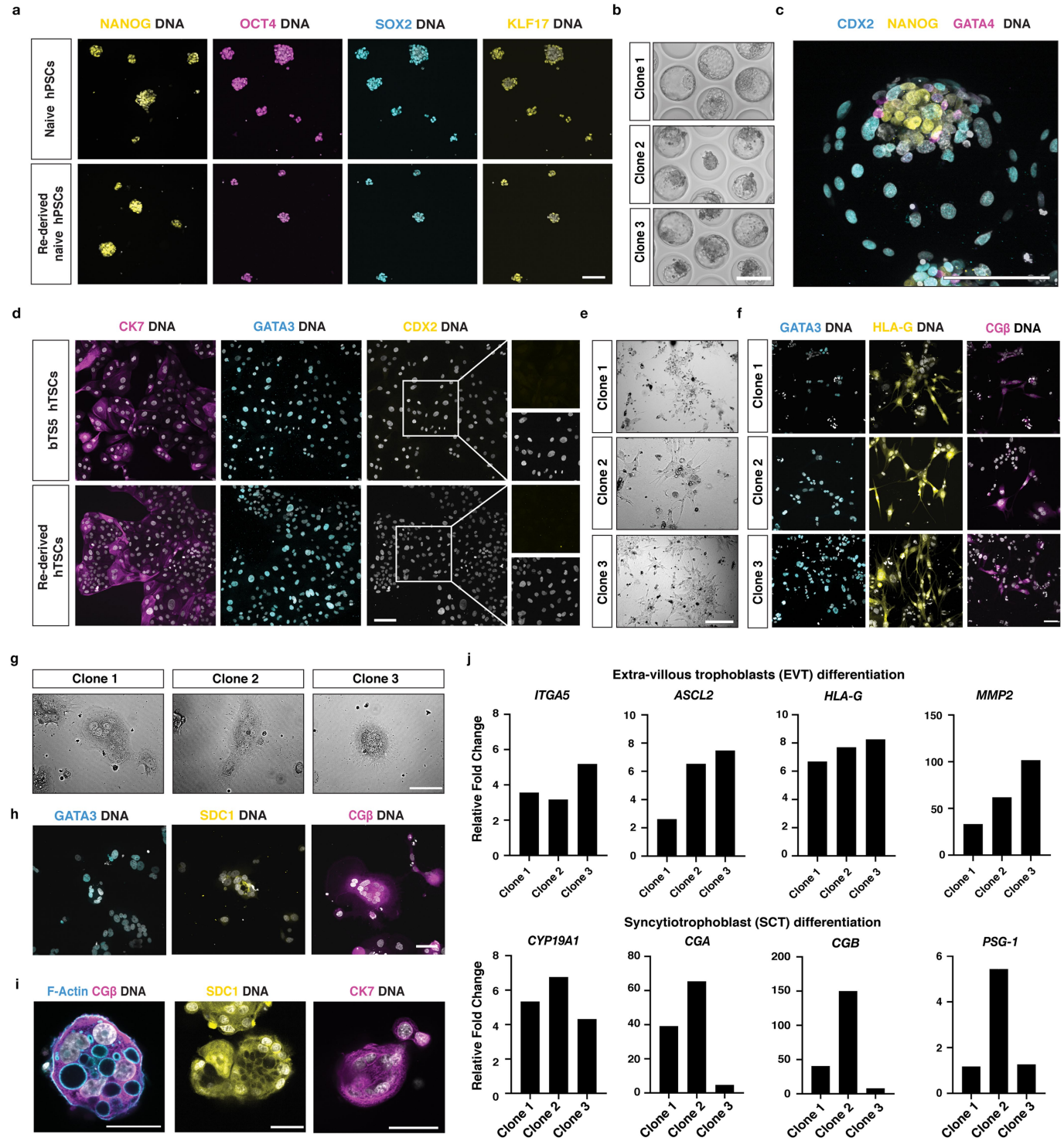
Extended Data Fig. 2 | Human blastocyst-like structures form analogs of pre-implantation lineages. a. Flow cytometry analysis plot of cells isolated from blastocyst-like structures and stained for lineage-specific surface markers PDGFRa (PrE) and TROP2 (TE). The gates were used to sort analogs of EPI (double negative), TE ($TROP2^{high}$) and PrE ($PDGFR\alpha^{high}$) to subsequently process for single cell RNA sequencing. Note that the gates did not exclude any cells. This analysis was performed to correlate RNA measures, while ensuring a representation of all cell types. **b**. UMAPs of the transcriptome of single cells isolated from blastocyst-like structures and displaying the expression levels of genes specific for each of the three blastocyst lineages (TE - Trophectoderm, EPI - Epiblast, and PrE- Primitive endoderm). **c-g**. UMAPs of single cells isolated from both blastocyst-like structures and from embryos ranging from E3 to E19. **c**. Coloration of cells originating from *In Vitro Fertilization* (IVF) embryos isolated on day 3 (E3) to day 7 (E7). This period comprises only pre-implantation stage embryos. **d**. Coloration of cells originating from IVF embryos isolated on day 6 (E6) to day 12 (E12). These blastocysts (E6) were cultured *in vitro*. Note that this annotation reflects the number of days in culture rather than the developmental stages. **e**. Coloration of cells originating from gastrulation-stage embryo isolated on day 17 (E17) to 19 (E19). **f**. The expression levels of genes specific for each of the three blastocyst lineages (EPI, TE, and PrE). **g**. Coloration of cells displaying their unsupervised cluster affiliation.



Extended Data Fig. 3 | Measurement of generation of off-target cells in human blastocyst-like structures and naive human pluripotent stem cells. **a, b.** UMAP of clusters formed from cells isolated from blastocyst-like structures (high-resolution clustering of 1, x50 as compared to Fig. 2b) (**a**) and displaying the expression levels of genes specific for amnion lineage (**b**). **c.** Origin of the cells composing cluster 11. **d-h.** UMAPs of naive hPSCs, primed hPSCs, cells isolated from blastocyst-like structures and cells isolated from a CS7 staged human embryo. **d.** Coloration of embryo cells based on previously proposed annotations[39]. **e.** Coloration of stem cells based on their origin. **f.** Display of the expression levels of genes specific for each of the three blastocyst lineages (EPI - Epiblast, TE - Trophectoderm, and PrE- Primitive endoderm). **g.** Coloration of individual cells based on their unsupervised cluster affiliation. **h.** Coloration of the cells previously identified as cluster 11 (see **a, b**). **i.** Quantification of the percentage of cells identified as abnormal based on the location in the UMAP in **h** (**top**) and on the cells annotations (**bottom**) for both naive hPSCs (**left**) and cells isolated from blastocyst-like structures (**right**). Similar results were obtained based on the location in the UMAP in (Extended Data Fig. 2c-e). **j.** Heatmap of previously proposed markers of different lineages differentially expressed in cells from blastocyst-like structures and gastrulation-stage embryo[10].

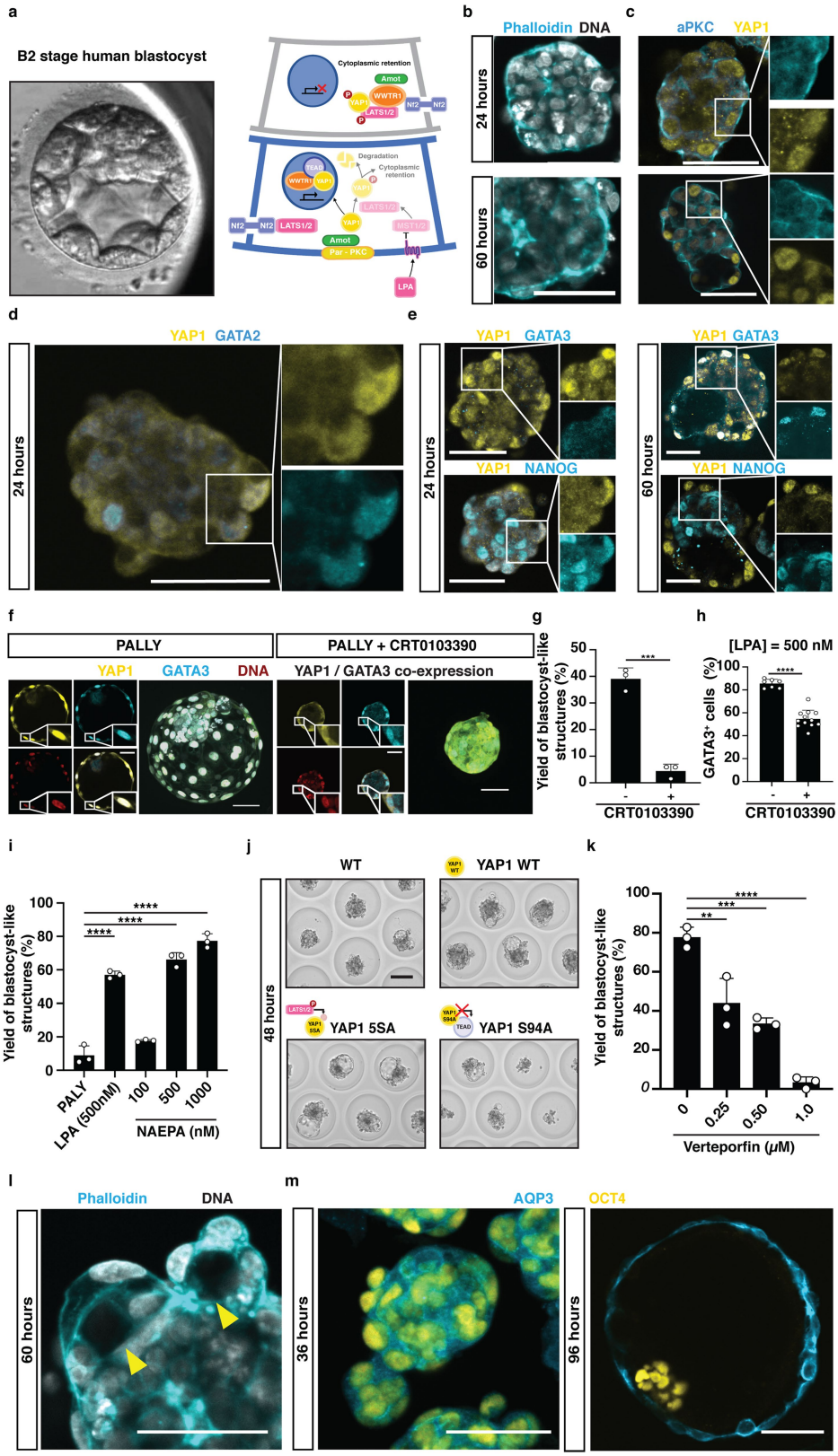


Extended Data Fig. 4 | Cells in human blastocyst-like structures are transcriptionally similar to pre-implantation lineages. a. Principal component analysis (PCA) plot with PC1 vs PC2 (**top**) or PC1 vs PC3 (**bottom**) computed with top 500 variable gene in the bulk transcriptome of individual lineages of blastocyst-like structures (EPI, TE and PrE); stem cell lines: naive and primed hPSCs; hTSCs: blastocyst derived hTSCs (bTS5)[20], primed hPSC derived hTSCs (BAP[14] and TM4 protocols[45]; PrE like stem cell lines (RACL or nEND cells[22]); naive PSC and TSCs rederived from blastocyst-like structures (see methods). **b**. Heatmap of key blastocyst and post-implantation lineage markers differentially expressed between TE analogs (TROP2^{+}) of the blastocyst-like structures and hTSCs in their bulk transcriptome. **c**. Pseudotime analysis of human mature TE markers *CGB5, CGB7, CGB8 and CGA*. Gene expression analysis was performed by using the public data analysis tool (https://bird2cluster.univ-nantes.fr/demo/PseudoTimeUI/). **d**. Heatmap of key pluripotency related genes differentially expressed between EPI analogs (PDGFR^{-}/ TROP2^{-}) in the blastocyst-like structures and primed hPSCs **e**. Heatmap of key pluripotency related genes or PrE markers differentially expressed between PrE analogs (PDGFRα^{+}) in the blastocyst-like structures, naive PSC derived PrE-like cells and nEND cells.



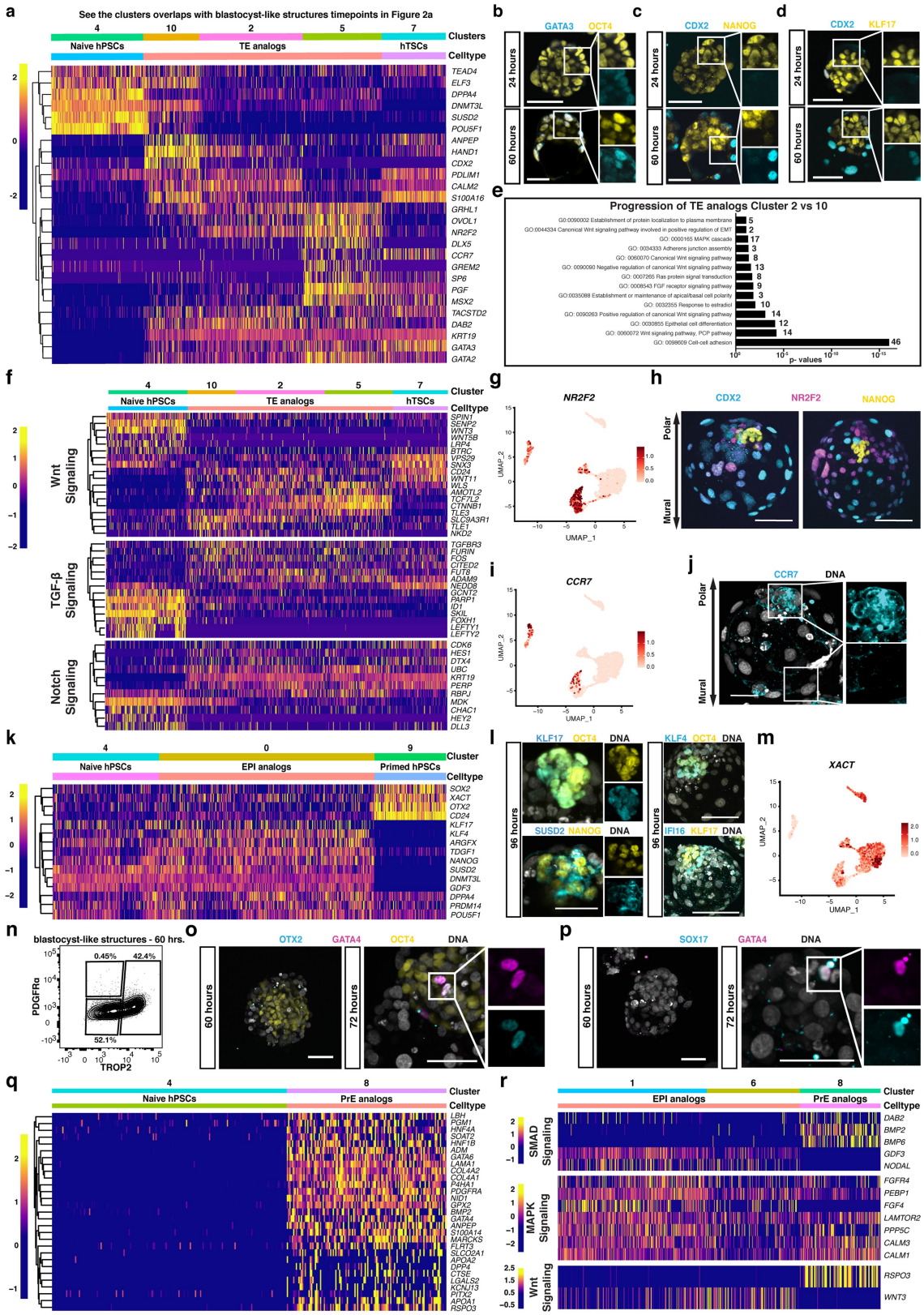
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Human blastocyst-like structures are permissive for derivation of stem cell lines. a. Immunofluorescence staining for pluripotency factors NANOG (Yellow), OCT4 (Magenta), SOX2 (Cyan) and for naive pluripotency factor KLF17 (Yellow) in naive hPSC controls (**top**) and naive hPSCs derived from blastocyst-like structures (**bottom**). Scale bar: 100 µm. **b**. Phase contrast images of blastocyst-like structures on microwell array formed from three rederived naive hPSC lines. Scale bar: 200 µm. **c**. Immunofluorescence stainings for EPI marker (NANOG), TE marker (CDX2) and primitive endoderm marker (GATA4) in representative second-generation blastocyst-like structures. Scale bar: 100 µm. **d**. Immunofluorescence staining for GATA3 (Cyan), post-implantation trophoblast marker CK7 (Magenta) and CDX2 (Yellow) in bTS5 hTSC (**top**) and hTSCs derived from blastocyst-like structures (**bottom**). Scale bar: 100 µm. **e**. Phase contrast images of day 6 EVT differentiations from three hTSC lines derived from blastocyst-like structures. Scale bar: 150 µm. **f**. Immunofluorescence stainings of trophoblast markers GATA3 (Cyan) and EVT marker HLA-G (Yellow) and CGβ (Magenta) of day 6 EVT analogs from three hTSC lines, derived from blastocyst-like structures. Scale bar: 100 µm. **g**. Phase contrast images of day 3 SCT analogs differentiated from three hTSC lines derived from blastocyst-like structures. Scale bar: 150 µm. **h**. Immunofluorescence stainings for trophoblast markers GATA3 (Cyan) and SCT marker SDC1 (Yellow) and CGβ (Magenta) of day 3 SCT analogs formed from hTSC line derived from blastocyst-like structure (Clone 1). Scale bar: 100 µm. **i**. Immunofluorescence stainings for CGβ (Magenta) counterstained with Phalloidin (Cyan) and Hoechst marking Actin and DNA respectively (**left**), SDC (Yellow), CK7 (Magenta) (**right**) counterstained with Hoechst marking DNA of day 6 trophoblast organoids formed from hTSC lines derived from blastocyst-like structures (Clone 1). Scale bar: 50 µm. **j**. Relative expression levels, as measured by RT-PCR, of day 6 EVT (**top**) and day 3 SCT analogs (**bottom**) with respective undifferentiated hTSCs lines derived from blastocyst-like structures. Expression levels were normalized to expression of *GAPDH*. n=1 biological replicate for three individual clones.



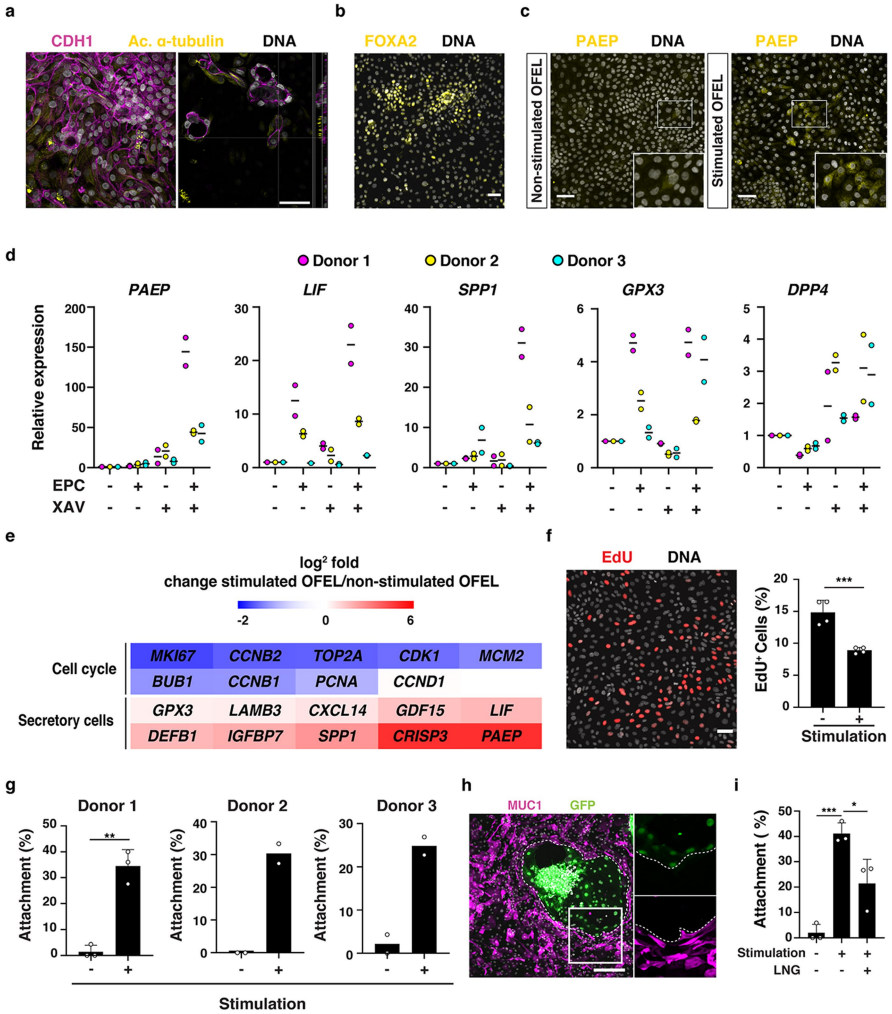
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | The development of the human trophectoderm analog depends on aPKC and Hippo elements. a. A frame from time-lapse microscopy of B2 stage human blastocyst (**left**). Schematic showing the differential Hippo activity in inner and outer cells of developing blastocyst and the molecular regulators of the Hippo signalling pathway (**right**). **b**. Phalloidin fluorescence (Cyan) stainings for F-actin in naive hPSCs aggregates cultured in PALLY medium for 24 h (**top**) and 60 h (**bottom**). Counterstain with Hoechst marking DNA. Scale bar: 50 µm. **c**. Immunofluorescence stainings for aPKC (Cyan) and YAP1 (Yellow) in aggregates of naive hPSCs cultured in PALLY medium for 24 h (**top**) and 60 h (**bottom**). Counterstain with Hoechst marking DNA. Scale bar: 50 µm. **d**. Immunofluorescence stainings for YAP1 (Yellow) with GATA2 (Cyan) in aggregates of naive hPSCs cultured in PALLY medium for 24 h. Scale bar: 50 µm. **e**. Immunofluorescence stainings for YAP1 (Yellow) and GATA3 (Cyan) (**top**) and YAP1 (Yellow) and NANOG (Cyan) (**bottom**) in naive hPSCs aggregates cultured in PALLY medium for 24 h (**left**) and 60 h (**right**). Counterstain with Hoechst marking DNA. Scale bar: 50 µm. **f**. Immunofluorescence staining for YAP1 (Yellow) and GATA3 (Cyan) in blastocyst-like structures cultured without (**top**) or with an aPKC inhibitor (2 µM CRT0103390, **bottom**). Counterstain with Hoechst marking DNA (Red). Insets: Individual and merge channels of YAP1 and GATA3 for a single optical section as well as maximum intensity projection of all the optical sections. Scale bar: 50 µm. **g**. Quantification of the yield of blastocyst-like structures upon the culture in PALLY medium or PALLY medium complemented with an aPKC inhibitor (2 µM CRT0103390). n=3 independent microwell arrays; mean± S.D.; Two tailed unpaired t-test. *** is P=0.0002. **h**. Quantification of the percentage of $GATA3^+$ cells in structures cultured in PALLY medium or in PALLY medium complemented with a aPKC inhibitor (2 µM CRT0103390). n=7 blastocyst-like structures for the group cultured in PALLY medium and n=12 aggregates for the group cultured in PALLY medium complemented with CRT0103390. Representative results from three independent experiments. Mean± S.D.; Two-tailed unpaired t-test. **** is P=1.79e-08. **i**. Quantification of the dose dependent effect of the LPA receptor agonist NAEPA on the yield of blastocyst-like structures. The PALY medium (thus without LPA) was complemented with NAEPA. n=3 independent microwell arrays; mean± S.D.; one-way Anova and Tukey's multiple comparisons test. **** is P<0.0001. **j**. Phase contrast images of representative naive hPSC aggregates cultured in PALLY medium complemented with Doxycycline (100 ng/ml) for 72 h and overexpressing different variants of YAP1. The naive hPSCs aggregates were cultured with an adjusted PALLY medium characterized by a reduced LPA concentration (5 nM). Scale bar: 100 µm. **k**. Measurement of the effect of Verteporfin (suppressor of the YAP1–TEAD complex) on the yield of blastocyst-like structures. n=3 independent microwell arrays; mean± S.D.; one-way Anova and and Dunnett's multiple comparisons test. ** is p=0.0010, *** is p=0.00019, **** is P<0.0001. **l**. Phalloidin fluorescence staining of F-actin (Cyan) in naive hPSCs aggregates cultured in PALLY medium for 60 h. Counterstain with Hoechst marking DNA. Yellow arrows: Formation of cavities. Scale bar: 50 µm. **m**. Immunofluorescence stainings for Aquaporin3 (AQP3, Cyan) and OCT4 (Yellow) in naive hPSCs aggregates cultured in PALLY medium for 36 (**left**) or 96 h (**right**, blastocyst-like structure stage). Scale bar: 50 µm.

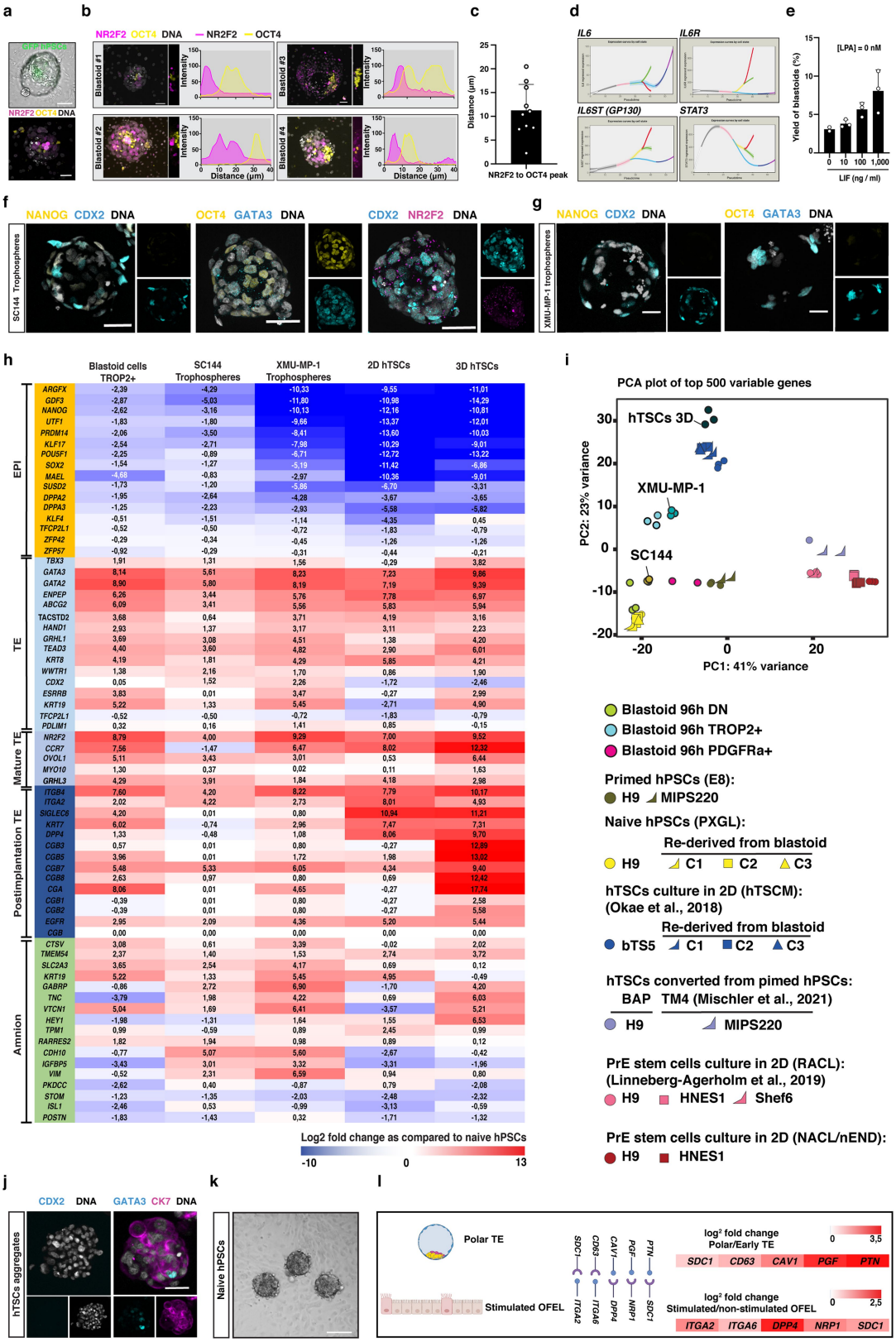


Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Blastocyst-like structures recapitulate the sequential specification of lineages occurring during blastocyst development. **a**. Heatmap of the average count values in the expression of TE genes upon formation of the blastocyst-like structures TE analogs. **b**-**d**. Immunofluorescence stainings for GATA3 (Cyan)) and OCT4 (Yellow) **(b)** or CDX2 (Cyan) and NANOG (Yellow) **(c)** or CDX2 (Cyan) and KLF17 (Yellow) **(d)** in naive hPSCs aggregates cultured in PALLY medium for 24 h **(top)** or 60 h **(bottom)**. Scale bar: 50 µm. **e**. Gene ontology terms associated with the genes differentially regulated in the late TE analog of blastocyst-like structures (cluster 10) as compared to the early TE (cluster 2). **f**. Heatmap of average count values of Wnt, TGF-β and Notch signaling-associated genes in cells from cluster 4 (naive hPSCs), 10, 2 and 5 (TE analogs) and 7 (TSC). **g**. UMAPs of single cells isolated from blastocyst-like structures and displaying the expression levels of polar trophectoderm specific gene: NR2F2. **h**. Immunofluorescence staining for CDX2 (Cyan), NR2F2 (Magenta) and NANOG (Yellow) in blastocyst-like structures. Scale bar: 100 µm. **i**. UMAPs of single cells isolated from blastocyst-like structures and displaying the expression levels of polar trophectoderm specific gene: CCR7. **j**. Immunofluorescence stainings for CCR7 (Cyan) in a blastocyst-like structures. Counterstain with Hoechst marking DNA. Scale bar: 50 µm. **k**. Heatmap of average count values of top differentially regulated genes in cells from cluster 4 (naive hPSCs), 0 (EPI analogs) and 9 (primed hPSCs). **l**. Immunofluorescence staining for KLF17 (Cyan) and OCT4 (Yellow) or KLF4 (Cyan) and OCT4 (Yellow) **(top)** and SUSD2 (Cyan) and NANOG (Yellow) or IFI16 (Cyan) and KLF17 (Yellow) **(bottom)** in blastocyst-like structures. Counterstain with Hoechst marking DNA. Scale bar: 100 µm. **m**. UMAPs of single cells isolated from blastocyst-like structures and displaying the expression levels of X chromosome activation-related gene-*XACT*. **n**. Flow cytometry analysis plot of cells isolated from blastocyst-like structures cultured in PALLY medium for 60 h and stained for lineage-specific surface markers PDGFRa (PrE) and TROP2 (TE)**.** **o, p**. Immunofluorescence stainings for OTX2 (Cyan), GATA4 (Magenta) and OCT4 (Yellow) (**o)** and SOX17 (Cyan) and GATA4 (Magenta) (**p**) in naive hPSCs aggregates cultured in PALLY medium for 60 h. Counterstain with Hoechst marking DNA. Scale bar: 50 µm. **q**. Heatmap of the average count values in the expression of PrE genes upon formation of the blastocyst-like structures PrE analogs. **r**. Heatmap of average count values of SMAD, MAPK and Wnt signaling-associated genes in cells from cluster 1, 6 (EPI analogs) and 8 (PrE analogs).

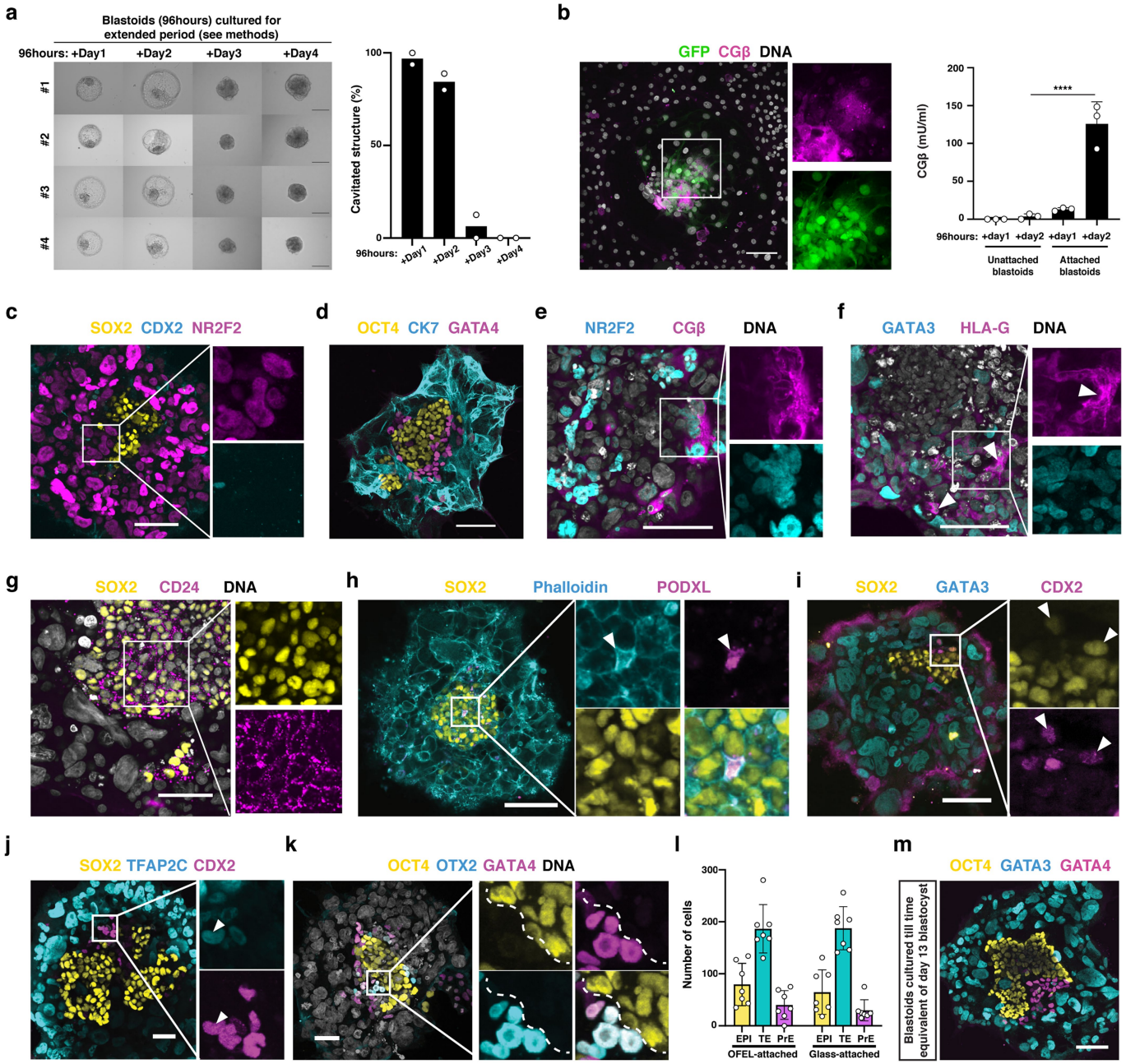


Extended Data Fig. 8 | Human blastoids recapitulate aspects of implantation. a. Immunofluorescence stainings for CDH1 (Magenta) and a ciliated cell marker acetylated α-tubulin (Yellow) in OFELs (**left**). Y-Z plane shows the apical location of the cilia (**right**). Scale bar: 50 µm. **b**. Immunofluorescence staining for FOXA2 (Yellow) marking the endometrial glandular cells in OFELs. Scale bar: 50 µm. **c**. Immunofluorescence staining for PAEP (Yellow) in non-stimulated (**left**) and stimulated (**right**) OFELs. **d**. qRT-PCR measurement of the expression levels of window-of-implantation markers in OFELs cultured with different media. Ctrl: Control medium, E: Estradiol, P: Progesterone, C: cAMP, X: XAV-939. Expression levels were normalized relative to the housekeeping gene GAPDH and the control condition. n = 2 independent experiments. The colors depict the data from 3 different donors. **e**. Heatmap of key cell cycle and secretory epithelial genes differentially expressed between stimulated and non-stimulated OFELs in bulk transcriptome. **f**. Staining for incorporated EdU (Red) reflective of cell proliferation in a stimulated OFEL (**left**). Scale bar: 50 µm. Quantification of the number of EdU$^+$ cells in non-stimulated and stimulated OFELs (**right**). Counterstain with Hoechst marking DNA. n=4 independent experiments. mean± S.D.; Unpaired two-tailed *t*-test, *** is P = 0.0009. **g**. Quantification of blastoid attachment onto OFELs prepared using endometrial organoids from 3 different donors. n=3 independent experiments for donor 1 and n=2 independent experiments for donor 2 and 3; mean± S.D.; Unpaired two-tailed *t*-test, ** is P =0.0011. **h**. Immunofluorescence stainings for MUC1 (Magenta), a glycoprotein that highly expresses at the luminal epithelial surface of endometrium in the receptive phase[46], with an attached GFP+ blastoid (48 h after deposition onto an OFEL). Dashed lines indicate the area that trophoblast cells repelled endometrial cells. Scale bar: 200 µm **i**. Quantification of blastoid attachment onto non-stimulated, stimulated OFELs, and OFELs additionally exposed to the contraceptive Levonorgestrel (LNG, 10 µM). n=3 independent experiments. mean± S.D.; one-way Anova and Tukey's multiple comparisons test, * is P = 0.0211, *** is P = 0.0006.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Trophectoderm state is crucial for interaction with endometrium during implantation. a. Representative images of human blastoids shortly after attachment to an OFEL. Dotted line outlines the inner cluster of blastoids that were formed using GFP^+ naive hPSCs (**top**, also see Supplementary Video 3). Immunofluorescence stainings for NR2F2 (Magenta) and OCT4 (Yellow) in blastoids shortly after attachment to an OFEL (**bottom**). **b**. Immunofluorescence stainings for NR2F2 (Magenta) and OCT4 (Yellow) and respective fluorescence intensity profiles of representative blastoids immediately after attachment onto OFEL. Profiles were measured perpendicular to the plane of attachment (**right**). Line width, 10 µm. Y axis shows normalized intensity. **c**. Quantification of the distance between the first peak of fluorescence intensity profiles of NR2F2 and OCT4. n=10 attached blastoids. mean± S.D. **d**. Pseudotime analysis of human pre-implantation development showing the expression of *IL6*, *IL6R*, *GP130* and *STAT3*. Gene expression analysis is performed by using the public data analysis tool (https://bird2cluster.univ-nantes.fr/demo/PseudoTimeUI/). **e**. Quantification of the dose dependent effect of LIF on the yield of blastoids. n=2 (without Lif) and n=3 (all other conditions) independent experiments. mean± S.D. **f**. Immunofluorescence staining for NANOG (Yellow) and CDX2 (Cyan) (**left**), OCT4 (Yellow) and GATA3 (Cyan) (**middle**) and CDX2 (Cyan) and NR2F2 (Magenta) (**right**) in representative trophospheres formed from a blastoid exposed to SC144. Scale bar: 50 µm. **g**. Immunofluorescence staining for NANOG (Yellow) and CDX2 (Cyan) (**left**), OCT4 (Yellow) and GATA3 (Cyan) (**right**) in representative trophospheres formed from a blastoid exposed to XMU-MP-1. Scale bar: 50 µm. **h**. Heatmap of key lineage specific genes differentially expressed in bulk transcriptome of the trophectoderm of blastoids (TROP2 positive cells), trophospheres (SC144 or XMU) and TSCs (2D or 3D) compared to naive hPSCs. **i**. PCA plot computed using bulk transcriptome of blastoid cells, hPSCs (naive, primed or blastoid rederived naive cell lines), TSCs (bTS5, blastocyst rederived lines or human stem cell derived TSC like cells) and pluripotent stem cell derived primitive endoderm like cells (RACL or NACL cells). **j**. Immunofluorescence stainings for CDX2 (Cyan) (**left**) and CK7 (Magenta) and GATA3 (Cyan) (**right**) in aggregates formed from bTS5 hTSCs. Counterstain with Hoechst marking DNA. Scale bar: 50 µm. **k**. Representative phase contrast images of aggregates of naive hPSCs, deposited onto stimulated OFELs. Scale bar: 100 µm. **l**. List of selected putative ligand-receptor pairs involved in cross-talk between polar trophectoderm and endometrial epithelial cells. The list was generated by in silico ligand receptor analysis of genes enriched in polar trophectoderm and stimulated OFEL, using Cellinker[47].



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Human blastoids recapitulate aspects of peri-implantation progression until day 13. a. Bright-field images of human blastoids (96 h) cultured for 4 additional days on a low attachment plate in post implantation culture condition (**left**). Each row shows a time series of an individual blastoid for 4 days. Note that, blastoids stably retain cavities at least for 2 days upon transferring to IVC media which has different osmolarity compared to the N2B27 media with PALLY. (See the methods for the composition of post implantation culture media.) Scale bar: 200 µm. Quantification of percentage of blastoids retaining cavities on each day of postimplantation stage culture (**right**). n=2 independent experiments. **b**. Immunofluorescence staining for the syncytiotrophoblast-associated marker CGβ (Magenta) in GFP^+ blastoids attached onto stimulated OFELs (48 h after deposition) (**left**). Counterstain with Hoechst marking DNA. Scale bar: 50 µm. ELISA measurements of the concentration of the protein CGβ secreted into the culture medium of unstimulated OFELs with unattached blastoids and stimulated OFELs with attached blastoids (24 and 48 h) (**right**). n=3 independent experiments. mean± S.D.; one-way Anova and Tukey's multiple comparisons test, **** is P = 0.00006. **c**. Immunofluorescence stainings for CDX2 (Cyan), NR2F2 (Magenta) and SOX2 (Yellow) in blastoids grown in postimplantation culture condition for 4 days. Scale bar: 100 µm. **d**. Immunofluorescence stainings for OCT4 (Yellow), CK7 (Cyan) and GATA4 (Magenta) in blastoids grown in postimplantation culture condition for 4 days. Scale bar: 100 µm. **e, f**. Immunofluorescence stainings for CGβ (Magenta) and NR2F2 (Cyan) (**e**) or HLA-G (Magenta) and GATA3 (Cyan) (**f**), in blastoids grown in postimplantation culture condition for 4 days (**e**) or 6 days (**f**). Counterstain with Hoechst marking DNA. Arrowhead points HLA-G positive EVT like cells. Scale bar: 100 µm. **g**. Immunofluorescence stainings for CD24 (Magenta) and SOX2 (Yellow) in blastoids grown in postimplantation culture condition for 6 days. Counterstain with Hoechst marking DNA. Scale bar: 100 µm. **h**. Immunofluorescence stainings for PODXL (Magenta) and SOX2 (Yellow) in blastoids grown in postimplantation culture condition for 4 days. Counterstain with Phalloidin marking F-actin (Cyan). Arrowhead points pro-amniotic-like cavity. Scale bar: 100 µm. **i-k**. Immunofluorescence stainings for SOX2 (Yellow), GATA3 (Cyan) and CDX2 (Magenta) (**i**), SOX2 (Yellow), CDX2 (Magenta) and TFAP2C (Cyan) (**j**), OCT4 (Yellow), GATA4 (Magenta) and OTX2 (Cyan) (**k**) in blastoids grown in postimplantation culture condition for 4 days. Counterstain with Hoechst marking DNA. Scale bar: 100 µm. **l**. Quantification of number of cells belonging to EPI, TE or PrE lineages in the blastoids cultured in postimplantation culture condition for four days on glass or OFEL. n=7 biological replicates. mean± S.D. **m**. Immunofluorescence stainings for OCT4 (Yellow), GATA3 (Cyan) and GATA4 (Magenta) in blastoids grown in postimplantation culture condition for 6 days corresponding to time equivalent of day 13 of cultured human blastocyst (**left**). Scale bar: 100 µm.

nature portfolio

Corresponding author(s): Nicolas Rivron

Last updated by author(s): Sep 14, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The phase contrast images of cultured cells or cell aggregates were acquired using Thermo Fisher scientific EVOS cell imaging system and inverted wide field microscope Axio VertA1.The fluorescent images and time-lapse images were acquired using Olympus IX83 microscope with Yokogawa W1 spinning disk (Software: CellSense 2.3 ; camera: Hamamatsu Orca Flash 4.0) or Nikon Eclipse Ti E inverted microscope, equipped with a Yokogawa W1 spinning disc (Software: Visiview 4.5.0.7 ; camera:Andor Ixon Ultra 888 EMCCD). Realtime PCR results were collected using CFX384 system (bio-rad). Single cell transcriptome libraries and bulk transcriptome libraries were sequenced using Illumina Novaseq 6000
Data analysis	The confocal images were analyzed using FIJI 1.53k, Bitplane Imaris 9.7.0 or Volocity 6.5 softwares.The cell-cell interactions were infered using the Cellinker Webpage (no version number available)FACS data were analysed using FACS DiVa 9.0.1All the statistical analysis were performed using Graphpad Prism 8.1.1 (330)Data was analyzed using trim-galore v0.6.6, hisat2 v2.2.1, htseq-count v0.13.5, RSEM v1.3.3, R v4.0.3, Seurat v4.0.1, computeSumFactors in scran package v1.18.7, multiBatchNorm in batchelor v1.6.3, SeuratWrappers v0.3.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw data for single cell and bulk RNA sequencing data of blastoids were deposited at the GEO repository under the accession number GSE177689. Published human embryo data are available in the GSE109555, E-MTAB-3929.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. For each experiment, the sample size is determined based on an interval of confidence equal or above 95% as described in the methods and in the text, and on the sampling error, which is estimated based on the standard deviations during pilot studies.
Data exclusions	In single cell sequencing data, based on initial evaluation of per-cell quality control metrics and outlier identification using the median absolute deviation algorithm, cells with <= 2000 detected genes or >= 12.5% mitochondrial gene percentage were filtered out.
Replication	All attempts at replication were successful over 3 independent experiments.
Randomization	Samples were randomly allocated to groups prior to treatments which prevented any bias in the interpretation of data
Blinding	The investigators were blinded by preparing in advance the different cocktails and allocating numbers to the different treatments used to stimulate the stem cells.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

anti-NANOG (Abcam #ab109250, clone EPR2027(2), dilution 1:100)
anti-CDX2 (Emergo Europe #MU392A-5UC, clone CDX2-88, dilution 1:100)
anti-GATA4 (Invitrogen #14-9980-82, clone eBioEvan, dilution 1:400)
anti-OCT4 (Santacruz Biotechnology #sc-5279, clone C-10, dilution 1:100)
anti-GATA3 (Santacruz Biotechnology #sc-9009, clone H-48, dilution 1:100)
anti-GATA3 (Invitrogen #14-9966-82, clone TWAJ, dilution 1:200)
anti-ZO-1 (Invitrogen #339100, clone ZO1-1A12, dilution 1:100)
anti-CDH1 (eBioscience #14-3249-82, clone DECMA-1, dilution 1:250)

anti-aPKC (Santacruz Biotechnology #sc-216, clone MC5, dilution 1:100)
anti-CK7 (Abcam #ab181598, clone EPR17078, dilution 1:300)
anti-KLF17 (Sigma #HPA002926, clone NA (polyclonal), dilution 1:200)
anti-YAP1 (Santacruz Biotechnology #sc-101199, clone 63.7, dilution 1:100)
anti-NR2F2 (Santacruz Biotechnology #ab211776, clone EPR18442, dilution 1:100)
anti-AQP3 (ac on lines #ABIN863208, clone NA (ABIN863208), dilution 1:100)
anti-CGB (Dako #A0231, clone NA (Polyclonal), dilution 1:300)
anti-CGB (Abcam #ab9582, clone 5H4-E2, dilution 1:200
anti-PAEP (Abclonal #A5751, clone NA (Polyclonal), dilution 1:500)
anti-FOXA2 (Cell signal #8186, clone D56D6, dilution 1:1000)
anti-Acetylated Tubulin (Sigma #T7451, clone 6-11B-1, dilution 1:500)
anti-IFI16 (novusbio #NBP1-83118, clone NA (Polyclonal), dilution 1:100)
anti-GATA2 (Abcam #ab109241, clone EPR2822(2), dilution 1:250)
anti-TROP2 (R&D systems #MAB650, clone 77220, dilution 1:100 (IHC) 1:50 (FACS))
anti-PDGFRα (R&D systems #AF307, clone NA (Polyclonal), dilution 1:100 (IHC) 1:50 (FACS))
anti-SOX17 (R&D systems #AF1924-SP, clone EPR20684, dilution 1:200)
anti-KLF4 (Sigma #HPA002926, clone NA (Polyclonal), dilution 1:200)
anti-OTX2 (R&D systems #AF1979, clone NA (Polyclonal), dilution 1:100)
anti-SUSD2 (Miltenyibiotec #130-117-682, clone W5C5, dilution 1:100)
anti-CCR7 (Thermo scientific #MA5-31992, clone SR36-04, dilution 1:200)
anti-SOX2 (Invitrogen #14-9811-80, clone Btjce, dilution 1:200)
anti-MUC1 (Invitrogen #MA1-35039, clone 115D8, dilution 1:200)
anti-CD24 (BD Biosciences #561644, clone ML5, dilution 1:100)
anti-PODXL (R&D systems #MAB1658, clone 222328, dilution 1:100)
anti-TFAP2C (R&D systems #AF5059, clone NA (Polyclonal), dilution 1:100)

Validation

Validation statements available from manufacturers:
anti-NANOG (https://www.abcam.com/nanog-antibody-epr20272-ab109250.html)
anti-CDX2 (https://www.labome.com/product/Biogenex/MU392A-5UC.html)
anti-GATA4 (https://www.thermofisher.com/antibody/product/Gata-4-Antibody-clone-eBioEvan-Monoclonal/14-9980-82)
anti-OCT4 (https://www.scbt.com/p/oct-3-4-antibody-c-10_
anti-GATA3 (https://www.scbt.com/p/gata-3-antibody-h-48?productCanUrl=gata-3-antibody-h-48&_requestid=3381940)
anti-GATA3 (https://www.thermofisher.com/antibody/product/Gata-3-Antibody-clone-TWAJ-Monoclonal/14-9966-82)
anti-ZO-1 (https://www.thermofisher.com/antibody/product/ZO-1-Antibody-clone-ZO1-1A12-Monoclonal/33-9100)
anti-CDH1 (https://www.thermofisher.com/antibody/product/CD324-E-Cadherin-Antibody-clone-DECMA-1-Monoclonal/14-3249-82)
anti-aPKC (https://www.scbt.com/p/pkc-zeta-antibody-c-20)
anti-CK7 (https://www.abcam.com/cytokeratin-7-antibody-epr17078-cytoskeleton-marker-ab181598.html)
anti-KLF17 (https://www.sigmaaldrich.com/AT/en/product/sigma/hpa002926)
anti-YAP1 (https://www.scbt.com/p/yap-antibody-63-7)
anti-NR2F2 (https://www.abcam.com/nr2f2-antibody-epr18442-ab211776.html)
anti-AQP3 (https://www.antibodies-online.com/antibody/863208/anti-Aquaporin+3+Gill+Blood+Group+AQP3+C-Term+antibody/)
anti-CGB (https://www.labome.com/product/Dako/A0231.html)
anti-CGB (https://www.abcam.com/hcg-beta-antibody-5h4-e2-ab9582.html)
anti-PAEP (https://abclonal.com/catalog-antibodies/PAEPPolyclonalAntibody/A5751)
anti-FOXA2 (https://www.cellsignal.at/products/primary-antibodies/foxa2-hnf3b-d56d6-xp-rabbit-mab/8186)
anti-Acetylated Tubulin (https://www.sigmaaldrich.com/AT/en/product/sigma/t7451?
gclid=Cj0KCQjw5uWGBhCTARIsAL70sLKfFm6GzCha4llea9SnzsZ5NvDKPNukmZNa61-V6xRBwJWw0GbCU6gaAmeSEALw_wcB)
anti-IFI16 (https://www.novusbio.com/products/ifi16-antibody_nbp1-83118)
anti-GATA2 (https://www.abcam.com/gata2-antibody-epr28222-ab109241.html)
anti-TROP2 (https://www.rndsystems.com/products/human-trop-2-antibody-77220_mab650)
anti-PDGFRα (https://www.rndsystems.com/products/human-pdgf-ralpha-antibody_af-307-na)
anti-SOX17 (https://www.rndsystems.com/products/human-sox17-antibody_af1924)
anti-KLF4 (https://www.sigmaaldrich.com/AT/en/product/sigma/hpa002926)
anti-OTX2 (https://www.rndsystems.com/products/human-otx2-antibody_af1979)
anti-SUSD2 (https://www.miltenyibiotec.com/AT-en/products/susd2-antibody-anti-human-w5c5.html#gref)
anti-CCR7 (https://www.thermofisher.com/antibody/product/CCR7-Antibody-clone-SR36-04-Recombinant-Monoclonal/MA5-31992)
anti-SOX2 (https://www.thermofisher.com/antibody/product/SOX2-Antibody-clone-Btjce-Monoclonal/14-9811-82)
anti-MUC1 (https://www.thermofisher.com/antibody/product/MUC1-Antibody-clone-115D8-Monoclonal/MA1-35039)
anti-CD24 (https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/alexa-fluor-647-mouse-anti-human-cd24.561644)
anti-PODXL (https://www.rndsystems.com/products/human-podocalyxin-antibody-222328_mab1658)
anti-TFAP2C (https://www.rndsystems.com/products/human-ap-2gamma-antibody_af5059)

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

embryonic stem cell lines: Shef6, HNES1 and iPSC lines: cR-nCRM2 and niPSC16.2.b were provided by the laboratory of Austin Smith
H9 and H9-GFP reset to naive state were provided by the Laboratory of Yasuhiro Takashima
hTSC line bTS5 was provided by the laboratory of Takahiro Arima
Endometrial organoids were provided by the laboratory of Hossein Baharvand

Authentication

H9 (primed and naive PSCs) and TSC-bTS5 were included in single cell sequencing analysis to authenticate their identity

Mycoplasma contamination	Cells were routinely tested for mycoplasma contamination. No contamination was detected.
Commonly misidentified lines (See ICLAC register)	no misidentified lines were used in the study

Human research participants

Policy information about studies involving human research participants

Population characteristics	Not applicable because no human subjects were involved in this research. We used donated embryos surplus to IVF treatment and endometrial organoids from a biobank.
Recruitment	Informed consent was obtained from all couples that donated spare embryos following IVF treatment and from people that previously donated endometrial samples that constituted the endometrial biobank used in this study.
Ethics oversight	The use of human embryos 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. Human endometrium samples were obtained from patients who signed an informed consent form and protocols approved by the Ethics Committee of Royan Institute (IR.ACECR.ROYAN.REC. 1397.93) and ethical approval from the Ethics Committee of the Shahid Beheshti University of Medical Sciences (IR.SBMU.MSP.REC. 1396.25). The Wicell line H9 was used under the agreement 20-WO-341 for a research program entitled 'Modeling early human development: Establishing a stem cell based 3D in vitro model of human blastocyst (blastoids)'. Blastoid generation was approved by the Commission for Science Ethics of the Austrian Academy of Sciences. All experiments complied with all relevant guidelines and regulations, including the 2021 ISSCR guidelines that forbid the transfer of human blastoids into an uterus.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were collected by dissociated blastoids with sequential treatment of 300units/ml collagenase IV for 30minutes and 0.5% Trypsin (10x) for 20minutes. Cells were stained with Trop2 and PDGFRa antibodies followed by the secondary antibodies
Instrument	FACS Aria III (BD)
Software	DiVa version on the F02 is 9.0.1.
Cell population abundance	Abundance of the cell populations of interest was determined by the appropriate negative control and the purity of sorted population was assessed by the post sort analysis.
Gating strategy	FSC-A/SSC-A and SSC-H/SSC-W gates were applied to remove debris, and non-single cell aggregates respectively. Dead cells were excluded by using DAPI signal. A example of FACS gating strategy is available at the Supplementary Figure 2A.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

1.3.3 Discussion

L'année 2021 a été marquée par la publication de plusieurs protocoles permettant la génération de modèles d'embryons humains appelés « blastoïdes ». Ces structures en 3D peuvent être obtenues à partir de cellules souches pluripotentes naïves déjà établies (Yanagida et al., 2021; Yu et al., 2021), ou par reprogrammation de cellules somatiques (Liu et al., 2021). L'objectif de ces études fut de générer des structures cellulaires comparables aux embryons tant au niveau du transcriptome que de la séquence d'expression des marqueurs du développement.

Le blastoïde se forme (presque) comme le blastocyste

Après avoir agrégé des cellules pluripotentes naives, l'inhibition de la voie HIPPO par de l'acide lysophosphatidique (LPA) s'avère essentielle pour la formation du blastoïde humain. Dans ces conditions, YAP1 est transloqué dans le noyau des cellules externes et va activer, comme dans l'embryon humain, le programme trophectoderme précédent la cavitation (Kagawa et al., 2021). Comme chez l'embryon humain, la formation du blastoïde est impossible si l'activité de la aPKC est inhibée. Les amas de cellules cohésives les unes aux autres, la polarisation, et la modulation de la voie HIPPO rappelent les mécanismes actifs dans la morula humaine au 4ème jour de développement (Gerri et al., 2020). Néanmoins, la comparaison s'arrête dès lors que l'on regarde le transcriptome des cellules du blastoïde. En début d'induction du blastoïde, ces dernières ne se groupent pas avec l'embryon pré-blastocyste. Curieusement, avant même de caviter (24-60h), certaines cellules du blastoïde co-expriment YAP1 et GATA2 mais pas GATA3 et CDX2 qui n'arrivent que plus tard (60h).

Chose qui n'arrive pas dans l'embryon humain, où GATA3 et CDX2 précèdent GATA2, et où ce dernier n'est pas détectable avant la cavitation (Meistermann et al., 2021).

En revanche, une fois formé, le blastoïde est bien composé d'analogues des trois lignées EPI, TE, PrE présentes dans l'embryon. Les signatures transcriptomiques des analogues TE et EPI sont détectables à 24h-60h, et celle du PrE entre 60h-96h. Ainsi, le PrE est spécifié en dernier dans le blastoïde, comme nous le supposons dans le blastocyste. Nous montrons de plus que l'analogue TE du blastoïde mature comme le TE du blastocyste, avec une expression de NR2F2 dans le TE polaire et CDX2 dans le TE mural (Kagawa et al., 2021, Figure 3F). Les blastoïdes ont une taille et un nombre de cellules comparables à ceux d'un blastocyste expansé B4.

Le protocole de Kagawa et collègues a un rendement de formation des blastoïdes de plus de 70%, contre 5-10% pour les autres protocoles publiés (Sozen et al., 2021; Yu et al., 2021). Cette différence est probablement en partie due à l'inhibition de la voie HIPPO par l'ajout de LPA. Sans LPA, le rendement diminue à moins de 10% comme pour les autres protocoles (Kagawa et al., 2021, Figure 3C).

Le blastoide : le meilleur modèle de l'implantation de l'embryon humain ?

Nous souhaitions montrer que ces blastoïdes avaient le même potentiel de culture postimplantatoire que les embryons humains. Nous avons cultivé en parallèle des embryons humains et des blastoïdes sur une monocouche de cellules endométriales primaires (Boretto et al., 2017). Il fut intéressant de constater que les blastoïdes n'ont pu s'attacher qu'aux cellules endométriales préalablement stimulées par des œstrogènes (E2) et progestérones (P4). *In vivo*, ces deux hormones sont exprimées pendant la phase de réceptivité, période du cycle menstruel durant laquelle l'endomètre est capapble d'accueillir l'embryon (Wang et al., 2020). Pour la co-culture,

nous avons utilisé du milieu CMRL modifié qui supporte la culture d'embryons de singes jusqu'à jour 20 (Ma et al., 2019). Nous n'avions pas assez de réplicas pour comparer l'efficicacité de culture des milieux CMRL et IVC. Malgré tout, les embryons humains ont pu survivre jusqu'à jour 10 et former la cavité amniotique au sein de l'épiblaste exprimant OCT4, comme les blastoïdes (Kagawa et al., 2021, Figure 4G). Les blastoïdes attachés sont capables de former du ST et de l'EVT, comme montré par la sécrétion de β-HCG et l'expression de HLA-G respectivement. Il serait intéressant de cultiver ces blastoïdes dans des conditions dites « 3D » (Xiang et al., 2020). Cette méthode de culture se veut plus proche de la réalité où, *in vivo,* l'embryon garde sa structure sphérique même implanté. Nous avons tenté au laboratoire de cultiver des embryons humains selon cette méthode, sans succès. Sur 20 embryons décongelés au stade blastocyste J5, aucun n'a survécu plus de 48h.

Pour résumer, le blastoïde humain modélise fidèlement certains aspects du développement péri-implantatoire humain comme l'expression des trois lignées EPI, TE, PrE, la maturation du trophectoderme, l'attachement *in vitro*, la formation de la cavité amniotique et la différenciation dans la lignée trophoblastique. En revanche, l'analyse transcriptomique et le profil d'expression des marqueurs moléculaires suggèrent que la formation du blastoïde est différente de celle d'un blastocyste qui se formerait depuis la morula. En d'autres termes, le blastoïde ne semble pas être adapté à l'étude de la première spécification, mais est un bon modèle d'étude du blastocyste à partir du stade B3 jusqu'en postimplantation.

Est-ce que le blastoïde est un « embryon artificiel » ?

Le blastoïde ressemble à l'embryon. Il en a la forme, la taille, le nombre de cellules. Comme l'embryon il cavite, croit, spécifie des lignées cellulaires différentes, est capable de modéliser l'implantation *in vitro.* Une fois attaché, il peut sécréter l'hormone de grossesse β-HCG*,* et former une cavité amniotique au sein de ses cellules pluripotentes.

La communauté scientifique s'est alors posée la question de son statut, et s'il fallait fixer des limites à son étude. Si le blastoïde est un embryon artificiel, ne devrait-il pas être soumis aux mêmes lois que l'embryon humain ? Et si c'est le cas, sa création ne reviendrait elle pas à une forme de clonage ? Car à la base de ce blastoïde, il y a des cellules dérivées d'embryons ou des cellules somatiques reprogrammées de patients. Ces derniers n'avaient sans doute pas conscience que leurs cellules serviraient à créer des structures ressemblant à un embryon.

Cette problématique s'est notamment illustrée en Australie où ont été générés les premiers blastoïdes humain (Liu et al., 2021). La Loi australienne désigne un embryon humain comme « … toute structure initiant un développement organisé d'une entité biologique avec un génome humain qui a le potentiel de se développer jusqu'à ou au-delà de l'apparition de la ligne primitive » (RIHE Act). Elle interdit également le clonage en désignant un embryon clone comme « un embryon qui est une copie génétique d'un autre humain mort ou vivant… ». Ces deux lois combinées ont temporairement stoppé les projets de recherche impliquant la génération de blastoïdes en Australie, qui sont désormais considérés comme des embryons à part entière. En France, la dernière révision de la Loi n° 2021-1017 du 2 août 2021 relative à la bioéthique autorise expressément la génération de blastoïdes humains.

Le statut du blastoïde dépend de notre définition de l'embryon. Si on entend par embryon le produit de la fusion des deux gamètes mâle et femelle, alors le blastoïde n'est pas un embryon. En revanche, nous pouvons aussi définir l'embryon par sa fonction et son devenir. Si l'embryon est une structure cellulaire pouvant engendrer une grossesse et une naissance, le statut du blastoïde est alors incertain. Pour mesurer son potentiel, il faudrait transférer un blastoïde humain dans un utérus humain. A l'heure actuelle, cette pratique est proscrite par les recommandations internationales (Lovell-Badge et al., 2021).

Les premiers blastoïdes de souris générés par agrégats de cellules souches pluripotentes et de cellules souches trophoblastiques n'étaient pas capable de former un fœtus (Rivron et al., 2018). Récemment, une étude a permis la génération d'embryons synthétiques de souris au-delà de la gastrulation, *ex utero,* uniquement à partir de cellules souches pluripotentes naïves (Tarazi et al., 2022). A l'intérieur d'un incubateur rotatif, « l'embryoïde » murin croit jusqu'à un équivalent E8.5, forme le tube neural et un cœur battant. L'utilisation de ce type de dispositif pour la culture de blastoïdes humains, sans mère porteuse, pourrait s'apparenter à un compromis éthique et scientifique dans notre compréhension des premiers stades de la vie humaine (Clark et al., 2021).

Conclusion et Perspectives

La recherche sur les embryons humains fut permise en France pour sa finalité médicale, afin de permettre une meilleure prise en charge des couples infertiles. Ce projet de recherche fut initié en 2013, au moment où la oi le permit, et où les technologies de séquençage ARN devenaient suffisamment performantes et adaptées à ce type de projet. En moins de dix années, le domaine s'est vu bousculé par la génération de modèles cellulaires mimant les lignées embryonnaires, la culture *in vitro* d'embryons humains postimplantatoires, et l'apport constant de nouvelles techniques d'analyses bioinformatiques.

Les résultats que nous avons publiés sont une synthèse pluridisciplinaire à la fois de recherche fondamentale, de biologie de la reproduction, et de bioinformatique. Nous avons reconstitué un patron transcriptomique et protéique de l'embryon humain qui nous a permis de situer la première spécification humaine, d'affiner notre compréhension de l'émergence de l'EPI et du PrE, et de caractériser la maturation du trophectoderme avant et après l'implantation de l'embryon. Ces travaux ont servi de référence à l'élaboration du modèle de blastoïde humain qui, pour être un bon modèle, doit d'être comparé et comparable à l'embryon humain.

Nous pensons que nos travaux seront utiles en clinique, notamment pour l'élaboration de meilleures conditions de culture. En reconstituant la séquence d'expression du génome embryonnaire, nous avons listé des voies de signalisation potentiellement actives ou activables. La modulation de ces voies pourrait améliorer le développement des lignées embryonnaires (Wamaitha et al., 2020), favoriser l'éclosion de l'embryon (Denizot et al., 2022), voir l'attachement et sa survie dans l'utérus. NR2F2 est un facteur de transcription qui marque le TE polaire juste avant l'implantation de

l'embryon. Si nous montrions dans des blastoïdes que l'expression de *NR2F2* était un facteur prédictif positif des chances d'implantation, il serait intéressant de pouvoir détecter son expression directement dans le milieu de culture. Nous pourrions par exemple identifier l'expression de facteurs sécrétés dont l'expression est corrélée à celle de *NR2F2*, et ainsi sélectionner uniquement les embryons les plus à même de s'implanter. Des stratégies de séquençage des protéines sécrétées dans les milieux de culture par spectrométrie de masse sont une piste de recherche en vue de l'amélioration de la sélection des embryons (Cortezzi et al., 2013; Eldarov et al., 2022; Poli et al., 2015).

Tout comme le scRNAseq le fut en son temps, il est concevable que la spectrométrie de masse en cellule unique (scMS) soit la prochaine grande évolution dans l'étude du développement embryonnaire (Tajik et al., 2022). Contrairement au scRNAseq qui mesure la présence d'ARNm mais pas de protéines, et aux immunofluorescences qui dépendent d'anticorps parfois indisponibles ou aspécifiques, la scMS est susceptible d'offrir un niveau d'analyse sans commune mesure avec nos moyens actuels. Elle pourrait permettre une cartographie non supervisée des protéines exprimées au cours du développement dans les cellules de l'embryon, et ainsi offrir une vision résolument fonctionnelle du développement humain.

Enfin, la disponibilité des modèles d'étude blastoïdes a fait entrer la biologie du développement humain à l'ère du haut débit. A partir de hNPSC, il est possible d'obtenir en moins de deux semaines autant de blastoïdes que nous avons obtenus d'embryons en sept ans. Le génome de ces blastoïdes peut être légalement et aisément modifié, ce qui permettrait une meilleure compréhension des mécanismes d'émergence et de régulation des lignées cellulaires, et de l'implantation. Ces programmes de recherche seront peut-être les prémices indispensables à tout essai clinique chez l'Homme, et seront peut-être amenés à complètement remplacer le

modèle d'étude murin qui, comparé aux autres mammifères, s'illustre de plus en plus par ses singularités.

Bibliographie

Aberkane A, Essahib W, Spits C, De Paepe C, Sermon K, Adriaenssens T, Mackens S, Tournaye H, Brosens JJ, Van de Velde H. 2018. Expression of adhesion and extracellular matrix genes in human blastocysts upon attachment in a 2D co-culture system. *Mol Hum Reprod* **24**:375–387. doi:10.1093/molehr/gay024

Acampora D, Omodei D, Petrosino G, Garofalo A, Savarese M, Nigro V, Di Giovannantonio LG, Mercadante V, 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 Rep* **15**:2651–2664. doi:10.1016/j.celrep.2016.05.041

Ahlström A, Westin C, Reismer E, Wikland M, Hardarson T. 2011. Trophectoderm morphology: an important parameter for predicting live birth after single blastocyst transfer. *Hum Reprod* **26**:3289–3296. doi:10.1093/humrep/der325

Alarcon VB. 2010. Cell polarity regulator PARD6B is essential for trophectoderm formation in the preimplantation mouse embryo. *Biol Reprod* **83**:347–358. doi:10.1095/biolreprod.110.084400

Allègre N, Chauveau S, Dennis C, Renaud Y, Meistermann D, Estrella LV, Pouchin P, Cohen-Tannoudji M, David L, Chazaud C. 2022. NANOG initiates epiblast fate through the coordination of pluripotency genes expression. *Nat Commun* **13**:3550. doi:10.1038/s41467-022-30858-8

Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology. 2011. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting†. *Human Reproduction* **26**:1270–1283. doi:10.1093/humrep/der037

Alter L, Boitrelle F, Sifer C. 2014. Comment sélectionner aujourd'hui le meilleur embryon à transférer ? *Gynécologie Obstétrique & Fertilité* **42**:515–525. doi:10.1016/j.gyobfe.2014.05.006

Artus J, Piliszek A, Hadjantonakis A-K. 2011. The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17. *Developmental Biology* **350**:393–404. doi:10.1016/j.ydbio.2010.12.007

Asami M, Lam BYH, Ma MK, Rainbow K, Braun S, VerMilyea MD, Yeo GSH, Perry ACF. 2022. Human embryonic genome activation initiates at the one-cell stage. *Cell Stem Cell* **29**:209-216.e4. doi:10.1016/j.stem.2021.11.012

Assou S, Boumela I, Haouzi D, Monzo C, Dechaud H, Kadoch I-J, Hamamah S. 2012. Transcriptome Analysis during Human Trophectoderm Specification Suggests New Roles of Metabolic and Epigenetic Genes. *PLOS ONE* **7**:e39306. doi:10.1371/journal.pone.0039306

Bak CW, Seok HH, Song S-H, Kim ES, Her YS, Yoon TK. 2012. Hormonal imbalances and psychological scars left behind in infertile men. *J Androl* **33**:181–189. doi:10.2164/jandrol.110.012351

Bartolacci A, Canto MD, Guglielmo MC, Mura L, Brigante C, Renzini MM, Buratini J. 2021. Early embryo morphokinetics is a better predictor of post-ICSI live birth than embryo morphology: speed is more important than beauty at the cleavage stage. *Zygote* **29**:495–502. doi:10.1017/S0967199421000253

Bertero A, Madrigal P, Galli A, Hubner NC, Moreno I, Burks D, Brown S, Pedersen RA, Gaffney D, Mendjan S, Pauklin S, Vallier L. 2015. Activin/nodal signaling and NANOG orchestrate human embryonic stem cell fate decisions by controlling the H3K4me3 chromatin mark. *Genes Dev* **29**:702–717. doi:10.1101/gad.255984.114

Bessonnard S, De Mot L, Gonze D, Barriol M, Dennis C, Goldbeter A, Dupont G, Chazaud C. 2014. Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. *Development* **141**:3637–3648. doi:10.1242/dev.109678

Binder NK, Evans J, Salamonsen LA, Gardner DK, Kaitu'u-Lino TJ, Hannan NJ. 2016. Placental Growth Factor Is Secreted by the Human Endometrium and Has Potential Important Functions during Embryo Development and Implantation. *PLOS ONE* **11**:e0163096. doi:10.1371/journal.pone.0163096

Blakeley P, Fogarty NME, Snell P, Christie L, Robson P, Niakan KK. 2015. Defining the three cell lineages of the human blastocyst by single- cell RNA-seq 34.

Bodri D, Kawachiya S, Brucker MD, Tournaye H, Kondo M, Kato R, Matsumoto T. 2014. Cumulative success rates following mild IVF in unselected infertile patients: a 3-year, single-centre cohort study. *Reproductive BioMedicine Online* **28**:572–581. doi:10.1016/j.rbmo.2014.01.002

Boivin J, Bunting L, Collins JA, Nygren KG. 2007. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Hum Reprod* **22**:1506–1512. doi:10.1093/humrep/dem046

Bontekoe S, Mantikou E, van Wely M, Seshadri S, Repping S, Mastenbroek S. 2012. Low oxygen concentrations for embryo culture in assisted reproductive technologies. *Cochrane Database Syst Rev* CD008950. doi:10.1002/14651858.CD008950.pub2

Boretto M, Cox B, Noben M, Hendriks N, Fassbender A, Roose H, Amant F, Timmerman D, Tomassetti C, Vanhie A, Meuleman C, Ferrante M, Vankelecom H. 2017. Development of organoids from mouse and human endometrium showing endometrial epithelium physiology and long-term expandability. *Development* **144**:1775–1786. doi:10.1242/dev.148478

Boroviak T, Stirparo GG, Dietmann S, Hernando-Herraez I, Mohammed H, Reik W, Smith A, Sasaki E, Nichols J, Bertone P. 2018. Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* **145**:dev167833. doi:10.1242/dev.167833

Bou G, Liu S, Sun M, Zhu J, Xue B, Guo J, Zhao Y, Qu B, Weng X, Wei Y, Lei L, Liu Z. 2017. CDX2 is essential for cell proliferation and polarity in porcine blastocysts. *Development* dev.141085. doi:10.1242/dev.141085

Braude P, Bolton V, Moore S. 1988. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* **332**:459–461. doi:10.1038/332459a0

Brons IGM, Smithers LE, Trotter MWB, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos 191–195.

Buehr M, Meek S, Blair K, Yang J, Ure J, Silva J, McLay R, Hall J, Ying Q-L, Smith A. 2008. Capture of Authentic Embryonic Stem Cells from Rat Blastocysts 1287–1298.

Carson SA, Kallen AN. 2021. Diagnosis and Management of Infertility: A Review. *JAMA* **326**:65–76. doi:10.1001/jama.2021.4788

Castel G, Meistermann D, Bretin B, Firmin J, Blin J, Loubersac S, Bruneau A, Chevolleau S, Kilens S, Chariau C, Gaignerie A, Francheteau Q, Kagawa H, Charpentier E, Flippe L, François-Campion V, Haider S, Dietrich B, Knöfler M, Arima T, Bourdon J, Rivron N, Masson D, Fournier T, Okae H, Fréour T, David L. 2020. Induction of Human Trophoblast Stem Cells from Somatic Cells and Pluripotent Stem Cells. *Cell Rep* **33**:108419. doi:10.1016/j.celrep.2020.108419

Cauffman G, De Rycke M, Sermon K, Liebaers I, Van de Velde H. 2009. Markers that define stemness in ESC are unable to identify the totipotent cells in human preimplantation embryos. *Human Reproduction* **24**:63–70. doi:10.1093/humrep/den351

Chadid ML, Carpio J, Valdivieso P, Zambrano M, García-Ferreyra J, Valdivieso-Mejía P. 2015. Percentage of Blastulation on the number and function of metaphase II oocytes. *JBRA Assisted Reproduction* **19**:111–113.

Chau K, Hennessy A, Makris A. 2017. Placental growth factor and pre-eclampsia. *J Hum Hypertens* **31**:782–786. doi:10.1038/jhh.2017.61

Chazaud C, Yamanaka Y. 2016. Lineage specification in the mouse preimplantation embryo. *Development* **143**:1063–1074. doi:10.1242/dev.128314

Chazaud C, Yamanaka Y, Pawson T, Rossant J. 2006. Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway. *Developmental Cell* **10**:615–624. doi:10.1016/j.devcel.2006.02.020

Cinkornpumin JK, Kwon SY, Guo Y, Hossain I, Sirois J, Russett CS, Tseng H-W, Okae H, Arima T, Duchaine TF, Liu W, Pastor WA. 2020. Naive Human Embryonic Stem Cells Can Give Rise to Cells with a Trophoblast-like Transcriptome and Methylome. *Stem Cell Reports* **15**:198–213. doi:10.1016/j.stemcr.2020.06.003

Clark AT, Brivanlou A, Fu J, Kato K, Mathews D, Niakan KK, Rivron N, Saitou M, Surani A, Tang F, Rossant J. 2021. Human embryo research, stem cell-derived embryo models and in vitro gametogenesis: Considerations leading to the revised ISSCR guidelines. *Stem Cell Reports* **16**:1416–1424. doi:10.1016/j.stemcr.2021.05.008

Cortezzi SS, Cabral EC, Trevisan MG, Ferreira CR, Setti AS, Braga DP de AF, Figueira R de CS, Iaconelli A, Eberlin MN, Borges E. 2013. Prediction of embryo implantation potential by mass spectrometry fingerprinting of the culture medium. *Reproduction* **145**:453–462. doi:10.1530/REP-12-0168

Cosemans G, Boel A, Bekaert B, Pascal E, Stamatiadis P, Stoop D, Chuva De Sousa Lopes SM, Menten B, Coucke P, Lluis F, Heindryckx B. 2022. O-216 CRISPR/Cas9 mediated knock-out (KO) reveals a divergent role for trophectoderm markers GATA2/3 in the mouse and human preimplantation embryo. *Human Reproduction* **37**:deac105.135. doi:10.1093/humrep/deac105.135

Datta J, Palmer MJ, Tanton C, Gibson LJ, Jones KG, Macdowall W, Glasier A, Sonnenberg P, Field N, Mercer CH, Johnson AM, Wellings K. 2016. Prevalence of infertility and help seeking among 15 000 women and men. *Human Reproduction* **31**:2108–2118. doi:10.1093/humrep/dew123

De Geyter C, Calhaz-Jorge C, Kupka MS, Wyns C, Mocanu E, Motrenko T, Scaravelli G, Smeenk J, Vidakovic S, Goossens V, The European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE). 2018. ART in Europe, 2014: results generated from European registries by ESHRE†: The European IVF-monitoring Consortium (EIM)‡ for the European Society of Human Reproduction and Embryology (ESHRE). *Human Reproduction* **33**:1586–1601. doi:10.1093/humrep/dey242

De Geyter C, Calhaz-Jorge C, Kupka MS, Wyns C, Mocanu E, Motrenko T, Scaravelli G, Smeenk J, Vidakovic S, Goossens V, The European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), Gliozheni O, Hambartsoumian E, Strohmer H, Petrovskaya E, Tishkevich O, Bogaerts K, Wyns Christine, Balic D, Sibincic S, Antonova I, Pelekanos M, Rezabek K, Markova J, Lemmen J, Sõritsa D, Gissler M, Pelkonen S, Pessione F, de Mouzon J, Tandler—Schneider A, Kalantaridou S, Urbancsek J, Kosztolanyi G, Bjorgvinsson H, Mocanu Edgar, Cloherty J, Scaravelli Giulia, de Luca R, Lokshin V, Karibayeva S, Magomedova V, Bausyte R, Masliukaite I, Petanovski Z, Calleja-Agius J,

Moshin V, Simic TM, Vukicevic D, Smeenk JMJ, Romundstad LB, Janicka A, Calhaz—Jorge C, Laranjeira AR, Rugescu I, Doroftei B, Korsak V, Radunovic N, Tabs N, Virant-Klun I, Saiz IC, Mondéjar FP, Bergh C, Berger-Menz E, Weder M, Ryan H, Baranowski R, Gryshchenko M. 2020. ART in Europe, 2015: results generated from European registries by ESHRE†. *Human Reproduction Open* **2020**:hoz038. doi:10.1093/hropen/hoz038

De La Rochebrochard É. 2018. 1 enfant sur 30 conçu par assistance médicale à la procréation en France: *Population & Sociétés* **N° 556**:1–4. doi:10.3917/popsoc.556.0001

De los Santos MJ, Gámiz P, Albert C, Galán A, Viloria T, Pérez S, Romero JL, Remohï J. 2013. Reduced oxygen tension improves embryo quality but not clinical pregnancy rates: a randomized clinical study into ovum donation cycles. *Fertil Steril* **100**:402–407. doi:10.1016/j.fertnstert.2013.03.044

De Paepe C, Cauffman G, Verloes A, Sterckx J, Devroey P, Tournaye H, Liebaers I, Van de Velde H. 2013. Human trophectoderm cells are not yet committed. *Human Reproduction* **28**:740–749. doi:10.1093/humrep/des432

De Vries WN, Evsikov AV, Haac BE, Fancher KS, Holbrook AE, Kemler R, Solter D, Knowles BB. 2004. Maternal β-catenin and E-cadherin in mouse development. *Development* **131**:4435–4445. doi:10.1242/dev.01316

Deglincerti A, Croft GF, Pietila LN, Zernicka-Goetz M, Siggia ED, Brivanlou AH. 2016. Self-organization of the in vitro attached human embryo. *Nature* **533**:251–254. doi:10.1038/nature17948

Dell'Oste V, Gatti D, Giorgio AG, Gariglio M, Landolfo S, De Andrea M. 2015. The interferon-inducible DNA-sensor protein IFI16: a key player in the antiviral response. *New Microbiol* **38**:5–20.

Denizot A-L, L'Hostis A, Sallem A, Favier S, Pierre R, Do Cruzeiro M, Guilbert T, Burlet P, Lapierre J-M, Robain M, Le Lorc'H M, Vicaut E, Chatzovoulou K, Steffann J, Romana S, Méhats C, Santulli P, Patrat C, Vaiman D, Ziyyat A, Wolf JP. 2022. Cyclic fertilin-derived peptide stimulates in vitro human embryo development. *F&S Science* **3**:49–63. doi:10.1016/j.xfss.2021.12.002

Dong C, Beltcheva M, Gontarz P, Zhang B, Popli P, Fischer LA, Khan SA, Park K, Yoon E-J, Xing X, Kommagani R, Wang T, Solnica-Krezel L, Theunissen TW. 2020. Derivation of trophoblast stem cells from naïve human pluripotent stem cells. *eLife* **9**:e52504. doi:10.7554/eLife.52504

Dumoulin JC, Meijers CJ, Bras M, Coonen E, Geraedts JP, Evers JL. 1999. Effect of oxygen concentration on human in-vitro fertilization and embryo culture. *Hum Reprod* **14**:465–469. doi:10.1093/humrep/14.2.465

Ebner T, Tritscher K, Mayer RB, Oppelt P, Duba H-C, Maurer M, Schappacher-Tilp G, Petek E, Shebl O. 2016. Quantitative and qualitative trophectoderm grading allows for prediction of live birth and gender. *J Assist Reprod Genet* **33**:49–57. doi:10.1007/s10815-015-0609-9

Eldarov C, Gamisonia A, Chagovets V, Ibragimova L, Yarigina S, Smolnikova V, Kalinina E, Makarova N, Zgoda V, Sukhikh G, Bobrov M. 2022. LC-MS Analysis Revealed the Significantly Different Metabolic Profiles in Spent Culture Media of Human Embryos with Distinct Morphology, Karyotype and Implantation Outcomes. *Int J Mol Sci* **23**:2706. doi:10.3390/ijms23052706

ESHRE Guideline Group on Good Practice in IVF Labs, De los Santos MJ, Apter S, Coticchio G, Debrock S, Lundin K, Plancha CE, Prados F, Rienzi L, Verheyen G, Woodward B, Vermeulen N. 2016. Revised guidelines for good practice in IVF laboratories (2015). *Hum Reprod* **31**:685–686. doi:10.1093/humrep/dew016

Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**:154–156. doi:10.1038/292154a0

Fischer B, Bavister BD. 1993. Oxygen tension in the oviduct and uterus of rhesus monkeys, hamsters and rabbits. *J Reprod Fertil* **99**:673–679. doi:10.1530/jrf.0.0990673

Fogarty NME, Abdelbaki A, McCarthy A, Devito L, Chen AE, Munusamy P, Blakeley P, Elder K, Snell P, Christie L, Serhal P, Odia RA, Sangrithi M, Niakan KK. 2021. Direct reprogramming of human embryonic to trophoblast stem cells. doi:10.1101/2021.08.18.456785

Fogarty NME, McCarthy A, Snijders KE, Powell BE, Kubikova N, Blakeley P, Lea R, Elder K, Wamaitha SE, Kim D, Maciulyte V, Kleinjung J, Kim J-S, Wells D, Vallier L, Bertero A, Turner JMA, Niakan KK. 2017. Genome editing reveals a role for OCT4 in human embryogenesis. *Nature* **550**:67–73. doi:10.1038/nature24033

Frum T, Halbisen MA, Wang C, Amiri H, Robson P, Ralston A. 2013. Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Dev Cell* **25**:610–622. doi:10.1016/j.devcel.2013.05.004

Frum T, Murphy TM, Ralston A. n.d. HIPPO signaling resolves embryonic cell fate conflicts during establishment of pluripotency in vivo. *eLife* **7**:e42298. doi:10.7554/eLife.42298

Gardner DK. 2016. The impact of physiological oxygen during culture, and vitrification for cryopreservation, on the outcome of extended culture in human IVF. *Reproductive BioMedicine Online* **32**:137–141. doi:10.1016/j.rbmo.2015.11.008

Gardner DK, Lane M, Stevens J, Schlenker T, Schoolcraft WB. 2000. Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer. *Fertility and Sterility* **73**:1155–1158. doi:10.1016/S0015-0282(00)00518-5

Gerri C, McCarthy A, Alanis-Lobato G, Demtschenko A, Bruneau A, Loubersac S, Fogarty NME, Hampshire D, Elder K, Snell P, Christie L, David L, Van de Velde H, Fouladi-Nashta AA, Niakan KK. 2020b. Initiation of a conserved trophectoderm program in human, cow and mouse embryos. *Nature* **587**:443–447. doi:10.1038/s41586-020-2759-x

Gerri C, McCarthy A, Scott GM, Regin M, Brumm S, Simon CS, Lee J, Montesinos C, Hassitt C, Hockenhull S, Hampshire D, Elder K, Snell P, Christie L, Fouladi-Nashta AA, Van de Velde H, Niakan KK. 2022. A conserved role of Hippo signaling in initiation of the first lineage specification event across mammals (preprint). Developmental Biology. doi:10.1101/2022.07.01.498418

Goedel A, Lanner F. 2021. A peek into the black box of human embryology. *Nature* **600**:223–224. doi:10.1038/d41586-021-03381-x

Goissis MD, Cibelli JB. 2014. Functional characterization of CDX2 during bovine preimplantation development in vitro. *Molecular Reproduction and Development* **81**:962–970. doi:10.1002/mrd.22415

Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND, Robson P. 2010. Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Developmental Cell* **18**:675–685. doi:10.1016/j.devcel.2010.02.012

Guo G, Stirparo GG, Strawbridge SE, Spindlow D, Yang J, Clarke J, Dattani A, Yanagida A, Li MA, Myers S, Özel BN, Nichols J, Smith A. 2021. Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* **28**:1040-1056.e6. doi:10.1016/j.stem.2021.02.025

Guo N, Li Y, Ai J, Gu L, Chen W, Liu Q. 2014. Two different concentrations of oxygen for culturing precompaction stage embryos on human embryo development competence: a prospective randomized sibling-oocyte study. *Int J Clin Exp Pathol* **7**:6191–6198.

Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, Soldner F, Cassady JP, Muffat J, Carey BW, Jaenisch R. 2010. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs 9222–9227.

Herbemont C, Labrosse J, Bennani-Smires B, Cedrin-Durnerin I, Peigne M, Sermondade N, Sarandi S, Vivot A, Vicaut E, Talib Z, Grynberg M, Sifer C. 2021. Impact of oxygen tension according to embryo stage of development: a prospective randomized study. *Sci Rep* **11**:22313. doi:10.1038/s41598-021-01488-9

Hertig AT, Rock J. 1949. Two human ova of the pre-villous stage, having a developmental age of about 8 and 9 days respectively. *Contrib Embryol* **33**:169–186.

Hertig AT, Rock J, Adams EC. 1956. A description of 34 human ova within the first 17 days of development. *American Journal of Anatomy* **98**:435–493. doi:10.1002/aja.1000980306

Home P, Ray S, Dutta D, Bronshteyn I, Larson M, Paul S. 2009. GATA3 Is Selectively Expressed in the Trophectoderm of Peri-implantation Embryo and Directly Regulates Cdx2 Gene Expression*. *Journal of Biological Chemistry* **284**:28729–28737. doi:10.1074/jbc.M109.016840

Hrdlickova R, Toloue M, Tian B. 2017. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip Rev RNA* **8**. doi:10.1002/wrna.1364

Hubert MA, Sherritt SL, Bachurski CJ, Handwerger S. 2010. Involvement of transcription factor NR2F2 in human trophoblast differentiation. *PLoS One* **5**:e9417. doi:10.1371/journal.pone.0009417

Hurlbut JB, Hyun I, Levine AD, Lovell-Badge R, Lunshof JE, Matthews KRW, Mills P, Murdoch A, Pera MF, Scott CT, Tizzard J, Warnock M, Zernicka-Goetz M, Zhou Q, Zoloth L. 2017. Revisiting the Warnock rule. *Nat Biotechnol* **35**:1029–1042. doi:10.1038/nbt.4015

Hyun I, Wilkerson A, Johnston J. 2016. Embryology policy: Revisit the 14-day rule. *Nature* **533**:169–171. doi:10.1038/533169a

Io S, Kabata M, Iemura Y, Semi K, Morone N, Minagawa A, Wang B, Okamoto I, Nakamura T, Kojima Y, Iwatani C, Tsuchiya H, Kaswandy B, Kondoh E, Kaneko S, Woltjen K, Saitou M, Yamamoto T, Mandai M, Takashima Y. 2021. Capturing human trophoblast development with naive pluripotent stem cells in vitro. *Cell Stem Cell* **28**:1023-1039.e13. doi:10.1016/j.stem.2021.03.013

Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lönnerberg P, Linnarsson S. 2014. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* **11**:163–166. doi:10.1038/nmeth.2772

James D, Levine AJ, Besser D, Hemmati-Brivanlou A. 2005. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* **132**:1273–1282. doi:10.1242/dev.01706

James JL, Carter AM, Chamley LW. 2012. Human placentation from nidation to 5 weeks of gestation. Part I: What do we know about formative placental development following implantation? *Placenta* **33**:327–334. doi:10.1016/j.placenta.2012.01.020

Jiang R, Sun T, Song D, Li JJ. 2022. Statistics or biology: the zero-inflation controversy about scRNA-seq data. *Genome Biol* **23**:31. doi:10.1186/s13059-022-02601-5

Johnson MH, Ziomek CA. 1981. The foundation of two distinct cell lineages within the mouse morula. *Cell* **24**:71–80. doi:10.1016/0092-8674(81)90502-x

Kagawa H, Javali A, Khoei HH, Sommer TM, Sestini G, Novatchkova M, Scholte op Reimer Y, Castel G, Bruneau A, Maenhoudt N, Lammers J, Loubersac S, Freour T, Vankelecom H, David L, Rivron N. 2021. Human blastoids model blastocyst development and implantation. *Nature* 1–6. doi:10.1038/s41586-021-04267-8

Kang M, Piliszek A, Artus J, Hadjantonakis A-K. 2013. FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Development* **140**:267–279. doi:10.1242/dev.084996

Kea B, Gebhardt J, Watt J, Westphal LM, Lathi RB, Milki AA, Behr B. 2007. Effect of reduced oxygen concentrations on the outcome of in vitro fertilization. *Fertility and Sterility* **87**:213–216. doi:10.1016/j.fertnstert.2006.05.066

Kigawa J. [Studies on the levels of pO2 and pCO2 in the uterine cavity and uterine tissue (author's transl)]. Nihon Sanka Fujinka Gakkai Zasshi. 1981 Oct;33(10):1646-54. Japanese. PMID: 6796630.

Kirkegaard K, Sundvall L, Erlandsen M, Hindkjær JJ, Knudsen UB, Ingerslev HJ. 2016. Timing of human preimplantation embryonic development is confounded by embryo origin. *Human Reproduction* **31**:324–331. doi:10.1093/humrep/dev296

Korotkevich E, Niwayama R, Courtois A, Friese S, Berger N, Buchholz F, Hiiragi T. 2017. The Apical Domain Is Required and Sufficient for the First Lineage Segregation in the Mouse Embryo. *Dev Cell* **40**:235-247.e7. doi:10.1016/j.devcel.2017.01.006

Kovacic B, Sajko MC, Vlaisavljević V. 2010. A prospective, randomized trial on the effect of atmospheric versus reduced oxygen concentration on the outcome of intracytoplasmic sperm injection cycles. *Fertil Steril* **94**:511–519. doi:10.1016/j.fertnstert.2009.03.077

Kuijk EW, van Tol LTA, Van de Velde H, Wubbolts R, Welling M, Geijsen N, Roelen BAJ. 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. doi:10.1242/dev.071688

La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, Fan J, Borm LE, Liu Z, van Bruggen D, Guo J, He X, Barker R, Sundström E, Castelo-Branco G, Cramer P, Adameyko I, Linnarsson S, Kharchenko PV. 2018. RNA velocity of single cells. *Nature* **560**:494–498. doi:10.1038/s41586-018-0414-6

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**:559. doi:10.1186/1471-2105-9-559

Le Guével R, Oger F, Martinez-Jimenez CP, Bizot M, Gheeraert C, Firmin F, Ploton M, Kretova M, Palierne G, Staels B, Barath P, Talianidis I, Lefebvre P, Eeckhoute J, Salbert G. 2017. Inactivation of the Nuclear Orphan Receptor COUP-TFII by Small Chemicals. *ACS Chem Biol* **12**:654–663. doi:10.1021/acschembio.6b00593

Li X, Wang C-Y. 2021. From bulk, single-cell to spatial RNA sequencing. *Int J Oral Sci* **13**:1–6. doi:10.1038/s41368-021-00146-0

Liang G, Zhang Y. 2013. Embryonic stem cell and induced pluripotent stem cell: an epigenetic perspective. *Cell Res* **23**:49–69. doi:10.1038/cr.2012.175

Lin S-CJ, Wani MA, Whitsett JA, Wells JM. 2010. Klf5 regulates lineage formation in the pre-implantation mouse embryo. *Development* **137**:3953–3963. doi:10.1242/dev.054775

Liu S, Bou G, Sun R, Guo S, Xue B, Wei R, Cooney AJ, Liu Z. 2015. Sox2 is the faithful marker for pluripotency in pig: Evidence from embryonic studies. *Developmental Dynamics* **244**:619–627. doi:10.1002/dvdy.24248

Liu X, Ouyang JF, Rossello FJ, Tan JP, Davidson KC, Valdes DS, Schröder J, Sun YBY, Chen J, Knaupp AS, Sun G, Chy HS, Huang Z, Pflueger J, Firas J, Tano V, Buckberry S, Paynter JM, Larcombe MR, Poppe D, Choo XY, O'Brien CM, Pastor WA, Chen D, Leichter AL, Naeem H, Tripathi P, Das PP, Grubman A, Powell DR, Laslett AL, David L, Nilsson SK, Clark AT, Lister R, Nefzger CM, Martelotto LG, Rackham OJL, Polo JM. 2020. Reprogramming roadmap reveals route to human induced trophoblast stem cells. *Nature* **586**:101–107. doi:10.1038/s41586-020-2734-6

Liu X, Tan JP, Schröder J, Aberkane A, Ouyang JF, Mohenska M, Lim SM, Sun YBY, Chen J, Sun G, Zhou Y, Poppe D, Lister R, Clark AT, Rackham OJL, Zenker J, Polo JM. 2021. Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* **591**:627–632. doi:10.1038/s41586-021-03372-y

Lovell-Badge R, Anthony E, Barker RA, Bubela T, Brivanlou AH, Carpenter M, Charo RA, Clark A, Clayton E, Cong Y, Daley GQ, Fu J, Fujita M, Greenfield A, Goldman SA, Hill L, Hyun I, Isasi R, Kahn J, Kato K, Kim J-S, Kimmelman J, Knoblich JA, Mathews D, Montserrat N, Mosher J, Munsie M, Nakauchi H, Naldini L, Naughton G, Niakan K, Ogbogu U, Pedersen R, Rivron N, Rooke H, Rossant J, Round J, Saitou M, Sipp D, Steffann J, Sugarman J, Surani A, Takahashi J, Tang F, Turner L, Zettler PJ, Zhai X. 2021. ISSCR Guidelines for Stem Cell Research and Clinical Translation: The 2021 update. *Stem Cell Reports* **16**:1398–1408. doi:10.1016/j.stemcr.2021.05.012

Ma H, Zhai J, Wan H, Jiang X, Wang X, Wang L, Xiang Y, He X, Zhao Z-A, Zhao B, Zheng P, Li L, Wang H. 2019. In vitro culture of cynomolgus monkey embryos beyond early gastrulation. *Science* **366**:eaax7890. doi:10.1126/science.aax7890

Ma S, Peng Y, Hu L, Wang X, Xiong Y, Tang Y, Tan J, Gong F. 2022. Comparisons of benefits and risks of single embryo transfer versus double embryo transfer: a systematic review and meta-analysis. *Reproductive Biology and Endocrinology* **20**:20. doi:10.1186/s12958-022-00899-1

Madeja ZE, Sosnowski J, Hryniewicz K, Warzych E, Pawlak P, Rozwadowska N, Plusa B, Lechniak D. 2013. Changes in sub-cellular localisation of trophoblast and inner cell mass specific transcription factors during bovine preimplantation development. *BMC Developmental Biology* **13**:32. doi:10.1186/1471-213X-13-32

Makhijani R, Grow DR. 2020. Donor egg is the best second choice for many infertile couples: real progress in overcoming age-related fertility is not here yet. *J Assist Reprod Genet* **37**:1589–1591. doi:10.1007/s10815-020-01880-w

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences* **78**:7634–7638. doi:10.1073/pnas.78.12.7634

Mastroianni L, Jones R. 1965. OXYGEN TENSION WITHIN THE RABBIT FALLOPIAN TUBE. *J Reprod Fertil* **9**:99–102. doi:10.1530/jrf.0.0090099

McConaha ME, Eckstrum K, An J, Steinle JJ, Bany BM. 2011. Microarray assessment of the influence of the conceptus on gene expression in the mouse uterus during decidualization. *Reproduction* **141**:511–527. doi:10.1530/REP-10-0358

Meintjes M, Chantilis SJ, Douglas JD, Rodriguez AJ, Guerami AR, Bookout DM, Barnett BD, Madden JD. 2009. A controlled randomized trial evaluating the effect of lowered incubator oxygen tension on live births in a predominantly blastocyst transfer program. *Hum Reprod* **24**:300–307. doi:10.1093/humrep/den368

Meistermann D, Bruneau A, Loubersac S, Reignier A, Firmin J, François-Campion V, Kilens S, Lelièvre Y, Lammers J, Feyeux M, Hulin P, Nedellec S, Bretin B, Castel G, Allègre N, Covin S, Bihouée A, Soumillon M, Mikkelsen T, Barrière P, Chazaud C, Chappell J, Pasque V, Bourdon J, Fréour T, David L. 2021. Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification. *Cell Stem Cell* S1934590921001855. doi:10.1016/j.stem.2021.04.027

Mekhoubad S, Bock C, de Boer AS, Kiskinis E, Meissner A, Eggan K. 2012. Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* **10**:595–609. doi:10.1016/j.stem.2012.02.014

Molè MA, Coorens THH, Shahbazi MN, Weberling A, Weatherbee BAT, Gantner CW, Sancho-Serra C, Richardson L, Drinkwater A, Syed N, Engley S, Snell P, Christie L, Elder K, Campbell A, Fishel S, Behjati S, Vento-Tormo R, Zernicka-Goetz M. 2021. A single cell characterisation of human embryogenesis identifies pluripotency transitions and putative anterior hypoblast centre. *Nat Commun* **12**:3679. doi:10.1038/s41467-021-23758-w

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**:621–628. doi:10.1038/nmeth.1226

Nardelli AA, Stafinski T, Motan T, Klein K, Menon D. 2014. Assisted reproductive technologies (ARTs): Evaluation of evidence to support public policy development. *Reproductive Health* **11**:76. doi:10.1186/1742-4755-11-76

Narsinh KH, Plews J, Wu JC. 2011. Comparison of Human Induced Pluripotent and Embryonic Stem Cells: Fraternal or Identical Twins? *Mol Ther* **19**:635–638. doi:10.1038/mt.2011.41

Ng KYB, Mingels R, Morgan H, Macklon N, Cheong Y. 2018. In vivo oxygen, temperature and pH dynamics in the female reproductive tract and their importance in human conception: a systematic review. *Human Reproduction Update* **24**:15–34. doi:10.1093/humupd/dmx028

Niakan KK, 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. doi:10.1016/j.ydbio.2012.12.008

Niakan KK, Han J, Pedersen RA, Simon C, Pera RAR. 2012. Human pre-implantation embryo development. *Development* **139**:829–841. doi:10.1242/dev.060426

Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Schöler H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 1998 Oct 30;95(3):379-91. doi: 10.1016/s0092-8674(00)81769-9. PMID: 9814708.

Nichols J, Silva J, Roode M, Smith A. 2009. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development* **136**:3215–3222. doi:10.1242/dev.038893

Nichols J, Smith A. 2009. Naive and Primed Pluripotent States. *Cell Stem Cell* **4**:487–492. doi:10.1016/j.stem.2009.05.015

Nishioka N, Inoue K, Adachi K, Kiyonari H, Ota M, Ralston A, Yabuta N, Hirahara S, Stephenson RO, Ogonuki N, Makita R, Kurihara H, Morin-Kensicki EM, Nojima H, Rossant J, Nakao K, Niwa H, Sasaki H. 2009. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev Cell* **16**:398–410. doi:10.1016/j.devcel.2009.02.003

Nishioka N, Yamamoto S, Kiyonari H, Sato H, Sawada A, Ota M, Nakao K, Sasaki H. 2008. Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech Dev* **125**:270–283. doi:10.1016/j.mod.2007.11.002

Ojosnegros S, Seriola A, Godeau AL, Veiga A. 2021. Embryo implantation in the laboratory: an update on current techniques. *Human Reproduction Update* **27**:501–530. doi:10.1093/humupd/dmaa054

Okae H, Toh H, Sato T, Hiura H, Takahashi S, Shirane K, Kabayama Y, Suyama M, Sasaki H, Arima T. 2018. Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* **22**:50-63.e6. doi:10.1016/j.stem.2017.11.004

Oshlack A, Wakefield MJ. 2009. Transcript length bias in RNA-seq data confounds systems biology. *Biology Direct* **4**:14. doi:10.1186/1745-6150-4-14

Ottosen LDM, Hindkaer J, Husth M, Petersen DE, Kirk J, Ingerslev HJ. 2006. Observations on intrauterine oxygen tension measured by fibre-optic microsensors. *Reprod Biomed Online* **13**:380–385. doi:10.1016/s1472-6483(10)61443-5

Parisi S, Russo T. 2011. Chapter Eighteen - Regulatory Role of Klf5 in Early Mouse Development and in Embryonic Stem Cells In: Litwack G, editor. Vitamins & Hormones, Stem Cell Regulators. Academic Press. pp. 381–397. doi:10.1016/B978-0-12-386015-6.00037-8

Petropoulos S, Edsgärd D, Reinius B, Deng Q, Panula SP, Codeluppi S, Plaza Reyes A, Linnarsson S, Sandberg R, Lanner F. 2016a. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165**:1012–1026. doi:10.1016/j.cell.2016.03.023

Plusa B. 2005. Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. *Journal of Cell Science* **118**:505–515. doi:10.1242/jcs.01666

Plusa B, Piliszek A, Frankenberg S, Artus J, Hadjantonakis A-K. 2008. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135**:3081–3091. doi:10.1242/dev.021519

Poli M, Ori A, Child T, Jaroudi S, Spath K, Beck M, Wells D. 2015. Characterization and quantification of proteins secreted by single human embryos prior to implantation. *EMBO Mol Med* **7**:1465–1479. doi:10.15252/emmm.201505344

Polvani S, Pepe S, Milani S, Galli A. 2019. COUP-TFII in Health and Disease. *Cells* **9**:101. doi:10.3390/cells9010101

Posfai E, Petropoulos S, de Barros FRO, Schell JP, Jurisica I, Sandberg R, Lanner F, Rossant J. 2017. Position- and Hippo signaling-dependent plasticity during lineage segregation in the early mouse embryo. *eLife* **6**:e22906. doi:10.7554/eLife.22906

Pribenszky C, Nilselid A-M, Montag M. 2017. Time-lapse culture with morphokinetic embryo selection improves pregnancy and live birth chances and reduces early pregnancy loss: a meta-analysis. *Reprod Biomed Online* **35**:511–520. doi:10.1016/j.rbmo.2017.06.022

Radley A, Corujo-Simon E, Nichols J, Smith A, Dunn S-J. 2022. Entropy sorting of single-cell RNA sequencing data reveals the inner cell mass in the human pre-implantation embryo. *Stem Cell Reports* **0**. doi:10.1016/j.stemcr.2022.09.007

Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S. 2006. Stochastic mRNA Synthesis in Mammalian Cells. *PLoS Biol* **4**:e309. doi:10.1371/journal.pbio.0040309

Ralston A, Cox BJ, Nishioka N, Sasaki H, Chea E, Rugg-Gunn P, Guo G, Robson P, Draper JS, Rossant J. 2010. Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development* **137**:395–403. doi:10.1242/dev.038828

Regin M, Essahib W, Demtschenko A, Dewandre D, David L, Gerri C, Niakan K, Verheyen G, Tournaye H, Sterckx J, Sermon K, Velde HV de. 2022. Lineage segregation in human pre-implantation embryos is specified by YAP1 and TEAD1. doi:10.1101/2022.09.29.509946

Reignier A, Lefebvre T, Loubersac S, Lammers J, Barriere P, Freour T. 2021. Time-lapse technology improves total cumulative live birth rate and shortens time to live birth as compared to conventional incubation system in couples undergoing ICSI. *J Assist Reprod Genet* **38**:917–923. doi:10.1007/s10815-021-02099-z

Rivron NC, Frias-Aldeguer J, Vrij EJ, Boisset J-C, Korving J, Vivié J, Truckenmüller RK, van Oudenaarden A, van Blitterswijk CA, Geijsen N. 2018. Blastocyst-like structures generated solely from stem cells. *Nature* **557**:106–111. doi:10.1038/s41586-018-0051-0

Roode M, Blair K, Snell P, Elder K, Marchant S, Smith A, Nichols J. 2012. Human hypoblast formation is not dependent on FGF signalling. *Dev Biol* **361**:358–363. doi:10.1016/j.ydbio.2011.10.030

Saelens W, Cannoodt R, Todorov H, Saeys Y. 2019. A comparison of single-cell trajectory inference methods. *Nat Biotechnol* **37**:547–554. doi:10.1038/s41587-019-0071-9

Saiz N, Williams KM, Seshan VE, Hadjantonakis A-K. 2016. Asynchronous fate decisions by single cells collectively ensure consistent lineage composition in the mouse blastocyst. *Nat Commun* **7**:13463. doi:10.1038/ncomms13463

Schrode N, Xenopoulos P, Piliszek A, Frankenberg S, Plusa B, Hadjantonakis A-K. 2013. Anatomy of a blastocyst: cell behaviors driving cell fate choice and morphogenesis in the early mouse embryo. *Genesis* **51**:219–233. doi:10.1002/dvg.22368

Sermon K, Capalbo A, Cohen J, Coonen E, De Rycke M, De Vos A, Delhanty J, Fiorentino F, Gleicher N, Griesinger G, Grifo J, Handyside A, Harper J, Kokkali G, Mastenbroek S, Meldrum D, Meseguer M, Montag M, Munné S, Rienzi L, Rubio C, Scott K, Scott R, Simon C, Swain J, Treff N, Ubaldi F, Vassena R, Vermeesch JR, Verpoest W, Wells D, Geraedts J. 2016. The why, the how and the when of PGS 2.0: current practices and expert opinions of fertility specialists, molecular biologists, and embryologists. *Mol Hum Reprod* **22**:845–857. doi:10.1093/molehr/gaw034

Shahbazi MN. 2020. Mechanisms of human embryo development: from cell fate to tissue shape and back. *Development* **147**:dev190629. doi:10.1242/dev.190629

Shahbazi MN, Jedrusik A, Vuoristo S, Recher G, Hupalowska A, Bolton V, Fogarty NNM, Campbell A, Devito L, Ilic D, Khalaf Y, Niakan KK, Fishel S, Zernicka-Goetz M. 2016. Self-organisation of the human embryo in the absence of maternal tissues. *Nat Cell Biol* **18**:700–708. doi:10.1038/ncb3347

Singh AM, Hamazaki T, Hankowski KE, Terada N. 2007. A Heterogeneous Expression Pattern for Nanog in Embryonic Stem Cells. *Stem Cells* **25**:2534–2542. doi:10.1634/stemcells.2007-0126

Siriwardena D, Boroviak TE. 2022. Evolutionary divergence of embryo implantation in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences* **377**:20210256. doi:10.1098/rstb.2021.0256

Slade P, O'Neill C, Simpson AJ, Lashen H. 2007. The relationship between perceived stigma, disclosure patterns, support and distress in new attendees at an infertility clinic. *Hum Reprod* **22**:2309–2317. doi:10.1093/humrep/dem115

Smith AG. 2001. Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol* **17**:435–462. doi:10.1146/annurev.cellbio.17.1.435

Smith ZD, Chan MM, Humm KC, Karnik R, Mekhoubad S, Regev A, Eggan K, Meissner A. 2014. DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**:611–615. doi:10.1038/nature13581

Sozen B, Jorgensen V, Weatherbee BAT, Chen S, Zhu M, Zernicka-Goetz M. 2021. Reconstructing aspects of human embryogenesis with pluripotent stem cells. *Nat Commun* **12**:5550. doi:10.1038/s41467-021-25853-4

Stamatiadis P, Cosemans G, Boel A, Menten B, De Sutter P, Stoop D, Chuva de Sousa Lopes SM, Lluis F, Coucke P, Heindryckx B. 2022. TEAD4 regulates trophectoderm differentiation upstream of CDX2 in a GATA3-independent manner in the human preimplantation embryo. *Human Reproduction* deac138. doi:10.1093/humrep/deac138

Strumpf D, Mao C-A, Yamanaka Y, Ralston A, Chawengsaksophak K, Beck F, Rossant J. 2005. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132**:2093–2102. doi:10.1242/dev.01801

Sun H, Gong T-T, Jiang Y-T, Zhang S, Zhao Y-H, Wu Q-J. 2019. Global, regional, and national prevalence and disability-adjusted life-years for infertility in 195 countries and territories, 1990–2017: results from a global burden of disease study, 2017. *Aging* **11**:10952–10991. doi:10.18632/aging.102497

Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**:472–474. doi:10.1126/science.1198817

Tajik M, Baharfar M, Donald WA. 2022. Single-cell mass spectrometry. *Trends in Biotechnology* **40**:1374–1392. doi:10.1016/j.tibtech.2022.04.004

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. 2007. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**:861–872. doi:10.1016/j.cell.2007.11.019

Takashima Y, Guo G, Loos R, Nichols J, Ficz G, Krueger F, Oxley D, Santos F, Clarke J, Mansfield W, Reik W, Bertone P, Smith A. 2014. Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell* **158**:1254–1269. doi:10.1016/j.cell.2014.08.029

Tanaka S, Kunath T, Hadjantonakis A-K, Nagy A, Rossant J. 1998. Promotion of Trophoblast Stem Cell Proliferation by FGF4. *Science* **282**:2072–2075. doi:10.1126/science.282.5396.2072

Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, Lao K, Surani MA. 2009. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**:377–382. doi:10.1038/nmeth.1315

Tao J, Tamis R, Fink K, Williams B, Nelson-White T, Craig R. 2002. The neglected morula/compact stage embryo transfer. *Human Reproduction* **17**:1513–1518. doi:10.1093/humrep/17.6.1513

Tarazi S, Aguilera-Castrejon A, Joubran C, Ghanem N, Ashouokhi S, Roncato F, Wildschutz E, Haddad M, Oldak B, Gomez-Cesar E, Livnat N, Viukov S, Lokshtanov D, Naveh-Tassa S, Rose M, Hanna S, Raanan C, Brenner O, Kedmi M, Keren-Shaul H, Lapidot T, Maza I, Novershtern N, Hanna JH. 2022. Post-gastrulation synthetic embryos generated ex utero from mouse naive ESCs. *Cell* **185**:3290-3306.e25. doi:10.1016/j.cell.2022.07.028

Teklenburg G, Macklon NS. 2009. In Vitro Models for the Study of Early Human Embryo-Endometrium Interactions. *Reprod Sci* **16**:811–818. doi:10.1177/1933719109334966

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RDG. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells 196–199.

Kilens S, Meistermann D, Moreno D, Chariau C, Gaignerie A, Reignier A, Lelièvre Y, Casanova M, Vallot C, Nedellec S, Flippe L, Firmin J, Song J, Charpentier E, Lammers J, Donnart A, Marec N, Deb W, Bihouée A, Le Caignec C, Pecqueur C, Redon R, Barrière P, Bourdon J, Pasque V, Soumillon M, Mikkelsen TS, Rougeulle C, Fréour T, David L. 2018. Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nature Communications* **9**. doi:10.1038/s41467-017-02107-w

Theunissen TW, Powell BE, Wang H, Mitalipova M, Faddah DA, Reddy J, Fan ZP, Maetzel D, Ganz K, Shi L, Lungjangwa T, Imsoonthornruksa S, Stelzer Y, Rangarajan S, D'Alessio A, Zhang J, Gao Q, Dawlaty MM, Young RA, Gray NS, Jaenisch R. 2014. Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell* **15**:471–487. doi:10.1016/j.stem.2014.07.002

Thompson SM, Onwubalili N, Brown K, Jindal SK, McGovern PG. 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. *J Assist Reprod Genet* **30**:1577–1581. doi:10.1007/s10815-013-0100-4

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. 1998. Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* **282**:1145–1147. doi:10.1126/science.282.5391.1145

Treff NR, Scott RT. 2013. Four-hour quantitative real-time polymerase chain reaction-based comprehensive chromosome screening and accumulating evidence of accuracy, safety, predictive value, and clinical efficacy. *Fertil Steril* **99**:1049–1053. doi:10.1016/j.fertnstert.2012.11.007

Turco MY, Moffett A. 2019. Development of the human placenta. *Development* **146**. doi:10.1242/dev.163428

Unterholzner L, Keating SE, Baran M, Horan KA, Jensen SB, Sharma S, Sirois CM, Jin T, Latz E, Xiao TS, Fitzgerald KA, Paludan SR, Bowie AG. 2010. IFI16 is an innate immune sensor for intracellular DNA. *Nat Immunol* **11**:997–1004. doi:10.1038/ni.1932

Vallier L, Alexander M, Pedersen RA. 2005. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* **118**:4495–4509. doi:10.1242/jcs.02553

Van de Velde H, Cauffman G, Tournaye H, Devroey P, Liebaers I. 2008. The four blastomeres of a 4-cell stage human embryo are able to develop individually into blastocysts with inner cell mass and trophectoderm. *Hum Reprod* **23**:1742–1747. doi:10.1093/humrep/den190

Van den Berge K, Roux de Bézieux H, Street K, Saelens W, Cannoodt R, Saeys Y, Dudoit S, Clement L. 2020. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat Commun* **11**:1201. doi:10.1038/s41467-020-14766-3

Vander Borght M, Wyns C. 2018. Fertility and infertility: Definition and epidemiology. *Clinical Biochemistry*, The Role of Biomarkers in Reproductive Health **62**:2–10. doi:10.1016/j.clinbiochem.2018.03.012

Vassena R, Boué S, González-Roca E, Aran B, Auer H, Veiga A, Izpisua Belmonte JC. 2011. Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Development* **138**:3699–3709. doi:10.1242/dev.064741

Veiga A, Calderon G, Barri PN, Coroleu B. 1987. Pregnancy after the replacement of a frozen-thawed embryo with >50% intact blastomeres. *Human Reproduction* **2**:321–323. doi:10.1093/oxfordjournals.humrep.a136542

Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, Park J-E, Stephenson E, Polański K, Goncalves A, Gardner L, Holmqvist S, Henriksson J, Zou A, Sharkey AM, Millar B, Innes B, Wood L, Wilbrey-Clark A, Payne RP, Ivarsson MA, Lisgo S, Filby A, Rowitch DH, Bulmer JN, Wright GJ, Stubbington MJT, Haniffa M, Moffett A, Teichmann SA. 2018. Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature* **563**:347–353. doi:10.1038/s41586-018-0698-6

Viukov S, Shani T, Bayerl J, Aguilera-Castrejon A, Oldak B, Sheban D, Tarazi S, Stelzer Y, Hanna JH, Novershtern N. 2022. Human primed and naïve PSCs are both able to differentiate into trophoblast stem cells. *Stem Cell Reports*. doi:10.1016/j.stemcr.2022.09.008

Warnock, M. et al. *Report of the Committee of Inquiry into Human Fertilisation and Embryology* (Her Majesty's Stationery Office, London, UK, 1984).

Waldenström U, Engström A-B, Hellberg D, Nilsson S. 2009. Low-oxygen compared with high-oxygen atmosphere in blastocyst culture, a prospective randomized study. *Fertil Steril* **91**:2461–2465. doi:10.1016/j.fertnstert.2008.03.051

Wamaitha SE, Grybel KJ, Alanis-Lobato G, Gerri C, Ogushi S, McCarthy A, Mahadevaiah SK, Healy L, Lea RA, Molina-Arcas M, Devito LG, Elder K, Snell P, Christie L, Downward J, Turner JMA, Niakan KK. 2020. IGF1-mediated human embryonic stem cell self-renewal recapitulates the embryonic niche. *Nat Commun* **11**:764. doi:10.1038/s41467-020-14629-x

Wamaitha SE, Valle I del, Cho LTY, Wei Y, Fogarty NME, Blakeley P, Sherwood RI, Ji H, Niakan KK. 2015. Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes Dev* **29**:1239–1255. doi:10.1101/gad.257071.114

Wang W, Vilella F, Alama P, Moreno I, Mignardi M, Isakova A, Pan W, Simon C, Quake SR. 2020. Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. *Nat Med* **26**:1644–1653. doi:10.1038/s41591-020-1040-z

Wei X, Ye J, Shang Y, Chen H, Liu S, Liu L, Wang R. 2017. Ascl2 activation by YAP1/KLF5 ensures the self-renewability of colon cancer progenitor cells. *Oncotarget* **8**:109301–109318. doi:10.18632/oncotarget.22673

Weinberger L, Ayyash M, Novershtern N, Hanna JH. 2016. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol* **17**:155–169. doi:10.1038/nrm.2015.28

Weltner J, Lanner F. 2021. Refined transcriptional blueprint of human preimplantation embryos. *Cell Stem Cell* **28**:1503–1504. doi:10.1016/j.stem.2021.08.011

Wu AK, Elliott P, Katz PP, Smith JF. 2013. Time costs of fertility care: the hidden hardship of building a family. *Fertil Steril* **99**:2025–2030. doi:10.1016/j.fertnstert.2013.01.145

Wu S-P, Yu C-T, Tsai SY, Tsai M-J. 2016. Choose your destiny: Make a cell fate decision with COUP-TFII. *The Journal of Steroid Biochemistry and Molecular Biology* **157**:7–12. doi:10.1016/j.jsbmb.2015.11.011

Xiang L, Yin Y, Zheng Y, Ma Y, Li Y, Zhao Z, Guo J, Ai Z, Niu Y, Duan K, He J, Ren S, Wu D, Bai Y, Shang Z, Dai X, Ji W, Li T. 2020. A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* **577**:537–542. doi:10.1038/s41586-019-1875-y

Yagi R, Kohn MJ, Karavanova I, Kaneko KJ, Vullhorst D, DePamphilis ML, Buonanno A. 2007. Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development* **134**:3827–3836. doi:10.1242/dev.010223

Yamanaka Y, Lanner F, Rossant J. 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**:715–724. doi:10.1242/dev.043471

Yamanaka Y, Ralston A, Stephenson RO, Rossant J. 2006. Cell and molecular regulation of the mouse blastocyst. *Developmental Dynamics* **235**:2301–2314. doi:10.1002/dvdy.20844

Yan L, Yang M, Guo H, Yang L, Wu J, Li Rong, Liu P, Lian Y, Zheng X, Yan J, Huang J, Li M, Wu X, Wen L, Lao K, Li Ruiqiang, Qiao J, Tang F. 2013. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* **20**:1131–1139. doi:10.1038/nsmb.2660

Yanagida A, Corujo-Simon E, Revell CK, Sahu P, Stirparo GG, Aspalter IM, Winkel AK, Peters R, De Belly H, Cassani DAD, Achouri S, Blumenfeld R, Franze K, Hannezo E, Paluch EK, Nichols J, Chalut KJ. 2022. Cell surface fluctuations regulate early embryonic lineage sorting. *Cell* **185**:777-793.e20. doi:10.1016/j.cell.2022.01.022

Yanagida A, Spindlow D, Nichols J, Dattani A, Smith A, Guo G. 2021. Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* **28**:1016-1022.e4. doi:10.1016/j.stem.2021.04.031

Yedwab GA, Paz G, Homonnai TZ, David MP, Kraicer PF. 1976. The temperature, pH, and partial pressure of oxygen in the cervix and uterus of women and uterus of rats during the cycle. *Fertil Steril* **27**:304–309. doi:10.1016/s0015-0282(16)41722-x

Ying Q-L, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. 2008. The ground state of embryonic stem cell self-renewal. *Nature* **453**:519–523. doi:10.1038/nature06968

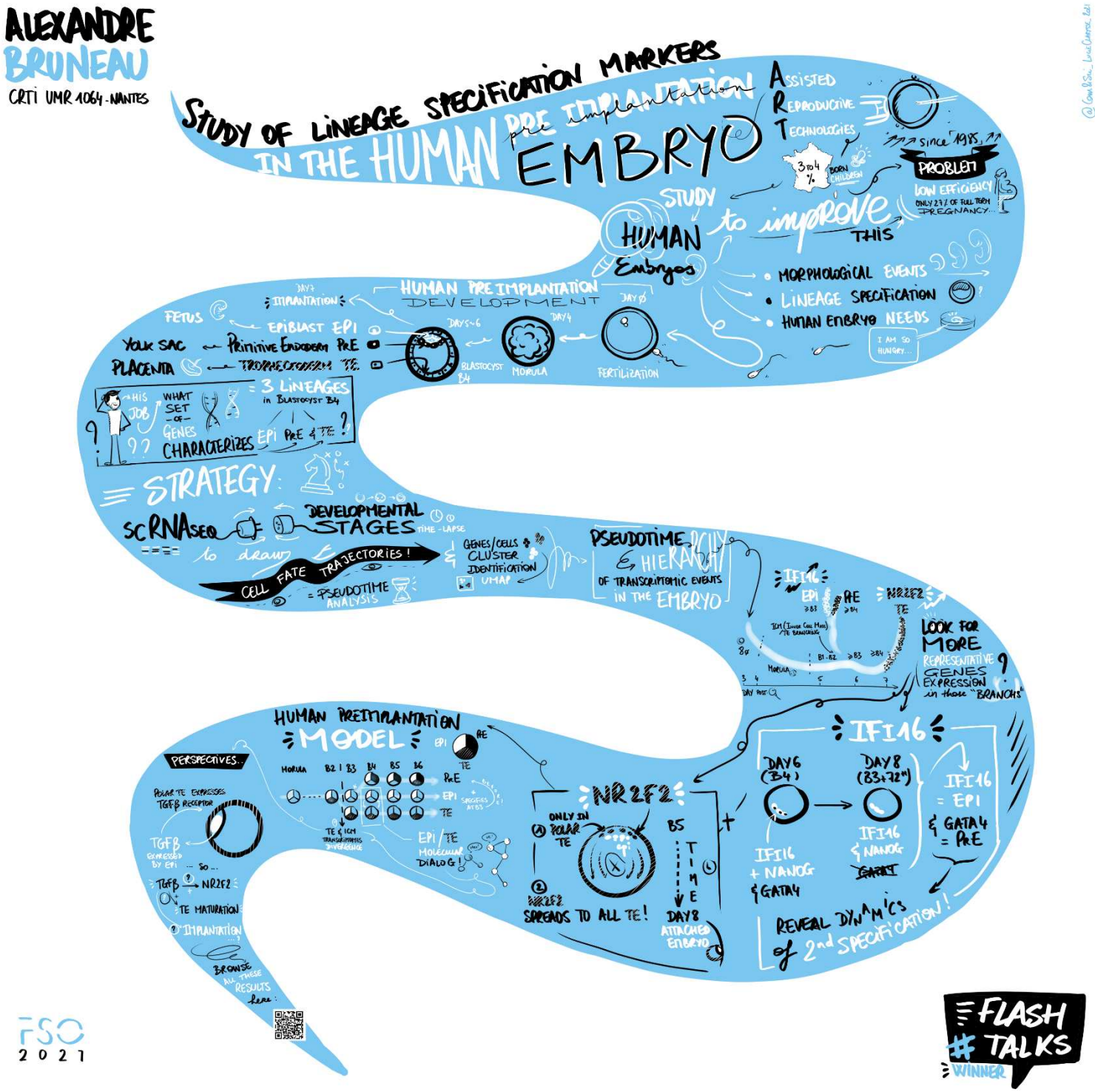
Yu F-X, Zhao B, Guan K-L. 2015. Hippo Pathway in Organ Size Control, Tissue Homeostasis, and Cancer. *Cell* **163**:811–828. doi:10.1016/j.cell.2015.10.044

Yu L, Wei Y, Duan J, Schmitz DA, Sakurai M, Wang L, Wang K, Zhao S, Hon GC, Wu J. 2021. Blastocyst-like structures generated from human pluripotent stem cells. *Nature* **591**:620–626. doi:10.1038/s41586-021-03356-y

Zhao Y-Y, Yu Y, Zhang X-W. 2018. Overall Blastocyst Quality, Trophectoderm Grade, and Inner Cell Mass Grade Predict Pregnancy Outcome in Euploid Blastocyst Transfer Cycles. *Chin Med J (Engl)* **131**:1261–1267. doi:10.4103/0366-6999.232808

Zhi X, Zhao D, Zhou Z, Liu R, Chen C. 2012. YAP promotes breast cell proliferation and survival partially through stabilizing the KLF5 transcription factor. *Am J Pathol* **180**:2452–2461. doi:10.1016/j.ajpath.2012.02.025

Zhou F, Wang R, Yuan P, Ren Y, Mao Y, Li R, Lian Y, Li J, Wen L, Yan L, Qiao J, Tang F. 2019. Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature* **572**:660–664. doi:10.1038/s41586-019-1500-0



Crédit : Lucie Clarisse

DOCTORAT BIOLOGIE BRETAGNE SANTE LOIRE

Nantes Université

Titre : Étude de la spécification des lignées cellulaires lors du développement péri-implantatoire humain

Mots clés : embryon humain, spécification, procréation médicalement assistée, bio-informatique

Résumé : L'embryon humain se développe suite à la fusion de l'ovocyte et du spermatozoïde. En quelque jours seulement, l'embryon passe d'une cellule unique à une structure multicellulaire complexe en trois dimensions, asymétrique, et dont les mécanismes de développement et de régulation sont encore largement méconnus.

Ce manque de connaissance fait notamment défaut en procréation médicalement assistée (AMP) pour les couples infertiles, où les taux de grossesse sont de moins de 30%.

Afin de mieux comprendre comment l'embryon humain se développe, nous avons mené un projet de recherche pluridisciplinaire melant bioinformatique, expertise hospitalière, et recherche fondamentale en biologie du développement.

Nous avons reconstitué le continuum transcriptomique des cellules de l'embryon humain au cours de ses premiers jours de développement.
Nous montrons que l'émergence de la masse cellulaire interne (ICM, futur foetus) et du trophectoderme (TE, futur placenta) a lieu peu après la formation du blastocyste, et non au stade morula comme chez la souris.
Nous montrons également que le trophectoderme mature plus vite du côté polaire, côté où il s'implante dans l'utérus.
Enfin, nous présentons de nouveaux marqueurs des lignées cellulaires de l'embryon avant et après son implantation *in vitro*.

Ce travail collaboratif ouvre de nouvelles perspectives de recherche dans l'étude du développement humain, de l'AMP, et la médecine régénérative.

Title : Study of cell lineage specification occuring during human peri-implantation development

Keywords : human embryo, specification, assisted reproductive technologies, bioinformatics

Abstract : The human embryo develops as a result of the fusion of the oocyte and the spermatozoon. In just a few days, the embryo develops from a single cell into a complex, three-dimensional, asymmetric, multicellular structure whose development and regulation mechanisms are still largely unknown.

This lack of knowledge is particularly lacking in assited reproductive technologies (ART) for infertile couples, where pregnancy rates are less than 30%.

In order to better understand how the human embryo develops, we conducted a multidisciplinary research project combining bioinformatics, hospital expertise, and fundamental research in developmental biology.

We have reconstructed the transcriptomic continuum of the human embryo cells during its first days of development.
We show that the emergence of the inner cell mass (ICM, future fetus) and the trophectoderm (TE, future placenta) takes place shortly after the formation of the blastocyst, and not at the morula stage as in the mouse.
We also show that the trophectoderm matures faster on the polar side, where it implants in the uterus.
Finally, we present new markers of the cell lineages of the embryo before and after its implantation in vitro.
This collaborative work opens new research perspectives in the study of human development, ART, and regenerative medicine.