Stem cell-derived embryo models in mouse and human to illuminate the “black box” of pre- to post-implantation development

Thesis by

Victoria Lynn Jorgensen

In Partial Fulfillment of the Requirements for

the Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2023

(Defended April 14th, 2023)

Victoria Jorgensen
ORCID: 0000-0002-4205-6198

ACKNOWLEDGEMENTS

When I began this PhD back in the fall of 2019, I always knew it would be hard. I didn't think it would be global-pandemic-takes-over-the-world hard, but it was. I didn't expect to lose my dad in the middle of it all, but I did. So 2019-Victoria, you were right, it was very hard, and harder than you could have predicted. But what I also knew is that I wouldn't have to endure this PhD alone. Both before and during grad school I have amassed the most amazing support system anyone could have asked for. To all of you, and you know who you are: From the bottom of my heart, thank you. I would not have survived this without you.

I want to thank my supervisor, Dr. Magdalena Zernicka-Goetz, for giving me the experience of a lifetime. In your lab I have gotten to work on two of the most amazing projects. Projects at the forefront of our field and that have pushed the boundaries of what I thought was even possible. I have had the privilege of meeting and collaborating with some of the most intelligent scientists I have ever met. I am grateful the opportunities I was given throughout grad school, and know that I am leaving here prepared and excited for my next steps.

Next I want to thank my thesis committee: Bruce Hay, Mitch Guttman, Matt Thomson, and Joe Parker. Trying to forge connections with faculty in a COVID-19 era was a bit daunting at first, but then you four made it easy. After each meeting with you four, I always felt reinvigorated for my project, and excited for the next steps. I am so appreciative for the conversations, discussions, feedback, and support that I have gotten from each of you.

Berna Sozen and Andy Cox, you were my first mentors in the lab and taught me nearly everything I know about mouse and human cell culture and blastoids. While our overlap in the MZG lab was short, it was incredible to have had the chance to work with you both.

Min Bao and Jamie Stover, thank you for being the greatest team, the best support crew, and the most amazing friends. Working alongside you both in the lab has definitely been a highlight of this PhD, and something I will always cherish. Min, I am so grateful to have worked alongside you, but I am even more grateful to have made a friend through our collaboration. Jamie, you've always been my biggest cheerleader. Your support during the

highs and lows were crucial, and I will always be grateful to you for cheering me on when I'm up and for picking me up when I'm down. Thank you both for helping me form the greatest trifecta that ever was.

Bailey Weatherbee, you are a superstar. Thank you for countless conversations, amazing scientific feedback, and for always being someone I could rely on even though you were thousands of miles away. Thank you for staying on the phone with me till 4 A.M. while I finished grant drafts, for helping me work through seemingly impossible situations, and for always being someone I could go to for help. I said it once, but I'll say it again, you are a superstar.

Daisy Wu, Sam Woo, and Reina Gabai, you are my people. Thank you for giving me a space to live and breathe away from my little science bubble. Having you three as lifelines during this PhD has saved me more times than I can count, and has helped more than you probably realize.

Matteo Guareschi, you've helped me through it all. I don't think words can really capture what your support has meant to me, but it's safe to say I would not have gotten through these last four years without you. It's been a crazy ride, thanks for strapping in next to me and helping me through all the ups and downs and twists and turns. Grazie per tutti, Matty. Sei il mio preferito e ti amo.

Dad, you're the reason I fell in love with science. From our backyard garden sessions, to neighborhood walks, to nature trips, you always pushed me to explore the world around me. You taught me how to be curious, to ask questions and to seek answers. You taught me how to not be so serious, and have fun—lessons that were very necessary these last four years. I wish you could be there for my graduation, but I know you'd be beyond proud to have another Dr. Jorgensen in the family. Thank you for always being my #1 fan.

Mom, you're the strongest person I know. You are the greatest role model, the best mentor, the proudest mom, and I am so grateful for all you've done for me. You taught me how to not sweat the small stuff, how to persevere through the big stuff, and how to hold my head

up high in good times and in bad. Your endless support has meant the world to me, and I hope you know I got through this because of you.

ABSTRACT

Mammalian development is a complex and highly regulated process by which a single cell, the totipotent zygote, gives rise to all lineages of the future organism. While incredible advancements have been made to study and understand the earliest events of our life, many questions are still unanswered. Moreover, the most precarious stage of development, implantation, remains a "black box" to researchers due to inaccessibility of the embryo within the uterus of the mother. In the last decade, however, the emergence of stem cell derived embryos represents an exciting alternative avenue to study these dynamic stages.

During my PhD, I worked to establish two pre-implantation stem cell models, one in human and one in mouse, to better understand the earliest days of mammalian development. These models replicate the blastocyst stage of development; at this point in time the embryo is ready to implant into the uterus and contains all embryonic and extra-embryonic tissues needed to form the future organism: the epiblast, the hypoblast, and the trophectoderm. Beginning with my human model, I demonstrate the ability of a single cell type, expanded potential stem cells (EPSCs), to give rise to structures that replicate the natural blastocyst in size, morphology, and initiation of lineage segregation. Furthermore, these human blastocyst-like structures can undergo the very beginning of post-implantation remodeling by forming an epiblast rosette and initiating lumenogenesis. Nevertheless, single cell RNA-seq (scRNA-seq) analysis reveals that lineages are not fully committed in this model, perhaps explaining why development is limited in these structures up to about Day 7/8. In the context of my mouse model, I combine not one but three distinct cell types to generate blastocyst-like structures: 1) wildtype embryonic stem cells (ESCs) to form the epiblast, 2) trophoblast stem cells (TSCs) to form the trophectoderm, and 3) Gata4-inducible ESCs to form the primitive endoderm. Again, these structures mimic the natural mouse blastocyst in morphology and lineage segregation and demonstrate the ability to transition to post-implantation stages. Development of the three blastocyst lineages was further confirmed via global scRNA-seq analysis comparing our Gata4i-Blastoids to natural embryos; importantly, however, this analysis also showed that differentiation of the mural trophectoderm, the tissue responsible

for uterine invasion, is lacking in our stem cell model and likely explains the inability for these blastoids to implant *in vivo*.

Altogether, this dissertation explains key aspects of pre- to post-implantation development and highlights the incredible power of stem cell-derived embryos to self-organize into structures that closely mimic the natural embryo.

PUBLISHED CONTENT AND CONTRIBUTIONS

Sozen, B.[#], **Jorgensen, V.[#],** Weatherbee, B. A. T., Chen, S., Zhu, M., & Zernicka-Goetz, M. (2021). Reconstructing aspects of human embryogenesis with pluripotent stem cells. *Nature Communications*, *12*(1), 5550. https://doi.org/10.1038/s41467-021-25853-4 **[#]co-first authors**

V.J. and B.S. designed and performed experiments and analyzed the data together with M.Z. V.J., S.C., and B.W. performed scRNA-seq data computational analysis, and statistics. V.J., B.S., and M.Z wrote the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS AND TABLES

Number *Page*

Chapter 1

EARLY MAMMALIAN EMBRYOGENESIS

1.1. Introduction

The initial days of mammalian development are critical to the overall health of the fetus, and are broken into three major stages: pre-implantation, peri-implantation, and post-implantation. Pre-implantation development encompasses zygote to blastocyst formation; this stage is characterized by fertilization and formation of the zygote, maternal to zygotic genome activation, cleavage stage divisions, compaction, lineage segregation, and finally generation of the blastocyst (Fig. 1.1). The blastocyst is composed of three distinct lineages: the embryonic epiblast, which will give rise to the embryo proper; the extra-embryonic primitive endoderm (mouse), or hypoblast (human), which generates the yolk sac; and the extra-embryonic trophectoderm, which will produce the placenta. At the peri-implantation stage, the blastocyst has hatched from the zona pellucida, the trophectoderm lineage further differentiates into polar and mural subtypes, and the embryo prepares for implantation into the maternal uterine tissue. Finally, during post-implantation stages, the embryo undergoes major morphological and transcriptional changes, ultimately leading to gastrulation and the formation of the three germ layers: ectoderm, endoderm, and ectoderm (Fig. 1.1).

Roughly 60% of human pregnancies fail during or prior to implantation, yet why this stage of development is so precarious remains largely unknown[1,2]. Understanding these early stages of development has immediate clinical significance, however, studying the first days of embryogenesis remains difficult due to inaccessibility of the embryo within the uterus of the mother. In recent years, the development of *ex utero* culture conditions for mouse[3–5] and human[6–8] open an exciting avenue to explore the mechanisms governing these stages, and thus better understand why a significant number of pregnancies fail during this period. Additionally, the rapidly developing field of stem cell-derived embryo models

is another powerful tool that could enable high-throughput and systematic studies to elucidate the dynamics of early development; this topic will be discussed further in subsequent chapters.

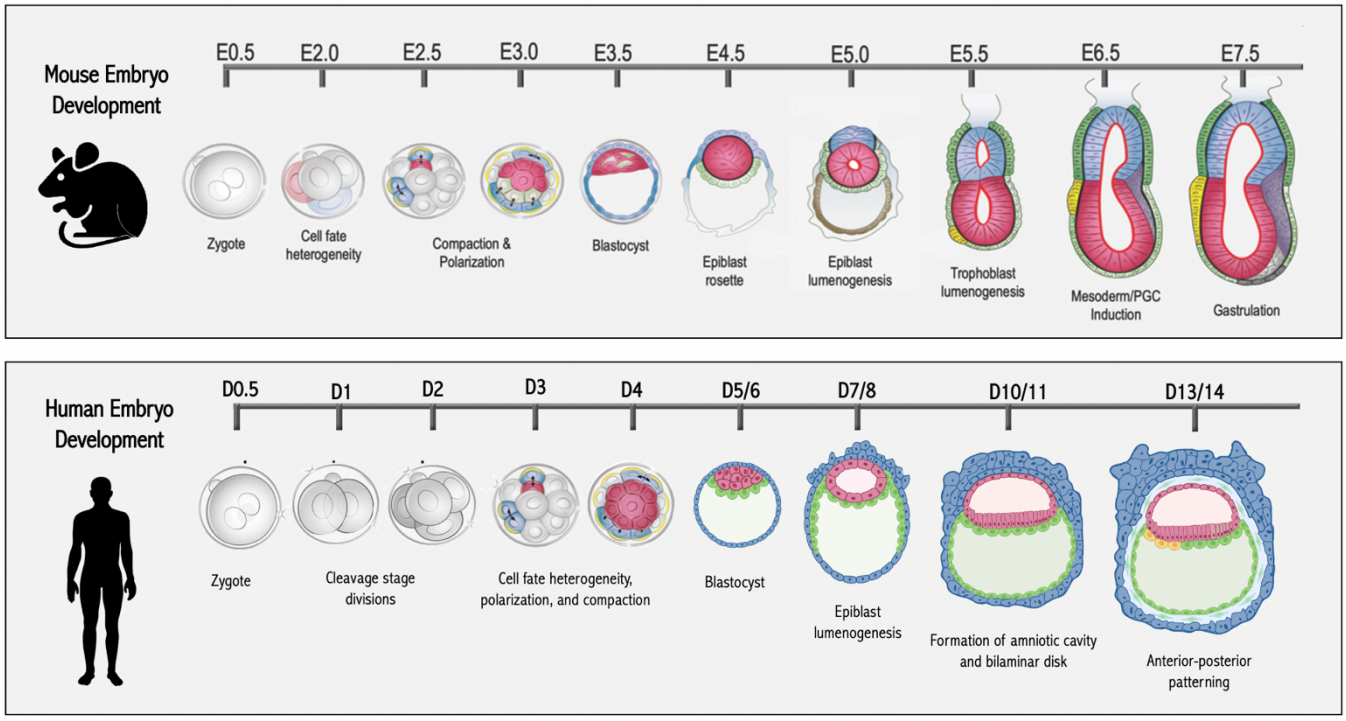


Figure 1.1: Schematic overview of mouse and human embryogenesis.
While mouse (top) and human (bottom) development share many commonalities, there also exists several major divergences. Embryogenesis commences upon fertilization of the female oocyte by the male sperm, forming the totipotent zygote. The zygote will then undergo a series of rapid cleavage stage divisions which will increase the number of cells within the embryo without altering the total volume. While cell fate heterogeneity is apparent in the mouse embryo by as early as the 4-cell stage, initial asymmetries in human embryos is believed to occur later. After cleavage divisions, the cells will undergo polarization and compaction; this will initiate inner and outer cell populations which designates the first lineage segregation event of the embryo, with outer cells giving rise to the trophectoderm (TE, blue) and inner cells giving rise to the inner cell mass (ICM, red). At embryonic day (E) 4.0 in mouse and day (D) 5/6 in human, the blastocyst has formed, and the second lineage segregation event has occurred with the ICM bifurcating into the epiblast (red) and primitive endoderm (green). Upon implantation, major morphological remodeling will take place forming the quintessential egg-cylinder in mouse, or bilaminar disk in human. In addition to morphological transformations, transcriptional changes leading to axis patterning, primitive streak formation, and ultimately gastrulation will take place. Mouse schematic adapted from illustrations by Andy Cox.

1.2. Pre-implantation development: Zygote to blastocyst

1.2.1. Fertilization, cleavage stage divisions, and ZGA

Many events of pre-implantation development, the period from zygote to blastocyst formation, are conserved between mouse and human. This stage commences upon fusion of the two gametes, the maternal oocyte and paternal sperm, leading to the formation of the zygote. This initial single cell is totipotent, meaning it can give rise to all embryonic and extra-embryonic tissues necessary to develop the fetus, placenta, and yolk sac. This totipotency is established due to the fact that the zygote is transcriptionally inactive, relying solely on maternally deposited proteins and mRNA that then remodel and activate the zygotic genome[3–5]. In mouse and human, the process of maternal to zygotic genome activation (ZGA) occurs in two waves, the minor and major ZGA, and these process happen with slightly different timescales in each organism[10]. For mouse, minor ZGA begins at the late zygote stage, and major ZGA at the two-cell stage[11]. In humans, both waves are delayed in comparison to the mouse with the minor and major waves starting at the four-cell stage and eight-cell stage, respectively[12]. While the early cells of the embryo, termed blastomeres, seem identical, evidence in the mouse shows that heterogeneities arise during the two-cell to four-cell stages[13–15] in relation to onset of ZGA[16], chromatin accessibility and remodeling[17,18], and partitioning and localization of cellular components[19]. While extensive studies on this heterogeneity has been carried out in the mouse, it remains unclear if this is also true for the context of human blastomeres. While some analysis has been done on this at the transcriptomic level[20], further studies will be necessary to analyze blastomere differences at the proteomic and epigenetic levels.

1.2.2. Compaction and polarization

Shortly after fertilization, and concomitant with ZGA, the single-celled zygote undergoes cleavage stage divisions; this is a series of rapid cell divisions that increase the number of cells without changing the total volume of the embryo. The cells of the embryo then undergo polarization and compaction, and this again occurs in a species-specific manner. In the mouse embryo, cell polarization and compaction occurs at the 8-cell stage[21,22] whereas in humans

this takes place during the 8- to 16-cell transition[23,24]. The establishment of inside-outside polarity is of critical importance to the embryo, as this designates the first cell fate decision. Outer cells will acquire an apical-basal cell polarity, accumulating atypical protein kinase C (aPKC) along the contact-free domain. These cells will also obtain nuclear expression of various markers such as components of the YAP/Hippo signaling pathway and the transcription factor GATA3 in humans[23,24] or Cdx2 in mice[25] (Fig. 1.3). Inner cells on the other hand remain unpolarized and will maintain expression of key pluripotency markers such as *Sox2*, *Nanog,* and *Oct4*[26,27]. Ultimately these polarized cells will give rise to the trophectoderm lineage whereas the inner unpolarized cells will give rise to the inner cell mass (ICM). It is important to note that key divergences exist between mouse and human in regard to this first lineage segregation event. One such example is that while Cdx2 is a key marker for designating the TE lineage in mouse, this factor does not become upregulated in human trophectoderm until the late blastocyst stage[28]. Thus, understanding how this first cell fate decision takes place in human remains elusive.

1.2.3. The blastocyst

The culmination of pre-implantation development leads to the formation of a structure known as the blastocyst. The blastocyst is first comprised of two major cell types: 1) the outer extra-embryonic trophectoderm (TE) which will generate the placenta, and 2) the bipotent inner cell mass (ICM) (Fig. 1.2). Eventually the ICM will diverge into the embryonic epiblast (EPI), which generates the embryo-proper, and the extra-embryonic primitive endoderm (PE) (also referred to as hypoblast (HYPO) in humans). The PE lineage will eventually give rise to the parietal endoderm (PaE) and the visceral endoderm (VE), and ultimately the yolk sac (Fig. 1.2). While the trophectoderm lineage is established early on during polarization, the bifurcation of the ICM occurs slightly later at around embryonic day (E) 3.0 - 3.5 in mouse[29] and day (D) 5/6 in human[30]. At approximately E4.75 in mouse, or D7/8 in human, the trophectoderm will undergo its own bifurcation event forming polar and mural TE subpopulations (Fig. 1.2).

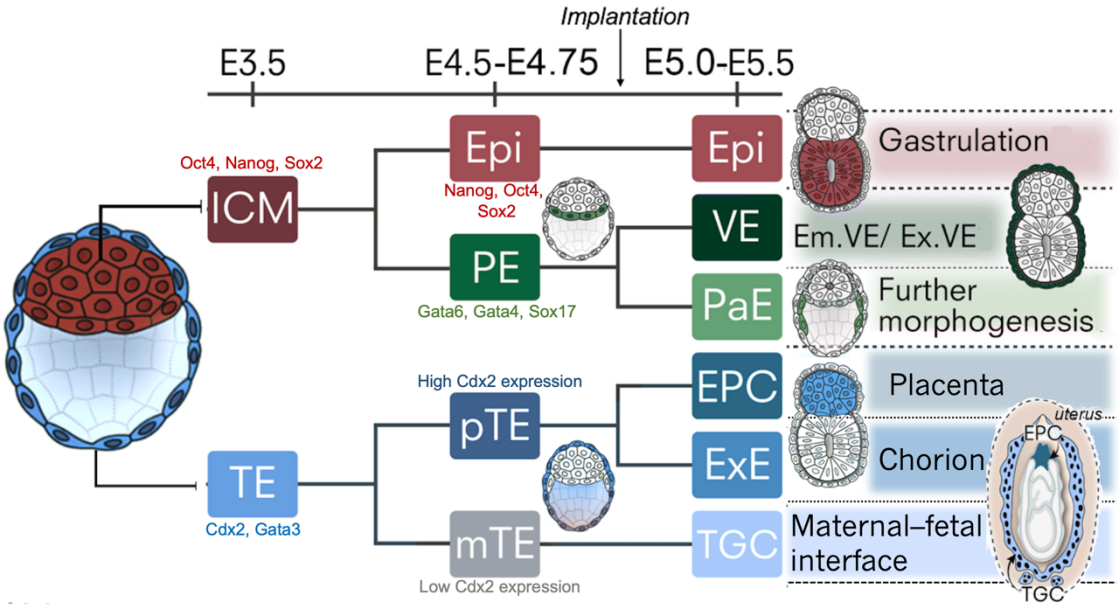


Figure 1.2: Differentiation of blastocyst lineages in mouse.
By E3.0 of embryogenesis, the blastocyst is initially composed of two lineages: the trophectoderm (TE, blue) and the bipotent inner cell mass (ICM, red). By E4.5, the ICM has diverged into two new lineages: the epiblast (EPI, red) which will give rise to the embryo-proper and undergo gastrulation, and the primitive endoderm (PE, green), which will give rise to the visceral endoderm (VE) and parietal endoderm (PaE). At this stage, the TE has also generated two unique subtypes: the polar TE (pTE), which will ultimately give rise to the ectoplacental cone (EPC) and extra-embryonic ectoderm (ExE) responsible for forming the placenta and chorion, and the mural TE (mTE) which will form trophoblast giant cells (TGCs) responsible for invasion into the maternal uterine tissues. Key lineage markers are written below each stage. Schematic adapted from Sozen *et al.,* 2019.

In the mouse embryo, it has been shown that cells internalized in the first wave of asymmetric divisions (establishing the inside-outside populations) strongly bias cell towards an epiblast fate; subsequently, cells internalized during the second wave of cell divisions are biased towards the primitive endoderm fate[31]. The bifurcation of the ICM is controlled by mutual exclusive expression of two master transcription factor regulators, Nanog and Gata6; the first leading to the formation of the epiblast and the latter to primitive endoderm[29] (Fig. 1.3). While these two factors are initially co-expressed starting at the 8-cell stage, they become gradually restricted in their expression in a "salt and pepper" like fashion. By E4.0, the ICM has fully diverged and the factors are restricted to their appropriate lineage. It was also shown that these early epiblast-fated cells demonstrate higher expression of *Fgf4* whereas the primitive-endoderm fated cells show higher expression of *Fgfr1* and *Fgfr2*[29]. FGFR1/2 positive cells exposed to increased levels of FGF4 ligand upregulated FGF/MAP signaling,

leading to subsequent expression of primitive endoderm specific markers *Gata6* and *Sox17*. Alternatively, cells producing FGF4 ligands, yet lacking FGFR1/2, will show higher expression of pluripotency markers *Nanog* and *Oct4*, committing them to the epiblast fate[29,31] (Fig. 1.3).

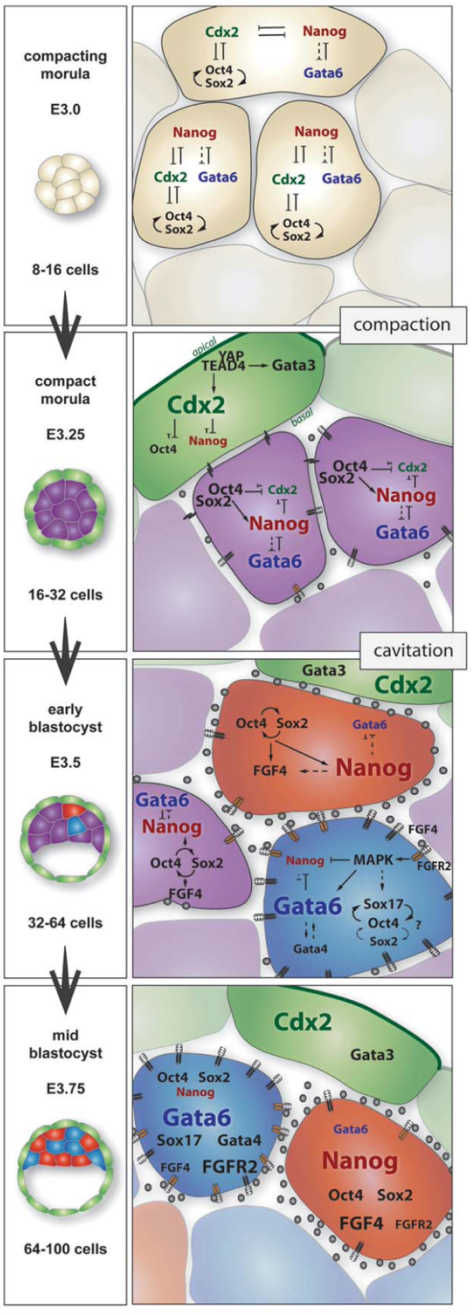


Figure 1.3: Signaling events responsible for early lineage fate determination in mouse. Key transcriptional programs have been identified as master regulators governing the earliest cell fate decisions. First, in the compact morula, outer cells (green) will begin to upregulate the transcription factor Cdx2, whereas inner cells (purple) will upregulate the transcription factors Nanog and Gata6. Outer cells will also upregulate YAP/TEAD4 signaling and initiate Gata3 expression. At roughly E3.5, the stage of early blastocyst, inner cells will diverge into either epiblast-fated cells (red), which show higher levels of Nanog, or primitive endoderm-fated cells (blue), which show higher levels of Gata6. The Nanog-producing epiblast cells will go on to express additional pluripotency markers such as Oct4 and Sox2, which in turn leads to the production of Fgf4. Conversely, the Gata6-producing primitive endoderm cells upregulate key endoderm fate determinants such as Sox17 and Gata4. Additionally, the primitive endoderm-like cells will produce Fgf receptors like Ffgr2 and Fgfr1 (not shown), which will respond to Fgf4 ligands from neighboring epiblast cells and trigger downstream MAPK pathways. By E3.75, the epiblast and primitive endoderm fates have been committed, and will sort accordingly in subsequent events.
Adapted from Schrode *et al.*, 2013.

In humans, the process of lineage segregation in the blastocyst shares many similarities, however there are key divergences. Firstly, human blastocyst formation occurs slightly later at ~D5/6 as opposed to E3.0-E3.5 in mouse[28]. Furthermore, unlike the mouse, the pluripotency marker OCT4 (also known as POU5F1) remains expressed in both the trophectoderm and ICM cells until ~D6 of embryo development; at this point, the trophectoderm lineage will begin to downregulate OCT4, whereas the hypoblast will retain expression up until implantation at ~D7[8,28,32]. Cells of the early ICM show co-expression of NANOG and GATA6. By late-blastocyst formation at D6, epiblast cells are designated by NANOG and high OCT4 expression whereas cells of the hypoblast are characterized by SOX17, GATA6, GATA4 and decreased expression of OCT4[8,33]. Interestingly, while FGF/ERK signaling has been shown to be crucial for primitive endoderm specification in mice, inhibition of this pathway in humans does not disrupt hypoblast formation[33]; this indicates that an alternative mechanism is likely responsible for this cell fate decision in humans and remains and exciting area of further investigation[34]. Further analysis will be necessary to fully understand the segregation of lineages in human embryos in the coming years.

1.3. Peri- to post-implantation development: Maturation of lineages

Peri-implantation refers to the stage of development during which the embryo begins to invade and interact with the uterine tissues. By this stage, the embryo has also fully segregated into the three distinct lineages necessary to produce the future organism: the trophectoderm, the epiblast, and the primitive endoderm/hypoblast. Each of these lineages will undergo extensive morphological changes that are coordinated by highly regulated transcriptional programs. Physical cues and mechanical forces have also been implicated as major factors contributing to remodeling at this stage. It is during this time that we also begin to see marked divergences between the mouse and human systems in a variety of aspects including implantation mechanisms, overall shape, and timing of events.

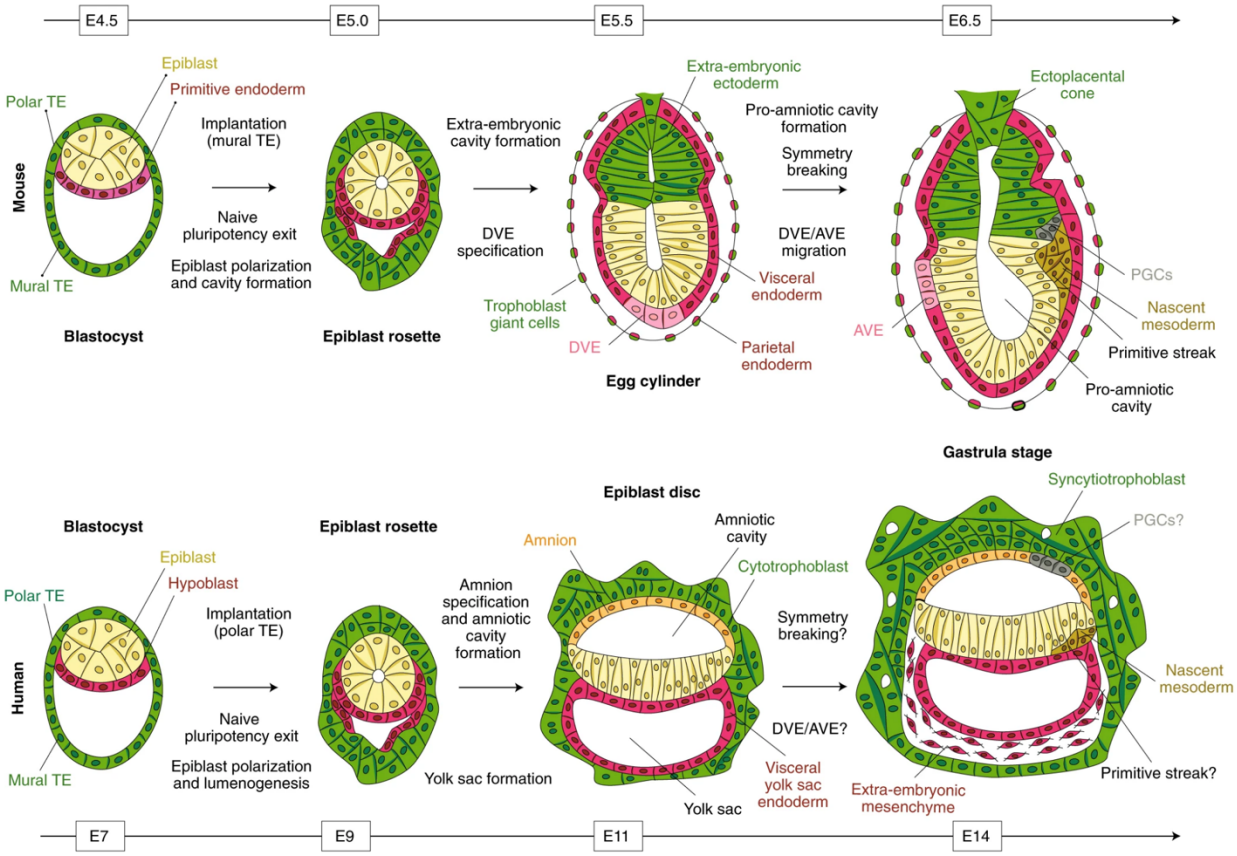


Figure 1.4: Schematic overview of post-implantation development in mouse and human.
Upon implantation, the blastocyst structure will undergo massive remodeling events and further lineage differentiation. Notably, the mouse embryo (top) will develop into the cup-shaped egg cylinder, which is composed of three key lineages: the epiblast (yellow), the extra-embryonic ectoderm (green), and the visceral endoderm (red). In human, the blastocyst will form what is referred to as the bilaminar disk, which is composed of the following: the epiblast and amnion (yellow), the visceral yolk-sac endoderm (red) and the outer cytotrophoblast (green). Upon this restructuring, events leading to symmetry breaking, axis patterning, and ultimately primitive streak formation will follow. Shahbazi *et al.,* 2018.

1.3.1. Trophectoderm

Upon maturation of the blastocyst, the embryo is ready to implant and initiates adhesion and invasion into the maternal tissue via the trophectoderm. At this point, the trophectoderm has specified into two major sub-types: the polar trophectoderm, which remains in close contact with the epiblast (embryonic pole), and the mural trophectoderm which is separated from the epiblast across the blastocoel cavity (abembryonic pole) (Fig. 1.2). Proximity to the epiblast and associated secreted factors are believed to be what drives segregation of these two cell fates[35,36]. Secretion of FGF4, as mentioned earlier, is critical

for segregation of the primitive endoderm lineage; expression of FGFR1/2 and reception of FGF ligands is also important for trophectoderm specification and is believed to drive the tissue boundary between polar and mural fates as well. Nevertheless, precisely how these two TE fates arise is not well understood and remains an active area of further investigation.

In the mouse, implantation occurs between E4.5 and E4.75 and is initiated by the mural trophectoderm. Just prior to implantation, the mural TE will begin to downregulate a key transcription factor CDX2, signifying onset of differentiation and a loss of multipotency[37]. These mural cells will undergo epithelial-to-mesenchymal transition and become trophoblast giant cells (TGCs) that are now capable of undergoing invasion into the endometrium[38,39]. The polar TE cells, however, will retain high expression of CDX2, and thus retain proliferative and self-renewing capabilities. These cells will later go on to develop the extra-embryonic ectoderm (ExE), ectoplacental cone (EPC), and the chorion (Fig. 1.4). For implantation and invasion, embryos become positioned in crypts located throughout the uterine lumen; the lumen will then close, encapsulating the embryo entirely. Once the lumen has sealed, TGCs will completely surround the conceptus; these cells are large and invasive polyploid cells that facilitate interactions between the embryo and the uterine wall[40].

In the human, there are key divergences compared to mouse with regards to implantation. First is timing, as human embryo implantation occurs slightly later, around D7/8[41,42]. Additionally, the process of invasion and implantation occurs via the polar trophectoderm, as opposed to the mural trophectoderm in mouse. Furthermore, the uterine wall in humans does not contain crypts, and the site of implantation is unknown and is not predetermined. Upon contact of the embryo with the luminal epithelial cells of the uterine wall, a process known as apposition, the polar trophectoderm will differentiate into polyploid syncytiotrophoblast (SCT) and mononuclear extra-villous cytotrophoblast (CTB)[40] (Fig. 1.4). The cytotrophoblast will form immediately abutting the cells of the epiblast, whereas the syncytiotrophoblast will form surrounding these cells and invading deeper into the

uterine wall. The invasive SCTs will produce various matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) in order to degrade the extra-cellular matrix of the endometrial epithelium and invade further into the tissue. Eventually the embryo will completely invade into the uterine stroma in order to make contact with the maternal blood supply[43].

Both mice and humans undergo placentation to form the placenta, the organ that is responsible for nutrient exchange between the mother and fetus. Nevertheless, key divergences exist between the two species with regard to formation and function of the placenta. In mouse embryos, implantation typically occurs between days 4 and 6 after fertilization, and placentation is complete by day 9[44]. In contrast, human embryos typically undergo implantation between days 7 and 10 after fertilization, and placentation is complete by day 17[44].

In mouse, once the embryo has been fully implanted, the polar trophectoderm will continue to proliferate due to signals of fibroblast growth factor 4 (FGF4) emanating from the epiblast. These cells will form a tightly packed epithelial column known as the extra-embryonic ectoderm (ExE), and retains some stem cell potency[44] (Fig. 1.4). Cells located furthest from the epiblast at the proximal tip of the ExE compartment will go on to form the ectoplacental cone (EPC) and secondary TGCs. The secondary TGCs are again invasive in nature, and will deeply penetrate the endometrial stroma to come in contact with maternal blood vessels. Lastly, upon gastrulation, the ExE compartment will differentiate to form the chorion, which will contribute to the fetal component of the placenta[44] (Fig. 1.4).

Unlike mouse, the timeline of events for placentation in humans is much less understood due to ethical and technical restrictions. It is known, however, that the invasive SCTs will eventually line the blastocyst to form a network of connective tissue, with interspersed blood vessels, and this constitutes the chorion. The chorion will then facilitate the exchange of nutrients and other materials between the mother and the fetus[42].

1.3.2. Epiblast

Upon implantation in mouse, the embryonic epiblast undergoes complex and highly coordinated reorganization, both morphologically and transcriptionally. Another important change is the transition of the epiblast from a naïve to primed pluripotency state[45]. During the naïve state, the epiblast maintains high self-renewing capacity and the ability to contribute to both soma and germ cells when reintroduced to chimeric blastocysts[46]. Additionally, the naïve epiblast maintains a key epigenetic signature: X chromosome activation. In XX embryos, the paternal X chromosome is silenced around the 4-cell stage, and remains silenced in all extra-embryonic tissues (trophectoderm and primitive endoderm). Interestingly, however, both X chromosomes will become re-activated leading up to implantation, but only in the embryonic epiblast lineage[47]. Once the embryo has implanted, the epiblast will convert to a single-layer columnar epithelium; the embryo know resembles a cup-shaped structure known as the egg-cylinder. At this point in time, cells are in what is known as the "primed state" and no longer maintain the ability to contribute to pre-implantation blastocyst chimeras[48]. These cells will also undergo random X-chromosome inactivation in female embryos[47]. In humans, the role of X-chromosome inactivation (XCI) is not as clearly understood. It is known, however, that unlike mouse XCI is not restricted in extra-embryonic lineages early on as the placenta, a trophectoderm derivative, shows extensive mosaicisms with regards to XCI[49]. It is predominantly believed that XCI does not occur until after implantation, or perhaps late blastocyst stage, but further investigations on this topic are necessary to fully understand XCI during early human development[50–52].

Another key morphological change characteristic of the post-implantation epiblast is epithelization, polarization, and lumenogenesis leading to the formation of the pro-amniotic cavity (Fig. 1.4). During this remodeling, the epiblast will form a rosette-like structure that is heavily dependent on basal membrane-stimulated integrin signaling[53]. This will lead to apical restriction and polarization of the epiblast cells, which is necessary for lumenogenesis. From this apical region, Rab11-dependent secretion of glycosylated

sialomucin proteins, most importantly Podocalyxin, lead to the emergence of the pro-amniotic cavity by creating membrane repulsion and separation[54].

In humans, similar processes are required for the formation of the amniotic cavity. Again, the epiblast will polarize and organize into a rosette structure, with secretions of Podocalyxin at the center to lead to lumen formation. Importantly, in order for lumenogenesis to take place in both mouse and human, epiblast cells must exit the naïve state and transition to a primed state of pluripotency[54]. Once the cavity has been formed within the epiblast, key morphological events will take place. In mouse, the epiblast will eventually give rise to a cup-shaped cylindrical structure and ultimately fuse with the cavity of the ExE compartment. In humans, however, this does not occur, and instead the epiblast will take on a disk-like structure (Fig. 1.4). This cavity will begin to expand and lead to separation of the epiblast epithelium and the overlaying amnion layer[55]. While the epiblast disc will eventually undergo axis patterning and gastrulation, the amnion will lead to generation of the amniotic sac; this sac surrounds the entirety of the conceptus and is critical for protecting the fetus from physical stress, temperature fluctuations, dehydration, and more. While little is known about generation of the amnion tissues, it is believed that segregation from the epiblast epithelium is mediated by downregulation of key pluripotency markers (OCT4, NANOG, SOX2), and subsequent upregulation of putative amnion markers (TFAP2A, ISL1, HAND1)[56]. BMP and WNT signaling has all been heavily implicated in amnion formation, expressing factors such as BMP4 and WNT3A[57].

1.3.3. Primitive endoderm/hypoblast

Similar to the epiblast and trophectoderm lineages, the primitive endoderm will also undergo key transcriptional and morphological changes throughout peri- to post-implantation development, however major questions remain in elucidating this process. In mouse, the primitive endoderm will also sub-divide into two distinct lineages. The first is the parietal endoderm (PaE), which alongside the mural trophectoderm, will contribute to Reichert's membrane; this is a thick basal lamina that facilitates deposition of extra-cellular matrix to facilitate post-implantation morphogenesis, as well as to protect the embryo from certain

stressors[58]. The cells at the line the epiblast, however, will eventually migrate to encompass both the epiblast and ExE compartments to give rise to the visceral endoderm (VE) lineage, completing the egg-cylinder structure characteristic of post-implantation development (Fig. 1.4). The visceral endoderm will eventually lead to formation of the visceral yolk sac as well as definitive endoderm which will form the inner lining of the gut[59,60]. In humans, very little is known about further development of the hypoblast lineage and remains an active area of research.

1.3.4. Proximal-distal axis patterning

At E5.0 of mouse development, the embryo consists of three predominant lineages: the epiblast, the ExE compartment, which directly abuts the epiblast at the proximal end, and the VE which encapsulates the two previous lineages. Additionally, it is believed that at this point in time, the epiblast remains unpatterned, expressing key pluripotency markers such as *Sox2*, *Nanog, Nodal, Sall4*, and *Oct4*[61–64]. In order to establish proximal-distal axis patterning, the ExE compartment secretes BMP4 ligands, which subsequently leads to expression of *Wnt3* and *Nodal* within the epiblast; Nodal signaling is further propagated via expression in the VE[65]. At ~E5.5, in response to Nodal signaling, cells at the distal tip of the embryo in the visceral endoderm undergo a transition from cuboidal to columnar morphology; this signifies generation of the distal visceral endoderm (DVE)[66]. In addition to the morphological changes that segregate the DVE from neighboring VE cells, key transcriptional programs are upregulated as well. Nodal will go on to activate SMAD2, and this Nodal-SMAD2 complex will initiate expression of key transcription factors such as *Foxa2* and *Lhx1*[65]. These factors, in unison with SMAD2, will lead to expression of Wnt and nodal antagonists such as Dkk1, Cer1, and Lefty1 within the DVE[65]. Initial axis patterning in the human context remains largely unexplored, however in the following section I outline current knowledge regarding anterior-posterior axis formation in humans.

1.3.5. Anterior-posterior axis patterning

Between E5.5 and E5.75, the DVE will then migrate unidirectionally to the proximal side of the embryo, directly adjacent to the EPI-ExE boundary; this will form the anterior visceral

endoderm (AVE), the signaling center responsible for inducing anterior patterning on the underlying epiblast[67,68]. The exact mechanisms leading to AVE migration are not fully understood, and in particular it is unclear at which point the anterior-posterior axis is first designated. Evidence shows that designation of AVE may occur as early as the blastocyst stage in mouse, with asymmetric expression of *Lefty1* in the primitive endoderm lineage[69]. Lineage tracing experiments using Hhex-GFP-, Lefty1-lacZ- and Cer1-GFP-expressing have also shown that these genes are present successively throughout development in the ICM, primitive endoderm, DVE and AVE, yet it remains unclear if they are clonally related[70,71]. In addition to *Lefty1,* the AVE marker *Cer1* is also seen upregulated in ICM cells prior to implantation. Nevertheless, when tracking Cer1-postivie ICM cells throughout development, it has been shown that individual cells contribute to only a fraction of the Cer1-GFP-expressing AVE. This would therefore suggest that the AVE is comprised of pre-implantation pre-cursors as well as cells with *de novo* gene expression upon implantation[70].

Culture conditions of human embryos *ex utero* remains challenging, and thus limited information is available regarding axis patterning. Importantly, while there have been some breakthroughs in these techniques, it is known that not all lineages are robustly supported; thus, conclusions regarding *in vitro* culture of human embryos may not accurately recapitulate *in vivo* events. Nevertheless, a study published in 2021 performs single cell analyses and functional assays in order to elucidate key events of human post-implantation development[72]. They first show that FGF signaling emanating from the epiblast lineage is critical for proliferation and maintenance of the extra-embryonic lineages. Additionally they identify an asymmetric population of cells in the hypoblast lineage which expresses inhibitors of the WNT, NODAL, and BMP pathways, such as CER1 and LEFTY1; they hypothesize that this is the anterior signaling center which pre-patterns the anterior-posterior axis within the adjacent epiblast. Interestingly, unlike the DVE in mouse which arises at the distal pole of the embryo in a distinct subset of cells, this signaling center occupies a widespread domain that becomes isolated to one side of the embryo over time. Thus, further studies investigating the mechanism for this localization and patterning will be required in the future.

1.3.6. Primitive streak formation

At ~E6.5, once anterior-posterior axis patterning has been established, the embryo is prepared to undergo gastrulation. Gastrulation is initiated by the formation of the primitive streak (PS), which occurs within the epiblast directly across from the AVE. The first morphological event signifying primitive steak formation is a thickening of the epiblast compartment, a process known as epithelial to mesenchymal transition (EMT). During this process, cells will detach from the pseudo-stratified epithelial cells of the epiblast to ingress and invaginate. Upon completion of gastrulation, all three germ layers of the future fetus will have been specified: the endoderm, which will predominantly form the gut; the mesoderm, which will generate bones, connective tissue, and skeletal muscles; and lastly the ectoderm, which will generate the skin, hair, and nervous system.

According to the "14 day rule" natural human embryos cannot be cultured beyond Day 14, or at any instance of primitive streak formation. Given this limitation, little is known about this process in humans and will likely remain a challenging area for research for the foreseeable future. Excitingly, several stem cell-derived embryo models for post-implantation human development have been established and offer a potential outlet to study this process outside the context of the embryo; more on this is found in Chapter 2.

Concluding remarks

In conclusion, early mammalian development is a complex and tightly organized process. Errors at any stage of lineage maturation can lead to congenital diseases, failure to implant, or even lethality to the embryo. While extensive research has been done on the earliest stages of mouse and human development, critical questions remain unanswered. A major hurdle preventing research in these areas is the ethical and technical restrictions that limit access to embryos at these stages. In the following chapters I will discuss several stem cell-derived embryo models that have the ability to recapitulate key developmental stages at a morphological and transcriptional level. The establishment of these models has the potential to revolutionize early embryo research by allowing large-scale and high-throughput studies

necessary for unraveling many questions in developmental biology that have eluded scientists for decades.

Chapter 2

STEM CELL-DERIVED EMBRYO MODELS IN MOUSE AND HUMAN

2.1. Introduction: The importance of stem cell-derived embryo models

Studying early embryogenesis, particularly in humans, has numerous technical and ethical restrictions. Moreover, the culture of human embryos *ex utero* is limited by the "14-day rule", by which embryos cannot be cultured passed 14 days, the point at which the primitive streak forms. In the last decade, stem cell-derived embryo models have emerged as a powerful tool to help overcome these limitations and to further elucidate the events of early development. In this chapter, I will summarize the current stem cell models for mouse and human, how these models recapitulate key morphological and transcriptional events of early development, and lastly where these models are lacking.

2.2. Stem cell-derived models: Mouse

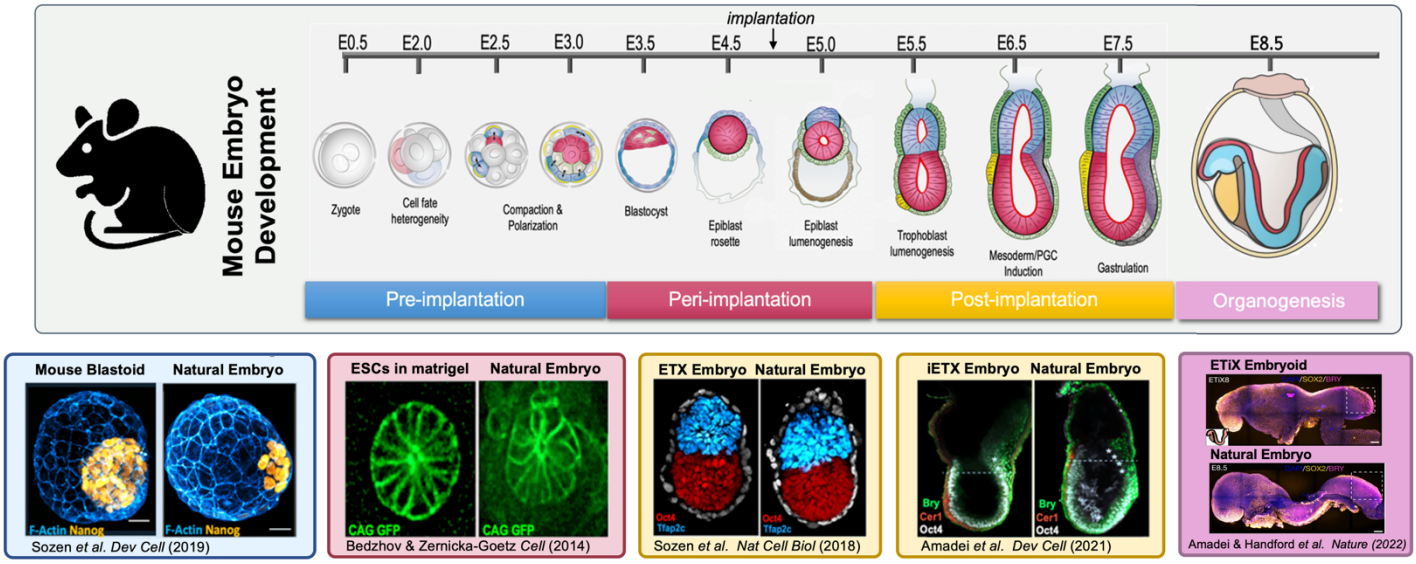


Figure 2.1: Overview of stem cell-derived models in mouse.
Top: Overview schematic of mouse embryo development, highlighting 4 major stages: pre-implantation, peri-implantation, post-implantation, and onset of organogenesis. **Bottom:** Examples of different stem cell-derived embryo systems in comparison to natural embryos. Each model corresponds to a distinct stage of development: Blastoids compared to the blastocyst (Sozen *et al.,* 2019); ESCs in matrigel compared to the epiblast rosette (Bedzhov & Zernicka-Goetz, 2014); ETX/iETX embryos compared to post-implantation embryos (Sozen *et al.*, 2018 & Amadei *et al.*, 2021); ETiX embryos compared to gastrulating embryos with signs of neurulation and organogenesis (Amadei & Handford *et al.*, 2022).

2.2.1. Stem cell derivation in mouse: ESCs, TSCs, and XEN cells

Mouse embryonic stem cells (mESCs) were first derived from the inner cell mass of blastocyst stage embryos in 1981[73,74]. Under distinct culture conditions, these cells exhibit properties of a naïve epiblast, demonstrating unlimited self-renewal, pluripotent potential, and the ability to contribute to epiblast lineages when re-incorporated into chimeric embryos[46]. A major breakthrough in the culture and maintenance of embryonic stem cells *in vitro* was the discovery of the "2iLif" condition, which utilizes three components to establish and stabilize the naïve pluripotent state[75,76]. These molecules are: 1) leukemia inhibitory factor (Lif) which inhibits differentiation of ESCs via activation of the transcription factor Stat3 (Smith, 2001); 2) CHIR99021, an inhibitor of glycogen synthase kinase 3 (GSK3) which ultimately activates Wnt pathways and promotes ESC viability and self-renewal; and 3) PD0325901, a MAP kinase/ERK kinase (MEK) inhibitor, that suppresses differentiation. Roughly 20 years later, culture conditions for isolating and maintaining trophoblast stem cells (TSCs) in culture were also established using a single factor, fibroblast growth factor 4 (Fgf4)[77]. In more recent years, culture conditions to mimic primitive endoderm-like cells, termed extraembryonic endoderm (XEN) cells, have also been defined via two methods: chemical induction by culturing ESCs with retinoic acid, or transcriptional manipulation by overexpressing GATA factors *Gata4* or *Gata6*[78,79] (Niakan, 2013; Fujikura, 2002).

2.2.2. Self-organization of the epiblast-rosette with ESCs

Given the ability to maintain mouse embryonic and extra-embryonic stem cells in 2D culture has led to the emergence of a rapidly developing field of stem cell-derived embryo models. One of the first models to arise shows that mouse ESCs grown in the presence of extra-cellular matrix (ECM) are able to recapitulate apical constriction and lumenogenesis characteristic of the peri-implantation epiblast rosette just prior to implantation[53].

2.2.3. ET Embryos

The next model to develop, termed ETS-embryos, combines ESCs, to mimic the epiblast (EPI), and TSCs, to mimic the trophectoderm (TE)/extra-embryonic ectoderm (ExE), in

matrigel; here the matrigel is used to substitute the primitive endoderm and subsequent lineages by providing an ECM scaffold[80]. Although this model is lacking one of the three major lineages, the system is still able to mimic distinct events of post-implantation development including: self-organization, polarization, and lumenogenesis of the embryonic epiblast and the extra-embryonic ectoderm (ExE) lineages, fusion of these two compartments to form a pro-amniotic-like cavity, asymmetric expression of mesoderm marker *T/Brachyury*, and specification of cells seemingly equivalent to primordial germ cells (PGCs) of the natural embryo[80]. Nevertheless, without contribution of the extra-embryonic primitive endoderm lineage, it calls into question how well this model can truly capture natural embryos and overall limits the developmental potential of these structures.

2.2.4. ETX Embryos

In order to overcome the limitations of ETS-embryos, ETX-embryos, are cultured without an exogenous ECM and combine three cell types to mimic post-implantation embryos: ESCs, TSCs, and now XEN cells, to mimic the PrE/VE lineage. In addition to the hallmarks observed in the ETS-embryo model, this system demonstrates encapsulation of the EPI and ExE compartments by a VE-like layer and formation of an anterior signaling center that establishes the anterior-posterior axis within the epiblast[81]. Lastly, while *T/Brachyury*-expressing cells show evidence of epithelial-to-mesenchymal (EMT) transition, ingression of cells into the mesodermal layer is incomplete and structures fail to fully initiate gastrulation. While the ETX-embryo can model many events of pre-gastrulation development, their potential is inherently limited; this is likely due to the fact that XEN cells do not sufficiently recapitulate the primitive endoderm lineage and its subsequent derivatives.

2.2.5. Inducible ETX Embryos

To overcome the limitations of previous systems, recent models, termed iETX or ETiX embryos, have replaced XEN cells with *Gata4*-inducible ESCs to drive the extra-embryonic endoderm fate. With this modification, iETX embryos are not only able to develop a robust anterior signaling center, but also undergo gastrulation to develop embryonic and extra-

embryonic mesoderm and definitive endoderm[82]. A major limitation of this system, however, is that structures were not able to be cultured beyond 6 days due to insufficient *in vitro* culture conditions. Recently, however, with the development of a roller culture system, natural and stem cell-derived embryos are able to be cultured until ~E8.5[5]. As such, ETiX embryos are able to progress two additional days, allowing for the formation of a brain, neural tube, heart, foregut, somite, allantois, primordial germ cells and yolk sac structures[83]. Two additional models, sEmbryos[84] and EiTiX-embryos[85] take this model one step further by utilizing not only *Gata4*-inducible ESCs, but also *Cdx2*-inducible ESCs to replace TSCs. These systems also produce structures that mimic ~E8.5 post-gastrulation embryos that have undergone advanced organogenesis. Unfortunately, all of these models are capped at their developmental potential, as culturing natural mouse embryos *ex utero* is limited to ~E8.5. Altogether, these models demonstrate the outstanding ability of stem cells to self-organize into complex ordered systems, but also highlight the necessity for improved culture conditions for embryos outside the uterine environment.

2.2.6. Mouse blastoids

In addition to the post-implantation models listed above, several models for pre-implantation development, known as "blastoids", have also been developed in recent years. The first model established combined ESCs and TSCs *in vitro* to mimic embryonic day 3.5[86]. Not only do these structures nicely mimic natural blastocysts in size, morphology, and cell number, the authors also performed immunostaining of key lineage markers to demonstrate proper initiation of all three cell types in the correct locations Nevertheless, efficiency to form blastoid structures, especially ones with all three lineages represented, occurs at low frequency; in particular, representation of primitive endoderm-like cells in this model seems lacking. This is likely due to the fact that mouse ESCs due not mimic the bipotent ICM, but rather a more differentiated epiblast lineage. In order to improve upon this model, it became evident that an alternative starting cell population was necessary in order to more robustly capture the primitive endoderm lineage.

In 2019, two models aimed to achieve this by utilizing extended pluripotent or expanded potential stem cells (EPSCs). By exposing ESCs to a specific chemical cocktail, it was believed that cells could be reprogrammed to a more naïve state, and expressed markers upregulated during the 2-cell to morula stages. Moreover, through chimera experiments injecting EPSCs into 8-cell mouse embryos, it was shown that a single EPSC could contribute to both embryonic and extra-embryonic lineages. By aggregating EPSCs alone[87] or EPSCs and TSCs in combination[88], two new blastoid models showed higher cavitation efficiency and higher instance of primitive endoderm formation. Sozen et al. also showed the ability to generate further derivatives of various lineages, with bifurcation of the TE to polar and mural fates as well as PrE to parietal and visceral endoderm-like cells. While these models offered significant improvement, neither model demonstrated the ability to fully transition to post-implantation morphology *in vitro* or *in vivo.* Although EPSCs were originally shown to have higher developmental potential and differentiate into all three lineages, newer evidence shows that this is not the case[89]p. In particular these cells do not fully recapitulate the extra-embryonic cell types, and is a likely explanation for why current blastoid models do not robustly develop to post-implantation stages, neither *in vitro* or *in vivo*.

Unlike post-implantation stem cell-derived embryo models, which lack a mural trophectoderm, blastoids have the theoretical potential to be implanted within the uterus of a mouse and develop *in vivo.* This capability is particularly exciting given that it could overcome the limitations of *ex utero* culture by harnessing the uterine environment and nutrients of the host animal. Nevertheless, several transfer experiments implanting blastoids into pseudopregnant mice have been performed, but with limited success. In many cases, structures are able to form deciduae, but fail to properly implant and develop the quintessential cup-shape morphology characteristic of post-implantation development. In order to overcome this developmental hurdle, models must first solidify starting cell populations that more closely mimic the natural embryo and have the potential to fully recapitulate the initial lineages and their derivatives.

2.3. Stem cell-derived models: Human

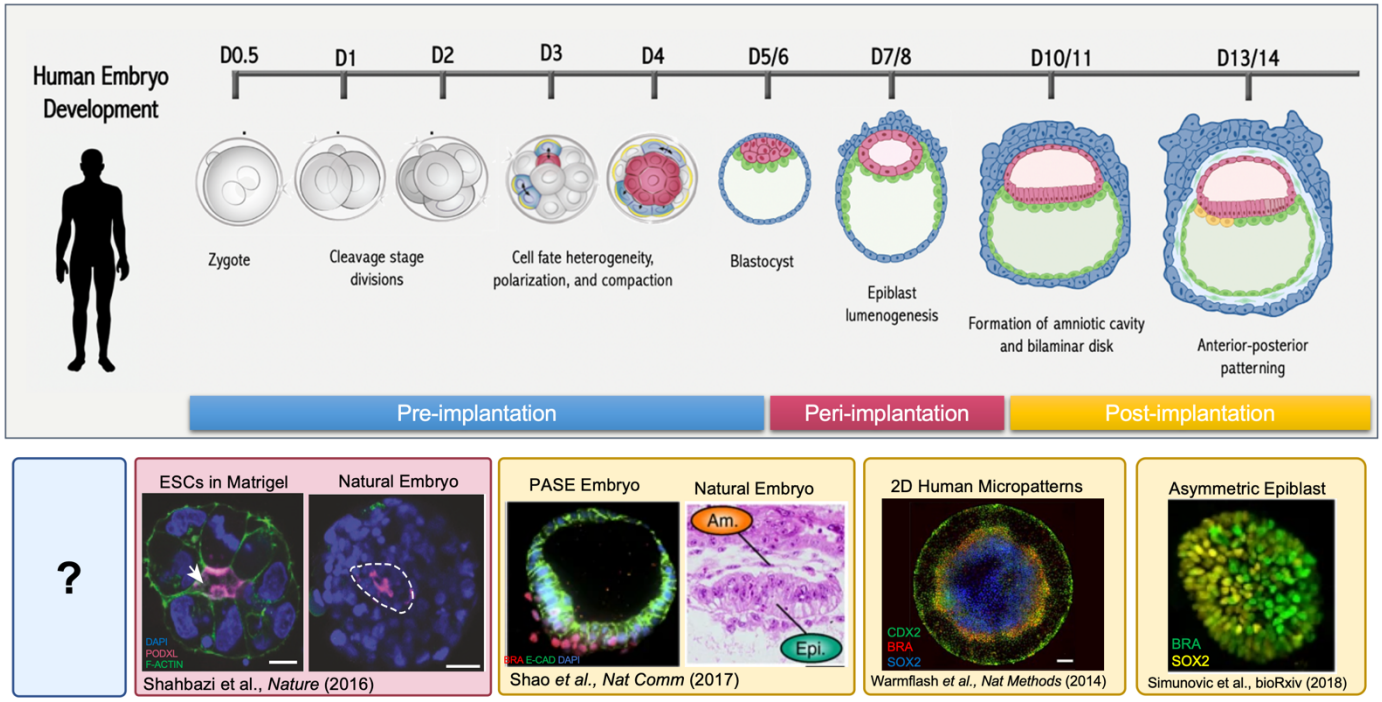


Figure 2.1: Overview of stem cell-derived models in mouse.
Top: Overview schematic of human embryo development, highlighting 3 major stages: pre-implantation, peri-implantation, post-implantation. **Bottom:** Examples of different stem cell-derived embryo systems in human; comparisons to natural embryo are included when possible, but post-implantation culture of human embryos is capped at Day 14. Each model corresponds to a distinct stage of development: ESCs in matrigel compared to the epiblast rosette (Shahbazi *et al.*, 2016); Post-implantation amniotic sac embryoids (PASE) compared to Carnegie Stage 5C natural embryos (Shao *et al.,* 2017); 2D micropatterns showing germ layer differentiation (Warmflash *et al.,* 2014); and asymmetric epiblasts, showing divergence of epiblast and amnion in stem cell aggregates (Simunovic *et al.*, 2018). Prior to 2020, no models for pre-implantation stages existed.

2.3.1. Stem cell derivation in human: ESCs, TSCs, TLCs, and HBLCs

In addition to the various mouse models, several human models have also been established in recent years. Human embryonic stem cells (hESCs) were first isolated from the epiblast of blastocyst-stage embryos in 1998[90]. Unlike the "2iLif" culture condition of mouse ESCs, reminiscent of a naïve pre-implantation epiblast, standard culture for hESCs captures a slightly later stage of development, known as the "primed state"[45]. These cells are believed to more closely mimic cells of the post-implantation epiblast and are typically cultured in media activating the TGFB and FGF signaling pathways.

In addition to hESCs, successful derivation of human trophoblast stem cells (hTSCs) from blastocyst and first trimester placental tissues[91]. Culture of hTSCs is characterized by

activation of Wnt and EGF pathways and simultaneous inhibition of TGF-β, histone deacetylase (HDAC), and Rho-associated protein kinase (ROCK)[91]. Nevertheless, even hTSCs derived from blastocyst-stage embryos appear to mimic a later stage of development resulting in post-implantation cytotrophoblast-like cells. While these cells show capabilities to differentiate into syncytiotrophoblast (STB) and extravillous trophoblast (EVT), it is unclear is they have the potential to contribute to pre-implantation trophectoderm lineages.

In more recent years, attempts have been made to convert hESCs to trophoblast-like cell (TLCs), in hopes that these cells would more closely mimic pre-implantation lineages as opposed to post-implantation cytotrophoblasts[92–94]. Additionally, derivation of hTSC lines is incredibly difficult and limited, leading to minimal genetic diversity and availability of these lines. To overcome these hurdles, several protocols have employed distinct media cocktails to drive differentiation of hESCs to trophectoderm-like states morphologically and transcriptionally. While early work demonstrated that hESCs could be driven towards TE differentiation upon BMP4 induction[95], newer protocols show that TGFB and ERK inhibition is sufficient to generate TLCs in a BMP4-independent manner[96]. This is important because it was recently shown that induction of BMP4 in primed hESCs drives differentiation to amnion-like cells, as opposed to trophectoderm lineages[96]. Another method, termed human trophectoderm stem cells (hTESCs), uses sphingosine-1 phosphate (S1P) to activate YAP signaling, a critical pathway required for trophectoderm specification[92]. S1P, in combination with BMP4 and the activin/nodal inhibitor SB431542, was thus shown to generate *CDX2* positive trophectoderm-like cells that distinctly diverge from standard hTSCs. Although each of these models claims to generate trophoblast-like cells, comparative analyses between models and the natural embryo will be necessary to fully unleash the potential of these systems.

In addition to hESCs and hTSCs, recent efforts to generate cell lines mimicking the extra-embryonic hypoblast have also been attempted. While hESCs and hTSCs can be directly isolated and propagated from blastocyst-stage embryos, hypoblast-like cells (HBLCs) are derived from hESCs by either over-expression of key transcription factors (GATA6[97],

SOX7, and SOX17[98]) or distinct media cocktails that activate Nodal and Wnt signaling, while downregulating FGF receptor signaling[99].

2.3.2. Naïve hESCs

In order to overcome the developmental limitations of primed hESCs, there has been immense efforts in the field to generate naïve hESCs that more closely mimic pre-implantation stages[100–111]. While each of these protocols generate slightly different "naïve" states, they are typically characterized by expression of key pluripotency markers (OCT4, KLF4, KLF2, KLF17 and NANOG), an overall decrease in DNA methylation, and maintenance of naïve reporters (typically OCT4). Some studies have also shown biallelic expression of X-linked transcripts, indicating reactivation of both X-chromosomes, a signature of the naïve pre-implantation epiblast[109]. Lastly, authors have investigated the expression of certain transposable elements, which are under tight epigenetic control during the earliest stages of development[109]. Altogether, these cells are believed to have a signature that greatly differs from primed hESCs, and thus represent a population of cells with greater differential potential similar to that of the pre-implantation morula/ICM.

With the explosion of naïve protocols in recent years, it has become necessary to evaluate the similarities and differences of each system to assess the optimal conditions. In 2017, Liu *et al.* performed a comprehensive study evaluating four conversion protocols, NHSM[100], RSeT, 5iLAF[101,104] and t2iLGöY[107] in the following areas: 1) reprogramming and morphology, verified by immunostaining of key pluripotency markers 2) global transcriptome analysis, 3) genomic stability, and lastly 4) differentiation capabilities[112]. Ultimately this study revealed that 5iLAF and t2iLGöY cells most closely mimic the human pre-implantation ICM at the transcriptional and epigenetic levels. Moreover, while they show robust upregulation of key pluripotency genes, they also show prominent downregulation of primed pluripotency markers.

An alternative protocol to those listed above is the development of expanded potential or extended pluripotent stem cells (EPSCs)[108,110]. While the naïve ground state in mouse has

been established using 2iLif conditions (PD0325901, CHIR 99021, and LIF), it has been shown that these factors are not alone sufficient to generate naïve cells in human. In this protocol, authors therefore performed a thorough chemical screen to identify additional molecules that could promote transition of *human* ESCs to the naïve state as indicated by activation of the *OCT4* distal enhancer (OCT4-DE), which is specific to pre-implantation development. Upon optimization, it was found that 4 additional molecules supplemented to the standard 2iLif condition could promote conversion to the expanded potential state: 1) (S)-(+)-dimethindene maleate (DiM), an antagonist of muscarinic M2 and histamine H1 receptors; 2) minocycline hydrochloride (MiH), a tetracycline antibiotic and PARP inhibitor; 3) IWR-1-endo (IWR-1), a Wnt pathway inhibitor believed to suppress spontaneous differentiation; and finally 4) Y27632, a ROCK inhibitor utilized to minimize apoptosis and maintain self-renewing capabilities. In addition to upregulating key naïve markers, these cells were shown to contribute to both the embryonic and extra-embryonic lineages of inter-species chimeras, indicating enhanced differential potential compared to primed hESCs. Lastly, authors performed global transcriptional analysis comparing hEPSCs to other naïve induction protocols and natural embryos at various stages; this analysis showed that hEPSCs upregulate a certain genes that are also present throughout 2-cell to morula stage development in human. It is important to note that a separate cluster of genes is clearly upregulated in hEPSCs that is not found at any stage in natural pre-implantation development; this indicates that while hEPSCs seem to contribute to lineages *in vivo*, clear divergences exist between these cells and the natural embryo. Further discussion on these cells can be found in subsequent chapters.

These protocols represent an exciting avenue to understand human naïve pluripotency and subsequent differentiation, yet major caveats remain. Firstly, many of these protocols demonstrate cell line specific dependencies, which makes reproducibility very difficult. Moreover, it is unclear if different cell lines produce different naïve signatures despite using identical conversion protocols; this is likely the case considering hESC lines have been shown to have significant genetic variation leading to variability in transcriptional stability, epigenetic states, chromosomal stability and more depending on precise culturing

conditions[113]. Another caveat of these systems is that sustained culture in naïve media has been shown to lead to karyotypic abnormalities and eventually cell lethality[104]. Lastly, while certain signatures of the pre-implantation embryo have been recapitulated in several of these systems, where these models diverge from the *in vivo* counterparts remains underexplored. In future studies, it will be critical to not only compare naïve signatures of these protocols across different cell lines and genetic backgrounds, but to also incorporate comparisons to the natural embryo to better understand these cells.

2.3.3. 2D Micropatterns

While primed hESCs have been shown to differentiate into all three germ layers, they most closely mimic a post-implantation epiblast, and therefore lack the ability to give rise to extra-embryonic lineages[114–116]. Nevertheless, several models have arisen using primed cells including 2D micropatterns and 3D embryoid bodies. In 2014, Warmflash *et al.* demonstrated that by culturing hESCs in a confined geometric environment in the presence of BMP4, cells will differentiate into three distinct rings: an outer trophectoderm-like layer, an inner ectodermal region, and lastly a mesendoderm population that expresses primitive streak markers throughout[117]. Other groups have used this system to demonstrate key pathways regulating epiblast patterning including BMP, Activin/Nodal, and Wnt[118]. 2D micropatterns have also demonstrated abilities to mimic certain aspects of neurulation when exposed to neural induction environments; as such, these systems represent a powerful model to understand neurogenesis and certain neurodegenerative diseases such as Huntington's Disease[119,120]. These systems are powerful tools for understanding the signaling dynamics necessary for epiblast organization, yet they ultimately ignore contributions from extra-embryonic lineages as well as 3D spatial patterning, and are thus limited in their potential.

2.3.4. Spheroids

Similar to the rosettes formed using mouse ESCs, human ESCs were also shown to mimic epiblast polarization and lumenogenesis upon embedding into extra-cellular matrix[54]. Using this model, it was shown that morphogenesis of the epiblast upon implantation is tightly

coupled with the naïve to primed pluripotency transition. If hESCs are maintained in naïve conditions, these cells were able to undergo polarization, yet lacked the ability to open a lumen. Upon naïve exit, these cells, under an Oct4-mediated transcriptional program, were then able to perform lumenogenesis successfully. Additionally, a subsequent study showed that a 48hr pulse of BMP4 is sufficient to induce symmetry breaking in these structures[121]. This was evidenced by unilateral expression of either SOX2 or BRA, denoting establishment of the anterior-posterior axis. Evidence of EMT was also shown via downregulation of the adherent junction protein E-CAD and increased expression of N-CAD in the BRA positive region. A major caveat of this study is that efficiency to generate such asymmetric structures is low, and limits the feasibility of using this model to better understand natural development. Together these systems represent a 3D model that is useful for studying the peri- to post-implantation epiblast; unfortunately, further maturation of the lineages is not observed in these structures such as specification of amnion or generation of primordial germ cells.

2.3.5. Post-implantation amniotic sac embryoids using microfluidics

In 2019, the previous models were greatly improved upon through the use of a microfluidic device. By generating gel pockets of a precise size and geometry within a PDMS chip, Zheng *et al.* were able to significantly improve the accuracy and efficiency of structure formation[122]. Additionally, this system demonstrates higher controllability, as key signaling molecules can be introduced to structures in a gradient dependent manner. These structures were not only able to undergo polarization and lumenogenesis, but also formation of a bipolar embryonic sac, specification of primordial germ cells (PGCs), and generation of primitive streak-like cells. While this model represents the potential for hESCs to mimic events of pre-gastrulation epiblast development, it is uncertain how closely these amnion-like cells replicate their *in vivo* counterparts in the natural embryo. Performing such experiments will remain difficult as access to these stages is incredibly limited, and *ex utero* culture of human embryos does not robustly capture amnion specified cells.

2.3.6. Human Blastoids

While extensive work has been done to model the post-implantation epiblast, many other aspects of early human development remain ignored. Extra-embryonic lineages, for example, and their contributions to epiblast organization remain largely unexplored. Additionally, models that capture events prior to implantation are severely lacking. With the emergence of naïve hESCs, cells that could replicate aspects of pre-implantation stages, many researchers began to wonder: can these cells be used to generate human blastocyst-like structures? Previous models had utilized primed hESCs, which can only differentiate to later derivatives of the embryonic epiblast; naïve hESCs, on the other hand, represent a population of cells that could give rise to embryonic *and* extra-embryonic lineages, and could thus potentially model the full human conceptus. At the onset of my PhD in 2019, there had been no published models that could generate blastocyst-like structures in human. In the following chapter I outline my work, done alongside Dr. Berna Sozen, in which we generate human blastocyst-like structures utilizing hEPSCs[123]. In most recent years, several additional models for human blastoids have been published; these are more thoroughly discussed in Chapter 5.

2.4. Concluding remarks

Stem cell-derived embryo models, in both mouse and human, represent exciting avenues for research in early embryogenesis. By generating robust models that faithfully recapitulate the events of mammalian development, we could overcome many of the ethical and technical limitations that hinder embryonic research today. In order for that to happen, however, several improvements must be made. First, many of these models show cell line dependencies and thus severely limit the reproducibility and efficiency of these models. Understanding how epigenetic states and genetic backgrounds affect differentiation of cell lines will be a critical step towards wide adoption of these protocols. Secondly, understanding where these models compare *and contrast* to the natural embryo will remain a top priority in evaluating the efficacy of these models. Finally, it is important to note that *in vitro* culture of even natural embryos will show divergences from *in vivo* development. Additional experiments understanding the artifacts of *in vitro* culture could therefore be an interesting direction to improve upon these systems.

Chapter 3

RECONSTRUCTING ASPECTS OF HUMAN EMBRYOGENESIS WITH PLURIPOTENT STEM CELLS

3.1 Abstract

Understanding human development is of fundamental biological and clinical importance. Despite its significance, mechanisms behind human embryogenesis remain largely unknown. Here, we attempt to model human early embryo development with expanded pluripotent stem cells (EPSCs) in 3-dimensions. We define a protocol that allows us to generate self-organizing cystic structures from human EPSCs that display some hallmarks of human early embryogenesis. These structures mimic polarization and cavitation characteristic of pre-implantation development leading to blastocyst morphology formation and the transition to post-implantation-like organization upon extended culture. Single-cell RNA sequencing of these structures reveals subsets of cells bearing some resemblance to epiblast, hypoblast and trophectoderm lineages. Nevertheless, significant divergences from natural blastocysts persist in some key markers, and signaling pathways point towards ways in which morphology and transcriptional-level cell identities may diverge in stem cell models of the embryo. Thus, this stem cell platform provides insights into the design of stem cell models of embryogenesis.

3.2. Introduction

Human life starts at fertilization with the union of the sperm and the egg to form the zygote. This unique totipotent cell undergoes continuous cleavage divisions without any increase in size, resulting in a sphere known as the morula four days after fertilization. Further division and cell differentiation results in a hollowed structure known as the blastocyst on the fifth day of development (Fig. 3.1A). By the blastocyst stage, two main groups of cells become defined: the inner cell mass (ICM), which will form the embryo proper and the first extra-embryonic tissue, the trophectoderm (TE), an epithelium that gives rise to the cells of the placenta. Just before embryo implantation on the sixth day, the inner cell mass starts to

differentiate into the epiblast (EPI) and hypoblast (HYPO), which will give rise to all embryonic cells and the extra-embryonic yolk-sac, respectively[30,124]. Following implantation, the EPI undergoes a series of morphological changes leading to the formation of a 3-dimensional (3D) rosette, which then forms a flattened disc-shaped structure that will initiate gastrulation on day fourteen.

The development of the human embryo occurs within the body of the mother, making it difficult to study in mechanistic detail. Although recent in vitro culture methods have advanced our abilities to study aspects of human embryo development ex-utero [6–8] surplus human embryos donated to research are rare and their use is subject to considerable ethical and legal restrictions . Due to these reasons, knowledge of the critical developmental steps allowing formation of the blastocyst stage embryo and its subsequent remodeling at early post-implantation remain largely unknown. Thus, the generation of several stem cell-derived models that recapitulate unique stages of mouse [80,81,86–88,126–130] human embryo [117,121,122,131–134] development in vitro have been invaluable.

Here, with these considerations in mind, we test the hypothesis that human pluripotent stem cells (hPSCs) under certain conditions could undergo self-organization into 3D embryo-like structures. Recent studies showed that PSCs can be reprogrammed to a molecular state, termed extended or expanded pluripotency (EP), that has developmental potency for both embryonic and extra-embryonic cell lineages[108,110,135,136]. We have therefore asked whether hPSCs that are grown under EP conditions (termed hEPSCs) and cultured with a combination of appropriate growth factors and/or inhibitors can capture aspects of early embryonic lineage development in 3D culture. We show that the resulting structures recapitulate some of the morphological and gene expression features of embryonic days 3 to day 9/10 of natural human embryogenesis with limited developmental potential. Single-cell RNA sequencing (scRNA-seq) further confirms that these structures recapitulate some aspects of blastocyst gene expression, with notable divergences. We anticipate that the future applications of this system can give insight into regulatory processes of cellular differentiation in human embryo

development whilst also highlighting ongoing challenges both in specifically understanding the multi-potency state of EPSCs and broadly modelling human embryogenesis in vitro.

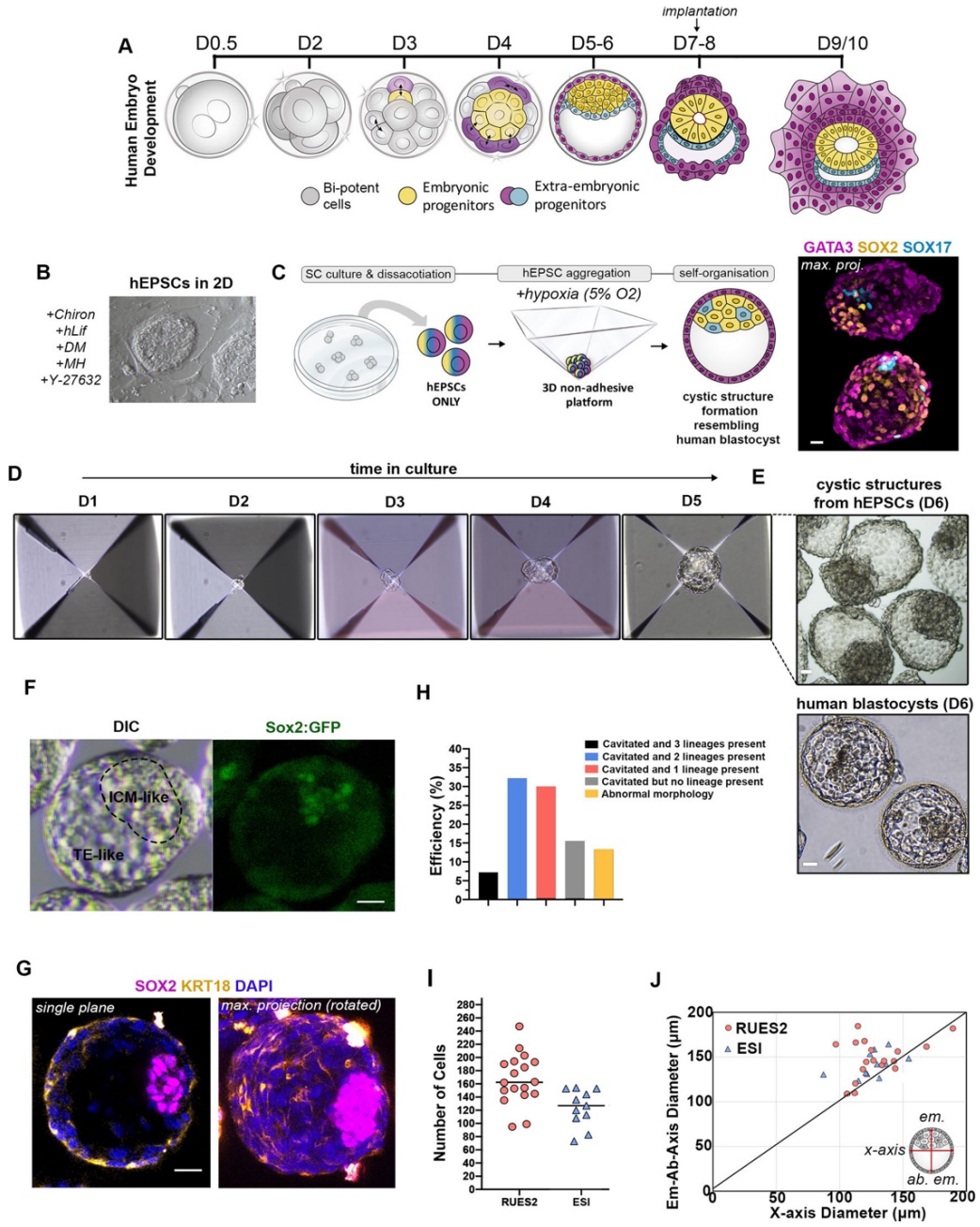


Figure 3.1: A 3D system from hEPSCs to mimic embryo-like events.
A. Scheme for natural human pre/peri-implantation embryo development. **B.** A representative dome-shaped naïve pluripotent hEPSC colony in 2D culture. Representative of at least 10 independent experiments. **C. Left**: Schematic of an AggreWell and 3D aggregation protocol with hEPSCs. **Right**: Representative examples of cystic structures generated from a typical experiment after 4 days demonstrate three lineages, resembling blastocyst stage natural embryo. Representative of at least 3 independent experiments. Lineage markers: SOX2, yellow; GATA3, magenta, and SOX17, cyan. n = 10 experiments. **D.** Representative phase-contrast images of hEPSC multicellular aggregates in AggreWell at the indicated time points during 3D culture. **E.** Phase-contrast images of cystic structures from hEPSCs at D6. (top) and natural human blastocysts at D6 (bottom). Representative of at least 3 independent experiments. **F.** A representative cystic structure generated from RUES2 hEPSC line with SOX2- fluorescent reporter. Representative of at least 2 independent experiments. G A representative structure immunostained for SOX2 in magenta, KRT18 in yellow to label inner compartment and outside epithelium, respectively. DAPI is shown in blue. Maximum projection image is shown on the right. n = 50 structures, 3 experiments. **H.** Efficiency quantification showing the number of structures with a cavity and identifiable lineage segregation. Present EPI-like and HYPO-like cells were determined by positive expression of SOX2 and SOX17, respectively, within inner compartment as seen by IF staining. Present TE-like cells were determined by positive expression of GATA3 or KRT18 in outer cells observed by IF staining. n = 186 structures, 2 experiments. **I.** Quantification for cell numbers in individual cystic structures generated from two established hEPSC lines, ESI017 (n = 11), RUES2 (n = 18). **J.** Measurements of axial diameters in cystic structures from ESI017 (n = 11), RUES2 (n = 18) hEPSC lines. Illustration on right shows the two axes measured. All scale bars in the figure indicate 20 µm.

3.3. Results

3.3.1. Self-organization of human EPSCs

We first converted hPSCs to hEPSCs through a minimum of 5 passages (Materials and Methods). The resulting cells acquired some morphological features characteristic of pluripotent cells in the naïve state of pluripotency, including dome-shaped colony formation, as supported by earlier observations [108] (Fig. 3.1B). However, we could also observe flat cell colonies, a morphological feature characteristic of pluripotent cells in the primed state, present in different ratios after each passage, suggesting the presence of a mixed population of cells in different pluripotent states under EP culture conditions (Fig. 3.2A). Using a multi-inverted-pyramidal microwell-based 3D culture system that we previously described to facilitate self-organization of mouse embryonic and extra-embryonic stem cells [81,82,88] we seeded small numbers of hEPSCs (5-6 cells per microwell) to enable their aggregation and subsequent self-organization (Fig. 3.1C). We first observed that the in vitro culture media normally used for the culture of natural human pre-implantation embryos promoted the formation of cavitated cystic structures (Fig 3.2B, see Methods). Aiming to support the maintenance of pluripotency and to promote TE-like differentiation, we mixed 2 parts of this medium with 1 part of EP[108] and 1 part of hTSC[91], two different stem cell base media (without the addition of any growth factors or inhibitors, see Materials and Methods). We observed that conditions of low oxygen tension (5% O2, similar to our previous conditions for mouse blastoid formation[88] and for the development of natural human blastocysts[137]) facilitated the formation of cavitated structures (Fig 3.2C).

We next screened various growth factors, cytokines, and small molecules at widely adopted concentrations[121,134] as previously published, and to identify conditions facilitating cavity and early lineage formation in these structures (Fig. 3.3A-E). We found that a combination of BMP4 (20 ng/ml), the WNT agonist CHIR99021 (2 μM), FGF2 (40 ng/ml), and ROCK inhibitor Y-27632 (5 μM) during the first 48 h of 3D culture enhanced cell survival and promoted formation of compact cellular aggregates (Fig. 3.3A-F). Additionally, we pulsed the cells with the ALK5 kinase inhibitor A83-01 (2 μM) to promote TE differentiation[91] for

the first 48 h of 3D culture, and removed this inhibitor after this time to avoid a complete loss in pluripotency. Concomitantly, the concentration of FGF2 was decreased by half (20 ng/ml) for the same purpose [121]. Using this optimized condition, we observed the emergence of cavitated structures, 3 to 4 days after cell seeding (Fig. 3.1D). By day 6 of 3D culture, the structures exhibited a blastocyst-like morphology, forming a cohesive single outside layer, with an enlarged cavity, and an internal acentric compartment (Fig. 3.1E), of which 7.2% expressed the markers of the three blastocyst lineages, as judged by immunofluorescence analysis of selected markers (Fig. 3.1F-H). The average cell number and diameter of these hEP-structures were comparable to those of human blastocysts[138] (Fig. 3.1I, J).

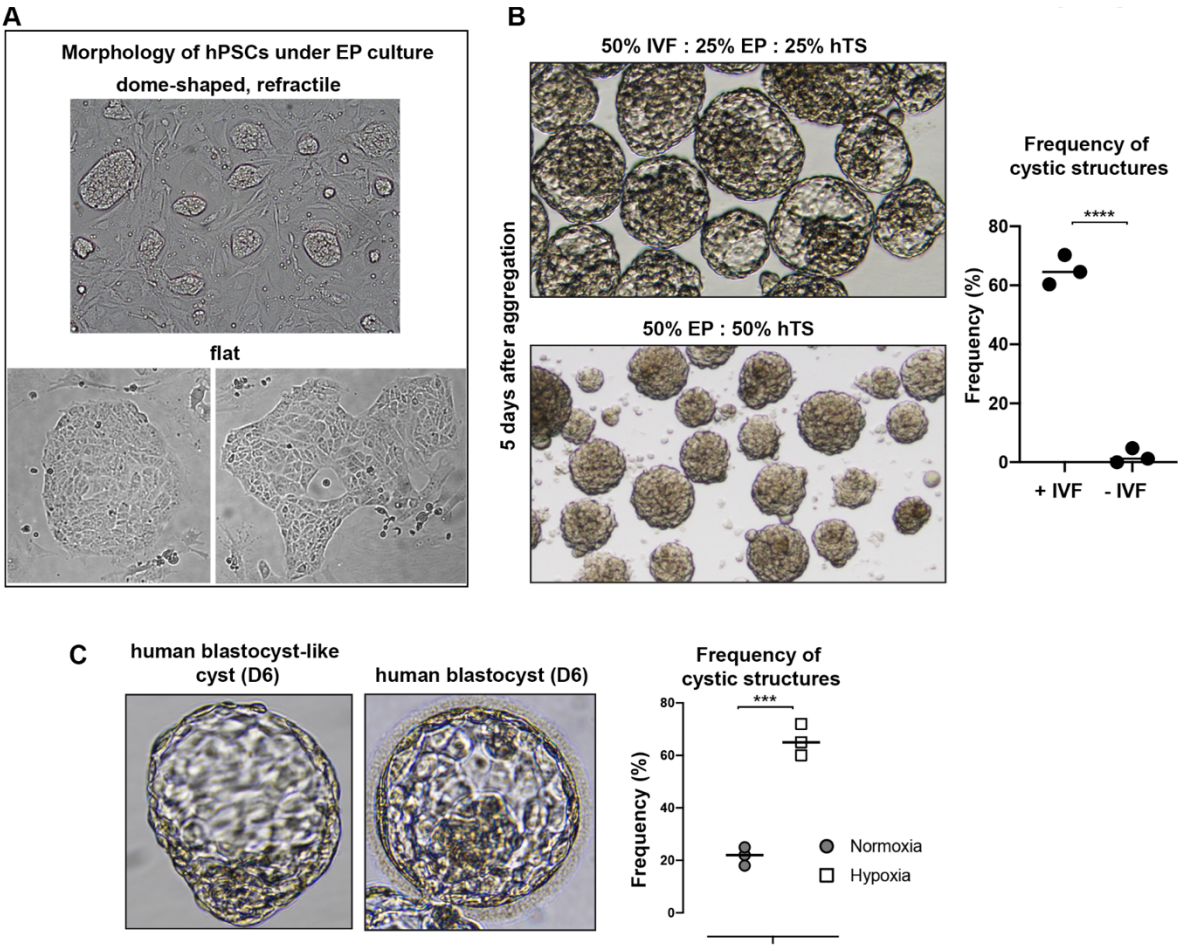


Figure 3.2: Characterization of cellular and structural morphologies under 2D and 3D culture conditions for hEPSCs.

A. DIC images showing varied colony morphology of hPSCs under EP culture conditions. The top image shows dome-shaped morphology, characteristic of pluripotent cells at the naïve state, while the bottom images show flat colonies, characteristic of cells in the primed state. Representative of at least 10 independent experiments. **B.** DIC images of hEP- structures grown for 5 days with (top) and without (bottom) IVF media. Quantification on the right shows the frequency of cystic structure formation in each condition (+/- IVF). Two-sided Student's t-test; ****$p<0.001$; 3 experiments. Error bars show S.E.M. **C.** Comparison of hEP- structures at day 6 (D6) of 3D culture to the natural human blastocyst at the same developmental time point. Quantification on the right shows the frequency of cystic structure formation under 20% O_2 (Normoxia; grey circle) and 5% O_2 (Hypoxia; white square) conditions. Two-sided Student's t-test; $p=0.005$; 3 experiments. Error bars show S.E.M. All scale bars in the figure indicate 20 um.

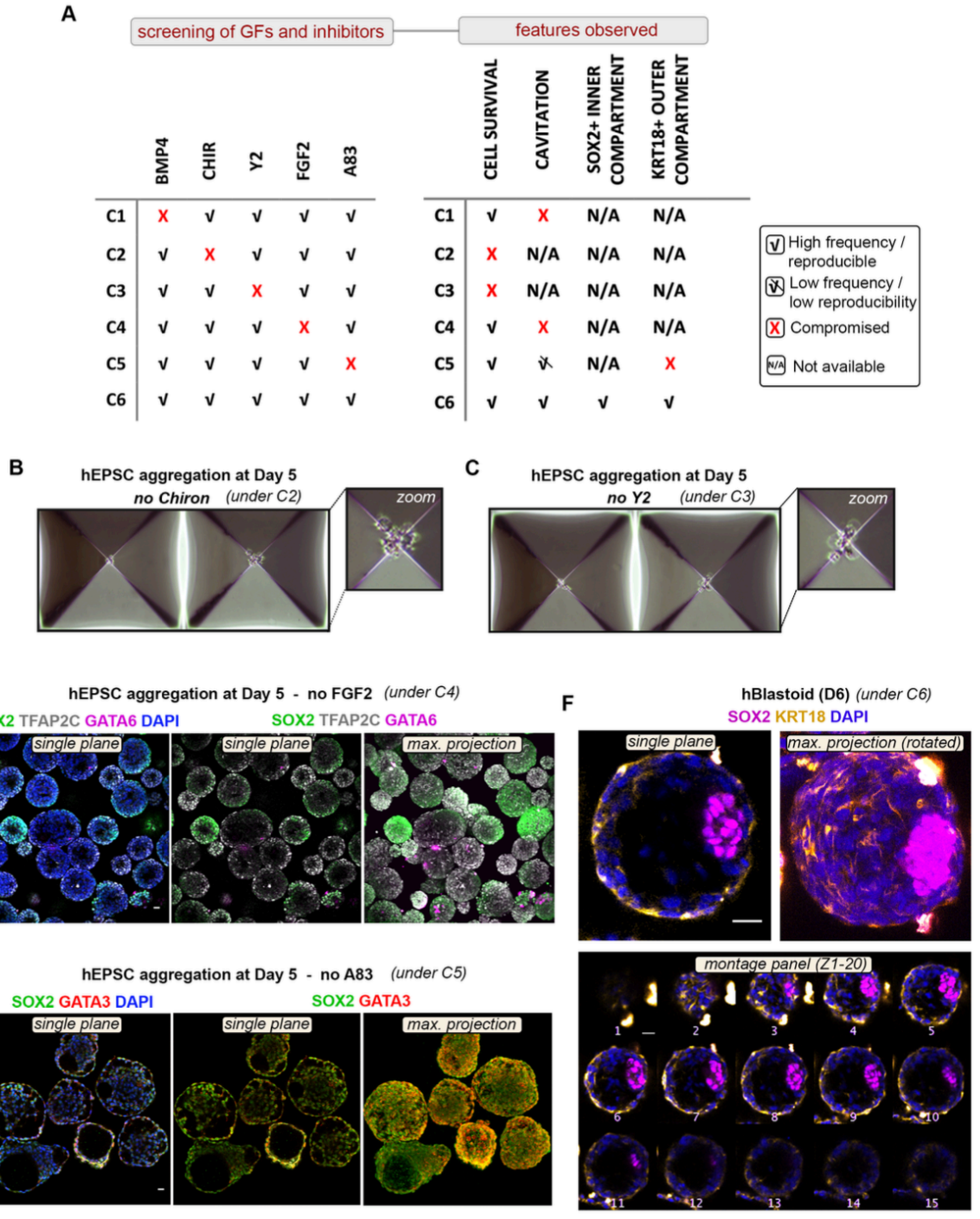


Figure 3.3: Screening of growth factors and inhibitors on hEP-derived structures. A. Table outlining tested media conditions with and without specific growth factors and inhibitors (left), and observed features from each of these conditions (right). **B, C.** Images show failure of cell survival grown for 5 days without CHIR99021 (Chiron, condition 2, C2) or Y - 27632 (Y2; condition 3, C3), respectively. **D.** Immunostaining showing expression of TFAP2C (white), and GATA6 (magenta) in structures grown under condition 4 (C4) for 5 days. DAPI shows nuclear staining in blue. Representative of at least 2 independent experiments. **E.** Immunostaining showing expression of GATA3 (red), and SOX2 (green) in structures grown under condition 5 (C5) for 5 days. DAPI shows nuclear staining in blue. Representative of at least 2 independent experiments. **F.** Immunostaining for SOX2 (magenta), and KRT18 (yellow) in representative structure grown under condition 6 (C6) until Day 6. Images are shown in a single plane view (top left), maximum projection (top right), and as a montage panel of Z-stack slice 1-20. All scale bars in the figure indicate 20 um. Representative of at least 2 independent experiments.

3.3.2. EPSC aggregates bear similarities to early human embryo

The first lineage segregation event begins with compaction and cell polarization in the mouse embryo at the 8-cell stage[139]. Only recently have studies begun to shed light on these events in human embryogenesis[23,24]. Hence, we utilized our platform to analyze the establishment and dynamics of cell polarization at the early timepoints of multicellular aggregate formation. We observed the assembly of intercellular junctions, characterized by basolateral localization of E-CADHERIN (Fig. 3.4A). At the apical surface, we found distinct enrichment of F-ACTIN and PARD6 within the first 48 h of cell aggregation (Fig. 3.4A), indicative of cell polarization in hEP-structures. Next, we analyzed spatiotemporal expression of the transcription factor GATA3, as a marker of TE specification in human embryogenesis. GATA3 was present in the nucleus within both polarized and non-polarized cells at day 2 and day 3 of 3D culture, although its intensity was significantly higher in polarized cells showing apical enrichment of PARD6 (Fig. 3.4A, B). These findings correlate with observations on natural human embryos at the morula stage (Fig. 3.4C).

We have recently shown that the PLC-Protein Kinase C (PKC) pathway controls cellular polarization at early stages of mouse embryo development[22]. We, therefore, treated our hEPSC 3D cultures with 2μM and 3μM of the PLC inhibitor, U73122. PLC inhibition resulted in a reduction of the nuclear GATA3 signal intensity (Fig. 3.4D, E, 3.5A), which correlated with a decrease in the apical enrichment of PARD6 (Fig. 3.4D, E, 3.5B). To confirm the relationship between polarization and outer cell commitment, we next used siRNA transfection to knockdown (KD) *PLCB1* in order to deplete PLC activity in cells during 3D aggregation (Fig. 3.4F). The depletion of PLC activity was confirmed by qRT-PCR that RNAi depletion of *PLCB1* was effective (Fig. 3.5C). In agreement with the previous data, we found a significant reduction in both PARD6 and GATA3 expression in hEPSC aggregates at day 3 (Fig. 3.4F). Thus, our results may suggest a role for the acquisition of apicobasal polarity in promoting the expression and nuclear localization of GATA3 to drive TE specification during development of hEP-structures.

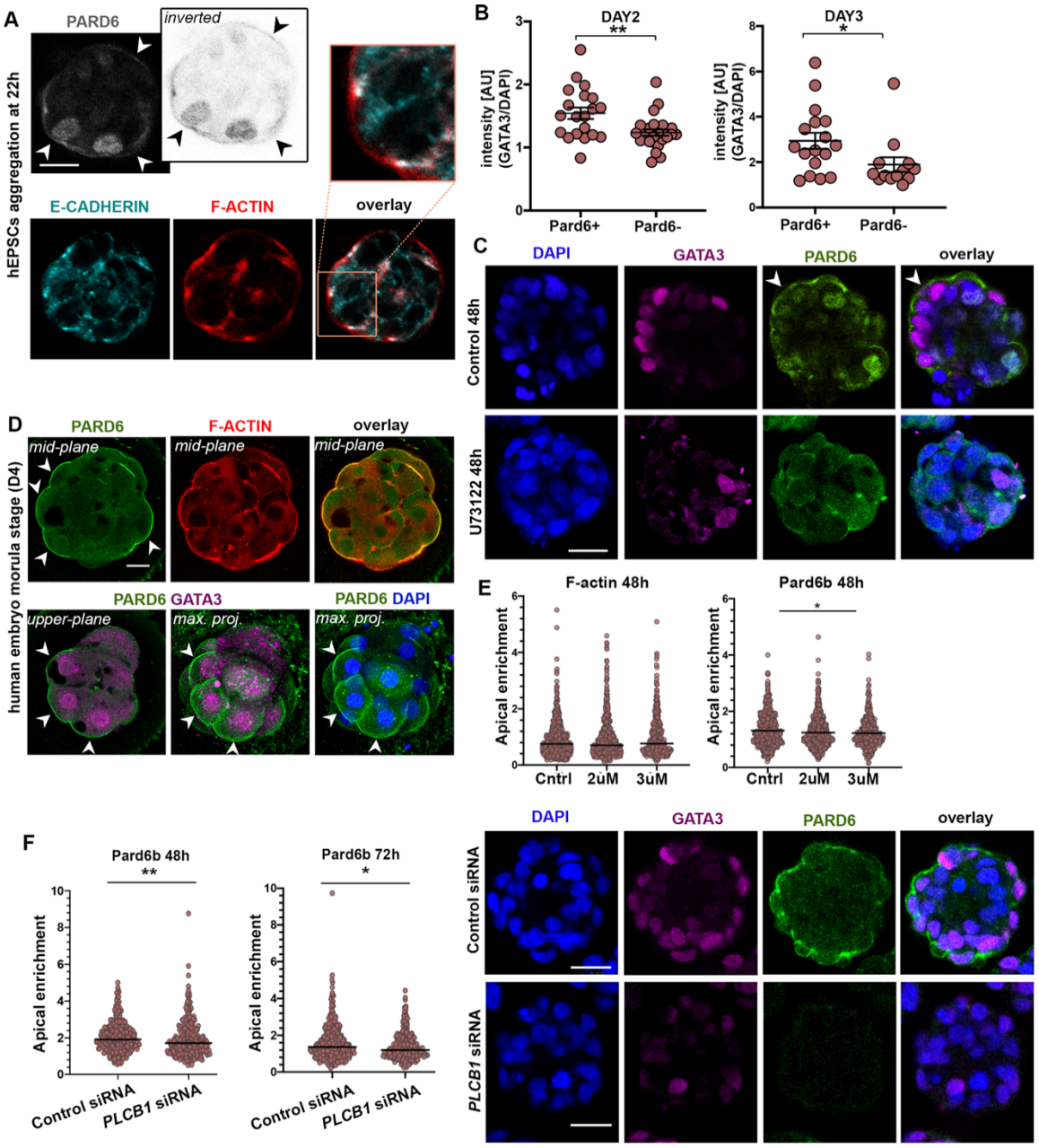


Figure 3.4: hEPSC aggregates show similarities to pre-Implantation embryo development.
A. Immunostaining of hEPSC aggregates at 22 h for PARD6 (grey), F-ACTIN (red), and E-CADHERIN (cyan). $n = 300$ aggregates, 3 experiments. **B.** Quantification of GATA3 expression in cells with or without PARD6 apical enrichment observed in cells within Day 2 and Day 3 of multicellular aggregates. All measurements normalized to DAPI. Two-sided Student's t-test; $p = 0.0033$ for Day 2; $p = 0.0433$ for Day 3; 3 experiments. Error bars represent S.E.M. **C.** Immunostaining of control and U73122-treated hEPSC aggregates at 48 h for PARD6 (green) and GATA3 (magenta). $n = 300$ aggregates, 3 experiments. **D.** A representative natural human embryo at morula stage (D4) stained for PARD6 (green), F-ACTIN (red), GATA3 (magenta). White arrowheads indicate apical PARD6 enrichment in the polarized cells with nuclear GATA3 expression. DAPI is shown in blue. **E.** Apical enrichment quantification of F-ACTIN and PARD6b at 48 h in multicellular structures with or without addition of PLC inhibitor (U73122). Control groups received no inhibitor, while the two experimental groups were treated with either 2 uM or 3 uM U73122. Each dot represents one analyzed cell. $p = 0.0333$, Kruskal-Wallis test with Dunn's multiple comparisons test. Data is shown as mean S.E.M. $n = 3$ experiments. Also see Extended Data Fig. 3b. **F.** Left: Quantification of Pard6b apical enrichment at 48 h ($p = 0.006$) and 72 h ($p = 0.0227$) in structures treated with either control siRNA or *PLCB1* siRNA. Each dot represents one analyzed cell. Two-sided Mann-Whitney test. Data is shown as mean S.E.M. $n = 3$ experiments. Right: Immunostaining of GATA3 (magenta) and PARD6 (green) in structures treated with either control siRNA (top) or *PLCB1* siRNA (bottom). DAPI is shown in blue. All scale bars in the figure indicate 20 um.

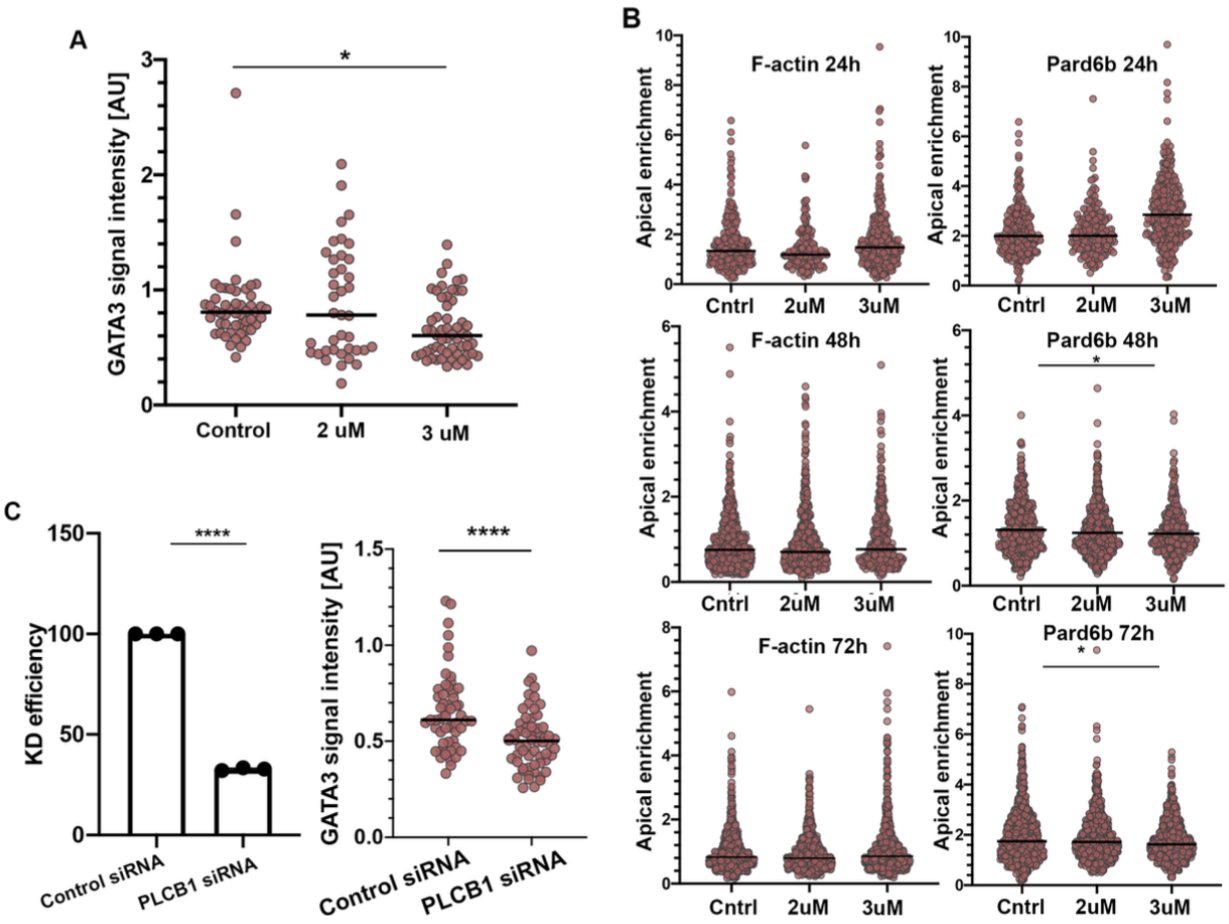


Figure 3.5: PLC-Protein Kinase C (PKC) pathway suppression affects polarity establishment. A. Quantification of nuclear GATA3 signal intensity in control and experimental groups treated with either 2uM or 3uM of selective PKC inhibitor, U73122. All measurements normalized to DAPI. Each dot represents one analyzed cell. p=0.0173, ANOVA with a multiple comparisons test. Data is shown as mean S.E.M. n=6 aggregates, 3 experiments. **B.** Apical enrichment quantification of F-ACTIN and PARD6b at 24, 48h and 72h in multicellular structures with or without addition of PLC inhibitor (U73122). Control groups received no inhibitor, while the two experimental groups were treated with either 2uM or 3uM U73122. Each dot represents one analyzed cell. p=0.0033 for Pard6b 48h; p=0.0267 for Pard6b 72h, Kruskal-Wallis test with Dunn's multiple comparisons test. Error bars show S.E.M. n= n=300 aggregates, 3 experiments. **C. Left:** Quantification of PLCB1 knock-down efficiency in groups treated with either control siRNA or *PLCB1* siRNA as determined by RT- PCR. Values were normalized against GAPDH. ****p<0.001, Two-sided Student's t-test. Approximately 800 structures were pooled for RT-PCR per group from 3D culture. n = 3 replicates. Error bars show S.E.M. **Right:** Quantification of nuclear GATA3 signal intensity in groups treated with either Control siRNA or *PLCB1* siRNA. All measurements normalized to DAPI. Each dot represents one analyzed cell. ****p<0.001, Two-sided Student's t-test. n=6 aggregates, 3 experiments. Error bars show S.E.M.

3.3.3. Differentiation into embryonic and extra-embryonic lineages

We next sought to investigate the formation of the blastocyst lineages upon cavitation of hEP-structures from day 4 onwards. We first used qRT-PCR to examine the expression level of core factors involved in establishing human blastocyst-like lineage identity (Fig. 3.6A). This analysis revealed that genes involved in TE specification, including *PLAC8, CDX2, KRT8,* and *KRT18*, were induced upon formation of cystic structures although *GATA3* showed only a marginal increase compared to other molecular determinants of TE identity (Fig. 3.6A). As expected, crucial transcription factors required for pluripotent EPI specification, including *NANOG* and *POU5F1*, showed similar levels of expression in cystic structures as in hEPSCs cultured in 2D, whereas *KLF4* was significantly upregulated in the cystic structures (Fig. 3.6A). Finally, we found that the expression of the core HYPO lineage determinant genes, *PDGFRA* and *GATA6*, were highly enriched in cystic structures although *SOX17* did not follow this trend (Fig. 3.6A).

In order to confirm these results spatially and on a protein level, we performed immunofluorescence analysis with some well-known lineage markers. In accordance with the findings from qRT-PCR, we observed enrichment for KRT18 in the outside, and expression of OCT4/SOX17 in the inner compartment (Fig. 3.6B). We further observed specification of the inner compartment with a second set of markers, SOX2/FOXA2[140] (Fig. 3.6C). At day 4, some structures displayed constitutive expression of GATA3 in the outer cell layer while maintaining expression of the hPSC/EPI marker OCT4 (10/23 structures scored) (Fig. 3.6D). At later time-points in culture (Day 6, see Materials and Methods), some structures maintained GATA3 expression in the outside layer, although this enrichment became mostly cytosolic rather than nuclear (53/135 structures scored) (Fig. 3.6E). These late-timepoint structures also showed poor expression of E-CADHERIN at day 6 (Fig. 3.6E). This result likely indicates a deficiency in junction assembly during the late cavitation process and may explain the compromised expression of some TE-specific markers as in vitro development progresses.

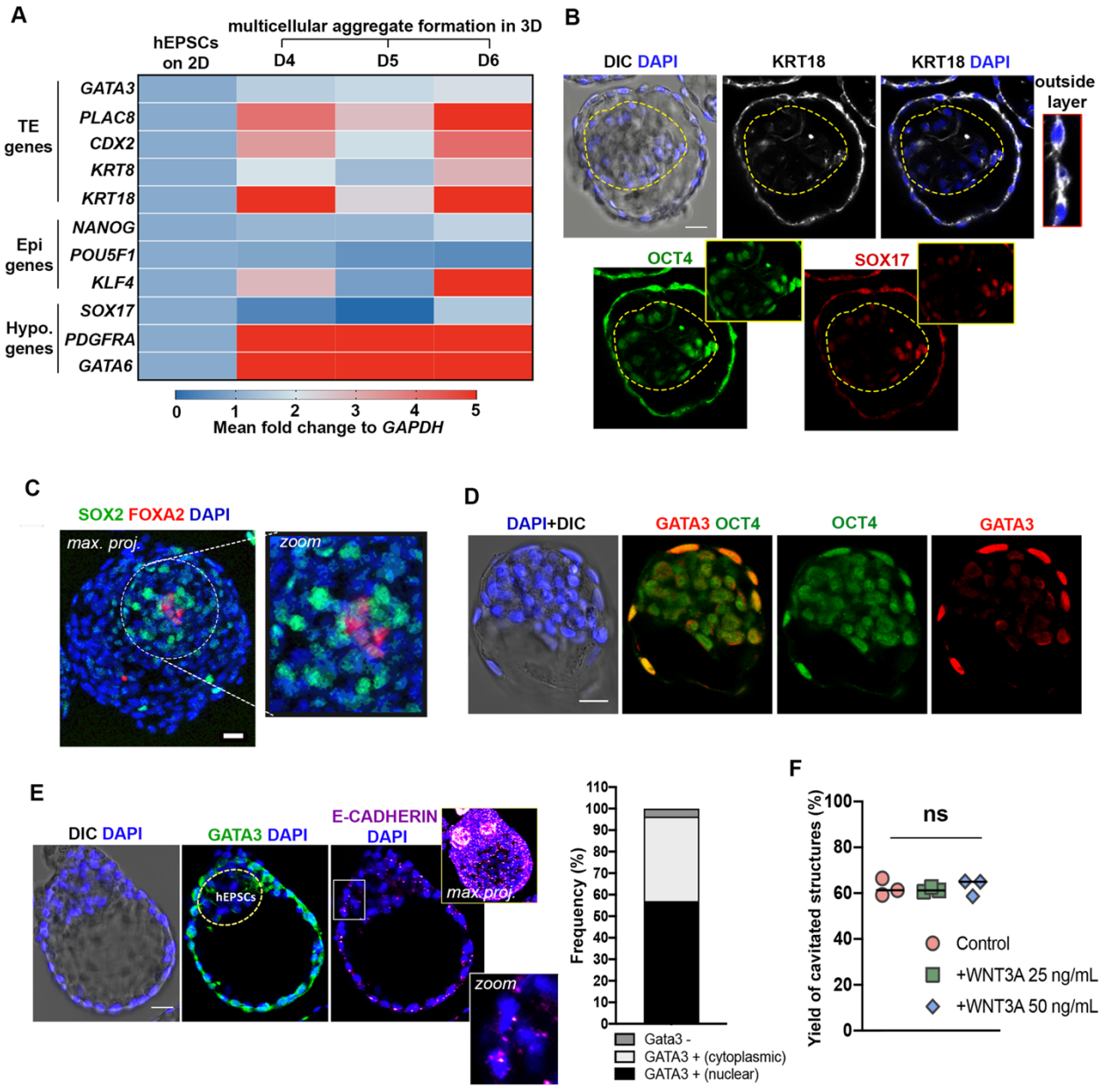


Figure 3.6: Specification of blastocyst lineages in hEP-structures.
A. Bulk qRT-PCR analysis of blastocyst lineage marker genes in EPSCs in 2D, and multicellular aggregates at day 4, 5, 6 formed in 3D represented as a heatmap of global ΔΔCt (fold-change) to GAPDH. 20 cystic structures were pooled per group from 3D culture and a minimum of 10 K hEPSCs were collected from the 2D culture. **B.** Immunofluorescence staining of structures generated from hEPSCs at day 5 for OCT4 (green), KRT18 (white) and SOX17 (red). Zoom image on the right shows cells with KRT18 expression. DAPI is shown in blue. n = 30 structures, 3 experiments. **C.** A representative structures generated from hEPSCs at day 5 stained for SOX2 (green) and FOXA2 (red) to suggest Epi/Hypo-like inner compartment (zoom on the right). Image presented as maximum projection. n = 20 structures, 2 experiments. **D.** Immunofluorescence staining of structures generated from hEPSCs at day 4 for OCT4 (green) and GATA3 (red). DAPI is shown in blue. n = 10/23 structures, 2 experiments. **E.** Left: Immunofluorescence staining of GATA3 (green) and E-CADHERIN (magenta) in a representative structure at day 6. Right: Quantification shows frequency of structures at day 6 of 3D culture showing Gata3 nuclear expression (57.03%, 77/135 structures scored); Gata3 cytoplasmic expression (39.25%, 53/135 structures scored); no detectable Gata3 expression (3.70%, 5/135 structures scored). **F.** Quantification shows frequency of cavitated structures in control and WNT3A-supplemented culture. WNT3A is applied in either at 25 or 50 ng/mL concentration. One-way ANOVA with multiple comparisons, p = 0.8473. ns, not significant. n = 450 for control; n = 483 for 25 ng/mL WNT3A; n = 419 for 50 ng/mL WNT3A (see source data). 3 independent experiments. Error bars show S.E.M. All scale bars in the figure indicate 20 um.

It has been suggested that WNT3A supplementation promotes the cavitation and thus TE-like lineage identity in in vitro mouse blastoid formation10, which correlates with canonical WNT expression in the TE lineage during mouse blastocyst development37. When we tested this possibility, we found that addition of WNT3A to the culture media did not make any significant difference to the yield of cavitated structures at day 6 (Fig. 3.6F). This suggests that WNT3A may function differently in human development than in mouse, supporting previous claims[133,141].

3.3.4. Post-implantation remodeling of hEP-structures

We next tested the developmental capacity of these hEP-structures to develop beyond implantation stages by culturing them in our previously established human embryo in vitro culture (IVC) platform4 (Fig. 3.7A). Within 24 h in IVC, the EP-structures reorganized into post-implantation-like morphology, 60% of which had a SOX2 positive EPI-like inner compartment surrounded by a KRT18 and GATA3 positive extra-embryonic-like compartment (Fig. 3.7B–D) with some structures also specifying a few FOXA2-expressing cells suggesting a HYPO-like specification (Fig. 3.7C). Significantly, we found that within 24 h in IVC, SOX2 positive cells in the EPI-like inner compartment became radially organized around a small central lumen (Fig. 3.7D). The formation of a small lumen was confirmed by PODXL expression (Fig. 3.7E). This indicates that hEP-structures are able to undertake some cellular rearrangements characteristic of early post-implantation human morphogenesis[6–8].

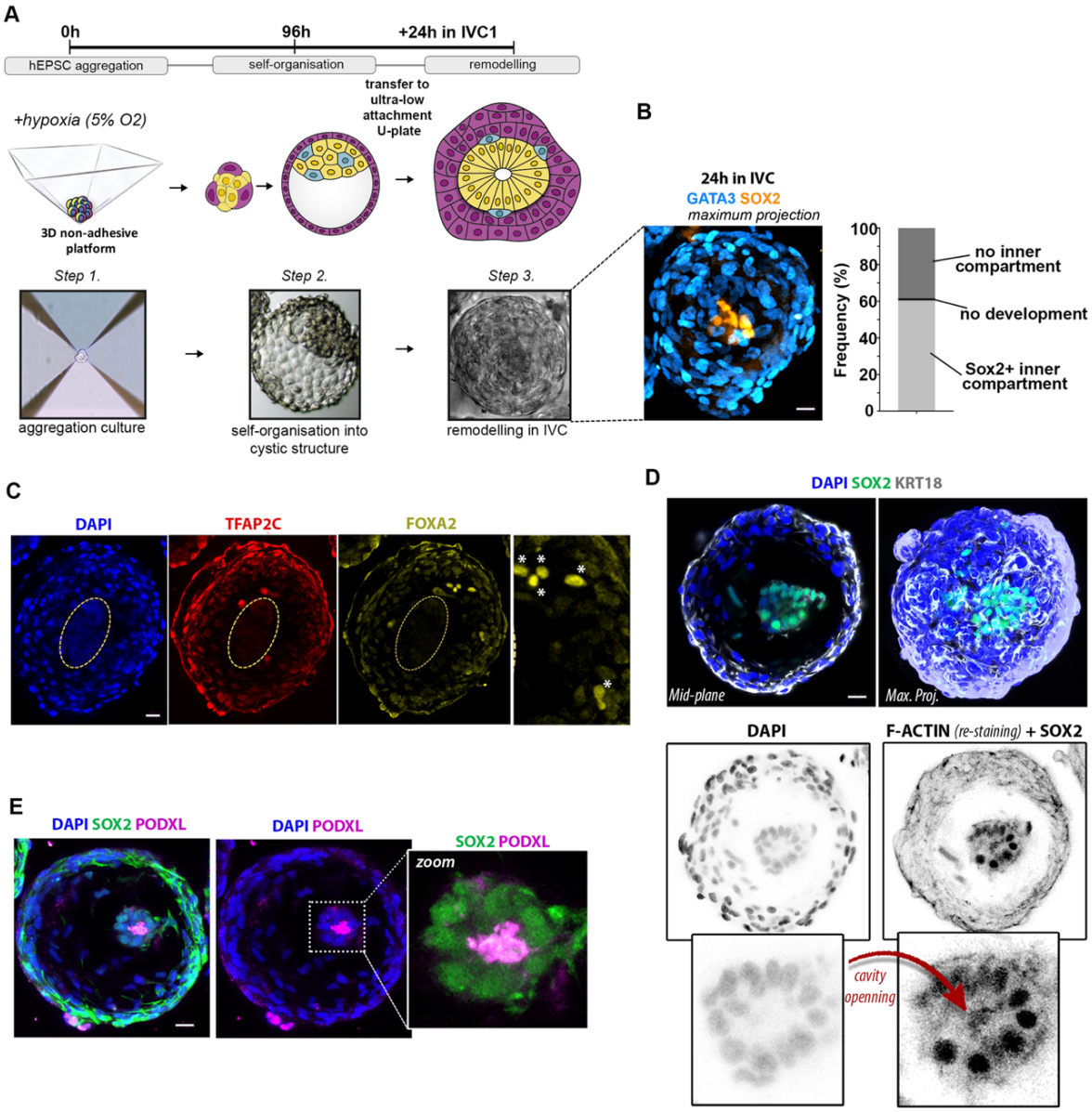


Fig. 3.7: Cultured hEPSC-derived cystic structures demonstrate implantation-like morphological remodeling.
A. Illustration detailing the process of in vitro structure formation in three steps: hEPSC aggregation, self-organization into cystic structures, and post-implantation re-organization in IVC media (see Methods). Below are phase-contrast images showing a representative structure at each of these steps. **B. Left:** Maximum projection of a representative post-implantation-like structure immunostained for GATA3 (cyan) to reveal extra-embryonic-like and SOX2 (yellow) for embryonic-like compartments. **Right:** quantification shows frequency of structures cultured in IVC media for 24 h that showed SOX2 + inner compartment (60.4%, light grey), no inner compartment (38.4%, dark grey), or no development (1.2%, black). n = 260 structures scored in 3 experiments. **C.** Immunostaining showing the expression of TFAP2C (red, an extra-embryonic marker) and FOXA2 (yellow, a HYPO marker) in a post-implantation-like structure cultured in IVC for 24 h. Zoomed image on the right shows FOXA2 expressing cells. (asterisks). Yellow dashed-lines demarcates embryonic inner compartment. Representative of at least 3 independent experiments. **D.** Immunostaining of post-implantation-like structure cultured in IVC media for 24 h. The top panels show a mid-plane (left) and maximum projection (right) view of a representative structure with the pluripotent compartment marked by SOX2 expression (green), surrounded by cells marked by KRT18 (white). DAPI for nuclear staining is in blue. The bottom panels show inverted images for better clarity of DAPI signal for cell nucleus on the left and F-ACTIN + SOX2 double-staining on the right. The opening of a cavity within the inner compartment (nuclear SOX2 expression) is marked by F-ACTIN. n = 20 structures, 3 experiments. **E.** Immunostaining showing the formation of a central cavity, as marked by PODXL (magenta), within the inner compartment, marked by SOX2 (green), of an hEPSC-derived structure after 24 h culture in IVC. n = 20 structures, 3 experiments. DAPI staining is in blue. All scale bars in the figure indicate 20 um.

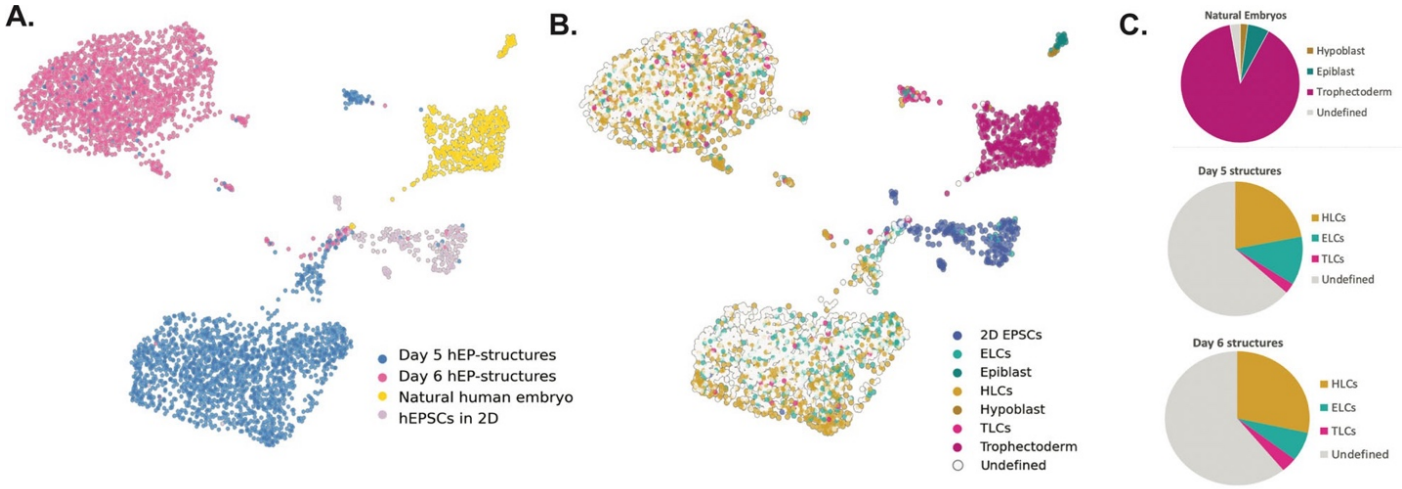
3.3.5. scRNA-seq analysis of hEP-structures

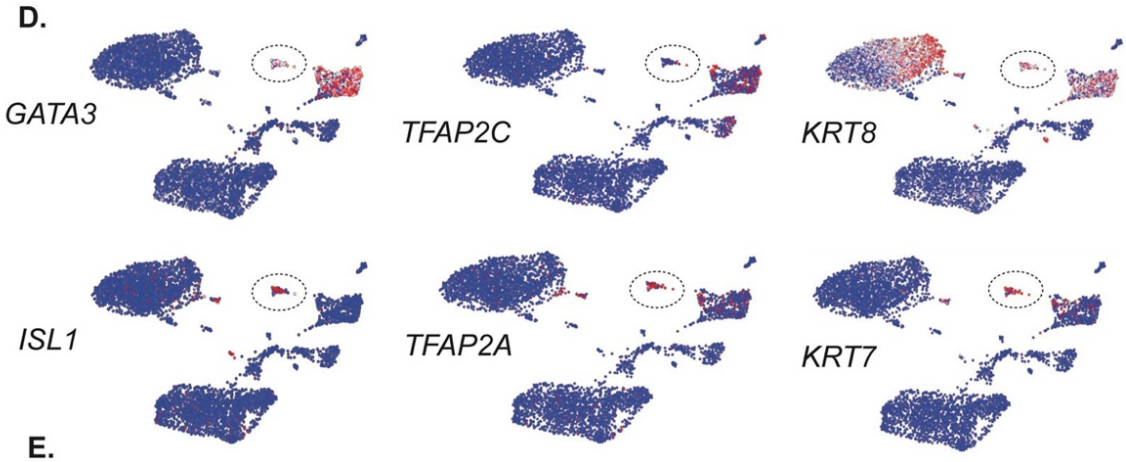
Finally, in order to further characterize the transcriptional programs of our hEP-structures, we performed scRNA-seq on hEPSCs grown in 2D (prior to 3D aggregation), hEP-structures at day 5, hEP-structures grown in IVC for 24 h, and natural human blastocysts at day 5/6. Cells clustered predominantly based on sample identity (Fig. 3.8A, B). We assigned a lineage score based on a list of well-defined lineage markers (see Supplementary Data 1 of Sozen and Jorgensen *et al.,*2021) and we observed signatures of EPI-Like Cells (ELCs), TE-Like Cells (TLCs) and HYPO-Like Cells (HLCs). However, these lineages fail to cluster in UMAP space. This analysis revealed that the hEP-structures were composed of a large portion of undefined cells (Fig. 3.8C), similar to two other 3D models of the human blastocyst[132,134]. In addition, we also found that there was an overrepresentation of HLCs compared to HYPO cells in the natural blastocyst (Fig. 3.8B, C). The hEPSCs grown in 2D showed high expression of pluripotency markers such as *TDGF1*, *NODAL*, and *POU5F1*, as expected (Fig. 3.9A) but there was considerable heterogeneity between cells. In day 5 hEP-structures we found a small subpopulation of *GATA3*-positive cells, which clustered close to the TE cluster of the natural embryo (Fig. 3.8D). This subcluster also expressed several amnion markers, such as *ISL1*, suggesting that it may share properties with the amnion rather than TE as has been reported in iBlastoids derived from reprogrammed hiPSCs[142] (Fig. 3.8D). Overall, while we do see expression of some key markers in our ELCs, HLCs and TLCs (Fig. 3.8E), there was a disproportionate representation of lineages in our hEP-structures as seen by the overrepresentation of HYPO-specific markers, with relatively little expression of TE-specific markers.

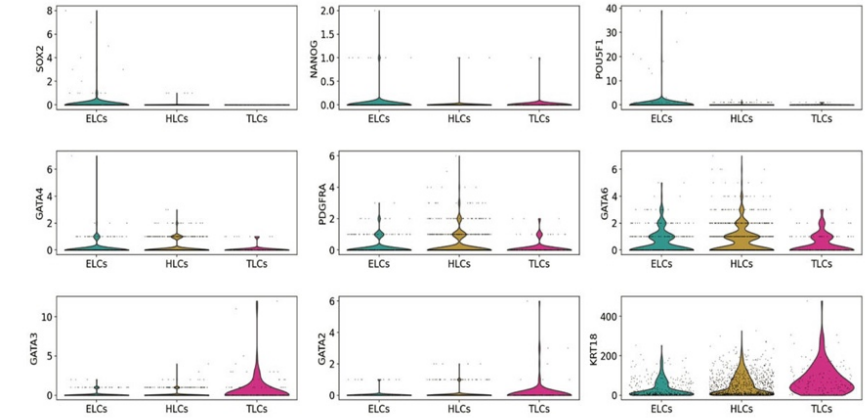
To better understand how the hEP-structures compared to natural human blastocysts at day 6, we compared the expression patterns of the key marker genes used to define the ELC, HLC and TLC signatures (Fig. 3.8F, 3.10A-C, see Supplementary Data 2 of Sozen and Jorgensen *et al.,*2021). This revealed that over half of the genes (52/96 at day 5, 50/96 at day 6) did not differ between ELCs and the human EPI, and a similar number of genes were expressed at comparable levels between the HLCs and HYPO (52/96 and 53/96). Interestingly, of those genes upregulated in HLCs compared to the HYPO, several were

extracellular matrix (ECM) proteins including *COL4A1, COL18A1*, and *FN1*. Importantly, these ECM genes were aberrantly expressed in ELCs and TLCs as well (Fig. 3.10D). In the TLCs, there was an observable shift between day 5 and day 6 in this signature where 43/96 genes were expressed at a similar level to the blastocyst TE at day 5, which decreased to 25/96 at day 6 (Fig. 3.8F, 3.10C, see Supplementary Data 2 of Sozen and Jorgensen *et al.*, 2021). In order to take a more global view of the differences between the hEP-structures and the natural blastocysts, we performed an unbiased comparative analysis of gene programs. This aims to identify clusters of cells which may upregulate blocks of gene programs. This revealed several gene programs upregulated in either the hEP-structures or natural blastocysts. These included ECM genes *COL3A1*, *COL4A2, COL4A1, COL1A1*, *COL6A3*, and *FN1*. Additionally, some genes purportedly related to the amnion—though also expressed in the TE—such as *POSTN* and *TPM1* were enriched in hEP-structures (Fig. 3.11A). These hEP-structures and natural blastocyst enriched genes were subjected to a simple gene set enrichment analysis using Reactome gene sets broadly related to signaling activities. This revealed an enrichment of several terms, including those related to PDGF, Interleukin, VEGF, PI3K, STAT3, and WNT signaling for the hEP-structures (Fig. 3.11B).

Taken together, we conclude that although the transcriptional machinery for blastocyst lineage programming appears to be initiated at some level in hEP-structures in our protocol, a continuum of cell fates then develops, indicating that complete trans-differentiation is not attained.


A.
Day 5 hEP-structures
Day 6 hEP-structures
Natural human embryo
hEPSCs in 2D
B.
2D EPSCs
ELCs
Epiblast
HLCs
Hypoblast
TLCs
Trophectoderm
Undefined
C.
Natural Embryos
Hypoblast
Epiblast
Trophectoderm
Undefined
Day 5 structures
HLCs
ELCs
TLCs
Undefined
Day 6 structures
HLCs
ELCs
TLCs
Undefined


D.
GATA3
TFAP2C
KRT8
ISL1
TFAP2A
KRT7


E.
SOX2
NANOG
POU5F1
GATA4
PDGFRA
GATA6
GATA3
GATA2
KRT18
ELCs
HLCs
TLCs

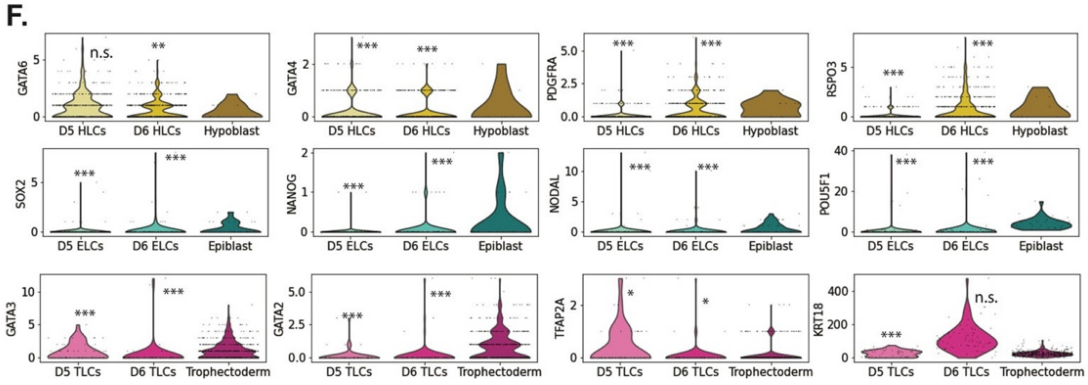
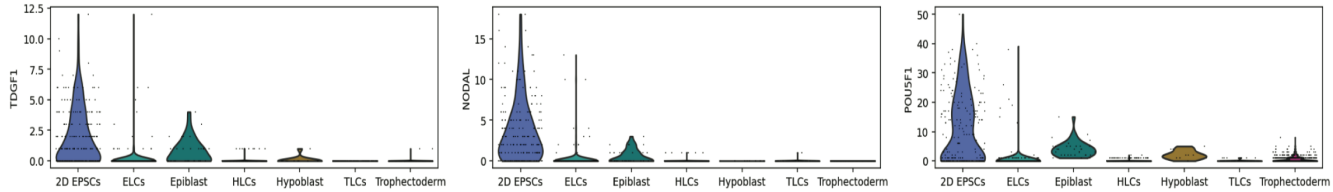
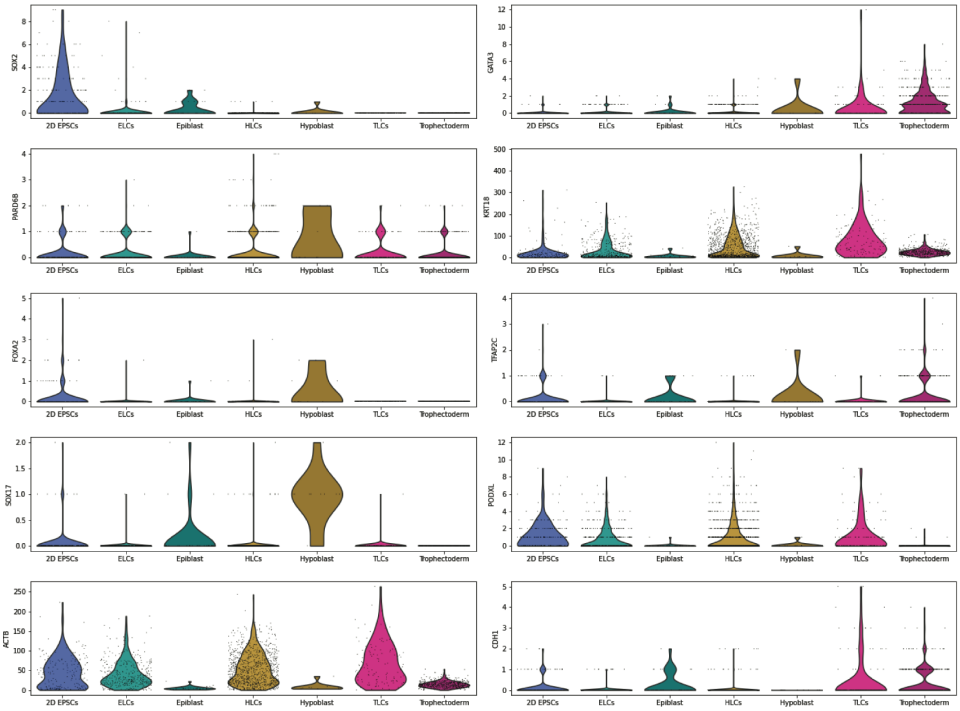

F.
GATA6
GATA4
PDGFRA
RSPO3
SOX2
NANOG
NODAL
POU5F1
GATA3
GATA2
TFAP2A
KRT18
n.s.
D5 HLCs
D6 HLCs
Hypoblast
D5 ELCs
D6 ELCs
Epiblast
D5 TLCs
D6 TLCs
Trophectoderm

Figure 3.8: scRNA-seq analysis of hEPSC-derived cystic structures.
A. Uniform manifold approximation and projection (UMAP) grouped by cell group: natural human embryos at D5/D6 (n = 542 cells), hEPSCs grown in 2D (n = 228 cells, 2 replicates), D5 cyst structures (n = 2013 cells, 3 replicates), and D6 cyst structures grown in IVC for 24 h (n = 2057 cells, 3 replicates). **B.** UMAP showing lineage scoring of all cell groups into EPI-like cells (ELCs, teal) and EPI (dark teal), HYPO-like cells (HLCs, yellow) and Hypoblast (dark yellow), TE-like cells (TLCs, magenta), and TE (dark magenta). Undefined cells appear in grey. **C.** Pie charts showing the distribution of lineage assignments for natural embryo, D5 structures, and D6 structures. **D.** Heat maps showing the expression of canonical TE markers enriched in natural cluster, as well as a subset of D5 TLCs (top row), and genes shown to be enriched exclusively in D5 TLC cluster (bottom row). A circle with a dotted line denotes D5 TLC cluster. **E.** Collection of violin plots showing the relative expression of certain key lineage markers in ELCs, HLCs, and TLCs. For EPI-related genes: SOX2, NANOG, POU5F1. For HYPO-related genes: GATA4, PDGFRA, and GATA6. For TE- related genes: GATA3, GATA2, and KRT18. **F.** Violin plots showing the expression of canonical markers for HYPO-related genes (yellow), EPI-related genes (teal) and TE-related genes (magenta) in Day 5 hEP-structures, D6 hEP-structures, and the natural embryo. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; (two-sided ad-hoc Dunn's multiple comparison test applied to an ANOVA), exact p-values are listed in Supp. Data 2). For D5 structures: HLCs, n = 445 cells; ELCs, n = 235 cells; TLCs, n = 50 cells. For D6 structures: HLCs, n = 586 cells; ELCs, n = 144 cells; TLCs, n = 73 cells. For natural blastocyst, HYPO, n = 11 cells; EPI, n = 32 cells; TE, n = 484 cells. All genes were taken from Liu *et al.* 2021, and are listed in Supp. Data 1.

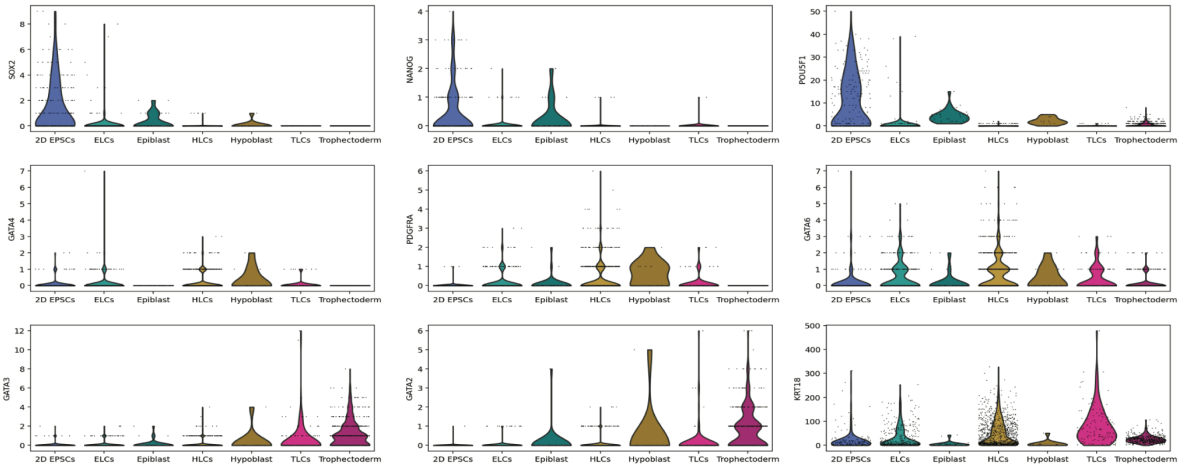
A. Pluripotency markers: Expression in EPSCs, Synthetic, Natural



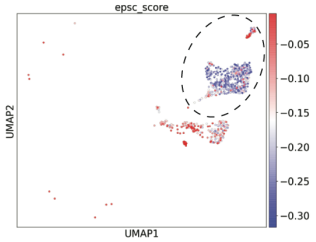
B. Antibodies: Expression in EPSCs, Synthetic, Natural



C. Genes from Fig. 5: Expression in EPSCs, Synthetic, Natural



D. EPSC score in 2D hEPSCs vs. Natural embryo



E. EPSC score in 2D hEPSCs, hEP-structures, and Natural embryo

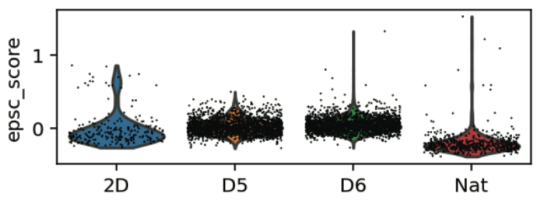
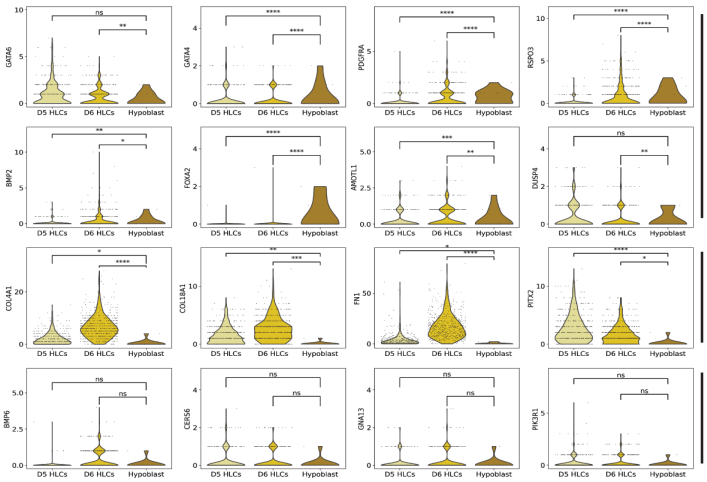


Figure 3.9: Characterization of hEPSCs in 2D culture with scRNA-seq analysis. A. Violin plots showing the expression of key pluripotency markers in 2D hEPSCs, lineages from hEP- derived structures, and lineages from natural embryos. **B.** Violin plots showing the expression of all genes corresponding to antibodies used throughout the main text in 2D hEPSCs, lineages from hEP- structures, and lineages from natural embryos. **C.** Violin plots showing the expression of genes used in Figure 5 in 2D hEPSCs, lineages from hEP- structures, and lineages from natural embryos. **D.** Heatmap of UMAP projection including cells from 2D hEPSCs and natural embryos showing relative "EPSC scores" (See Methods). Cells from natural embryo are surrounded by the circle with a grey dotted line. **E.** Violin plots comparing the "EPSC score" in 2D hEPSCs, D5 hEP-structures, D6 hEP- structures, and natural embryo. For 2D hEPSCs: n = 228 cells, 2 replicates. For natural embryos: n = 542 cells, 6 embryos.

A. Hypoblast genes: Natural vs. hEP-Structures

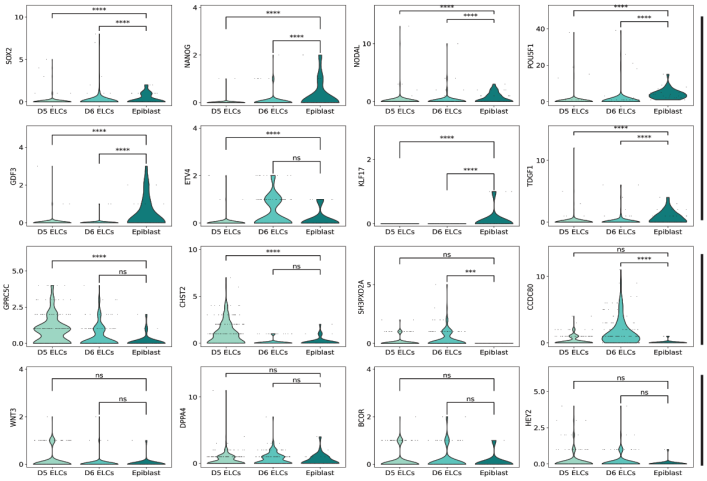


Downregulated
Day 5: 35/96
Day 6: 36/96

Upregulated
Day 5: 10/96
Day 6: 8/96

Not significant
Day 5: 52/96
Day 6: 53/96

B. Epiblast genes: Natural vs. hEP-Structures

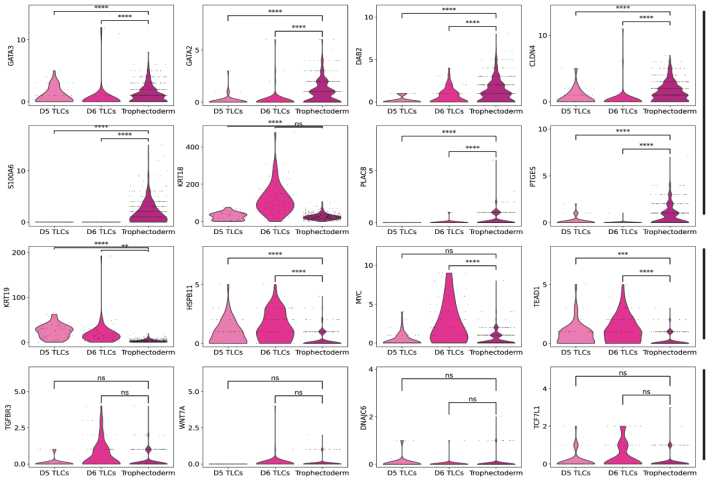


Downregulated
Day 5: 41/96
Day 6: 41/96

Upregulated
Day 5: 3/96
Day 6: 5/96

Not significant
Day 5: 52/96
Day 6: 50/96

C. Trophectoderm genes: Natural vs. hEP-Structures



Downregulated
Day 5: 48/96
Day 6: 59/96

Upregulated
Day 5: 9/96
Day 6: 16/96

Not significant
Day 5: 43/96
Day 6: 25/96

D. Extra-cellular matrix genes: Natural vs. hEP-Structures

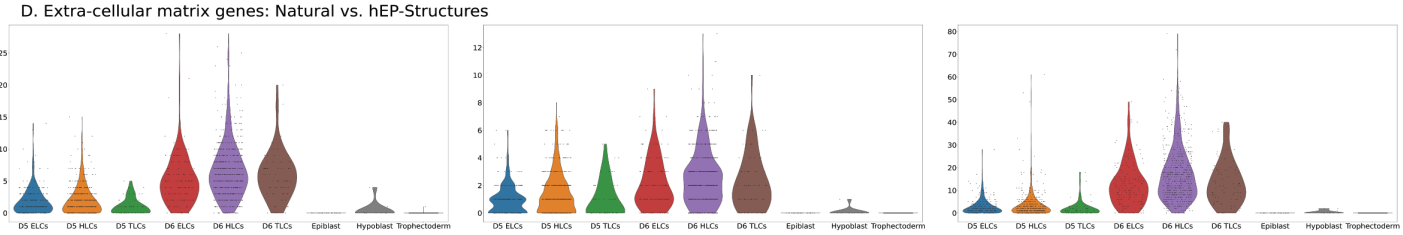


Figure 3.10. Characterization and statistical significance of assigned lineage scores compared to natural embryo. A. Violin plots showing the expression of HYPO-related genes in Day 5 HLCs, Day 6 HLCs, and natural blastocyst HYPO lineage. **B.** Violin plots showing the expression of EPI-related genes in Day 5 ELCs, Day 6 ELCs, and natural blastocyst EPI lineage. **C.** Violin plots showing the expression of TE-related genes in Day 5 ELCs, Day 6 ELCs, and natural blastocyst TE. **D.** Violin plots showing the expression of extra- cellular matrix genes in defined lineages of D5/D6 structures and lineages of natural embryos. Statistical analysis (two-sided ad-hoc Dunn's multiple comparison test applied to an ANOVA) was performed on 96 genes for each lineage, and the fraction of downregulated, upregulated, and not significant genes in comparison to the natural embryo are shown to the right of violin plots. (*$p < 0.05$, **$p < 0.01$, ***$p<0.001$; exact p-values are listed in Supp. Data 2). For Day 5 structures, n = 2013 cell, 3 replicates. For Day 6 structures, n = 2057 cell, 3 replicates. For natural blastocyst, n = 542 cell, 6 embryos.

Figure 3.11: Global analysis and comparison of hEP-structures, natural embryo, and previously published blastocyst-like models. A. Heat maps showing global gene expression patterns in Day 5 hEP-structures, Day 6 hEP-structures, natural blastocyst, iBlastoids (Liu *et al.*, 2021), StemBlastoids (Yu *et al.*, 2021), and 2D hEPSCs. An identical list of genes is denoted to the left of each heatmap, with a color block corresponding to related gene groups. Color bars below each heatmap denotes lineage assignments, as reported in their original publications. **B.** Gene set enrichment analysis based

on the hypergeometric test on Day 5 hEP-structures, Day 6 hEP- structures, and natural embryos was performed using Reactome- signaling gene sets to identify specific signaling pathways that are up-regulated in each.

3.3.6. Co-assembly of EPSCs and human trophoblast stem cells

The restricted specification of the TE-like lineage and potential questions regarding the differentiation of TE versus amnion fate in these structures generated solely from hEPSCs led us to question whether these cells possess the capacity to generate the TE-like lineage. We were recently able to correct a similar shortcoming in mouse blastoids by combining EPSCs with mouse trophoblast stem cells (TSCs)[88]. As a recent study has shown it is possible to derive human TSCs from first trimester human placental samples and blastocysts[91], we asked whether such TSCs could participate alongside hEPSCs in the assembly of human blastoids (Fig. 3.12A). To this end, we first allowed hEPSCs to aggregate for 24 h after which we added hTSCs[88]. The first 24 h of hEPSC aggregation demonstrated the increased potential to generate both embryonic EPI-like and extra-embryonic HYPO-like lineages in the EP condition compared to the Rset condition that has been shown to promote naïve pluripotency[101] (Fig. 3.12B). Within 4 days of co-culture, we observed the formation of cystic structures having internal acentric hEPSC compartments. We also observed robust expression of some of the blastocyst lineage markers, including GATA3, SOX2, and SOX17 (Fig. 3.12C). Nevertheless, it appeared that these structures struggled to form a cohesive TE-like epithelium and displayed not one but multiple cavities (Fig. 3.12C). Thus, in contrast to our experience with mouse TSCs, this human TSC line appears unable to rescue the shortcomings of human ESPCs to generate the TE lineage. Thus, future work will be required to generate human TSC lines having the ability to recapitulate pre-implantation development in order to enhance this system further.

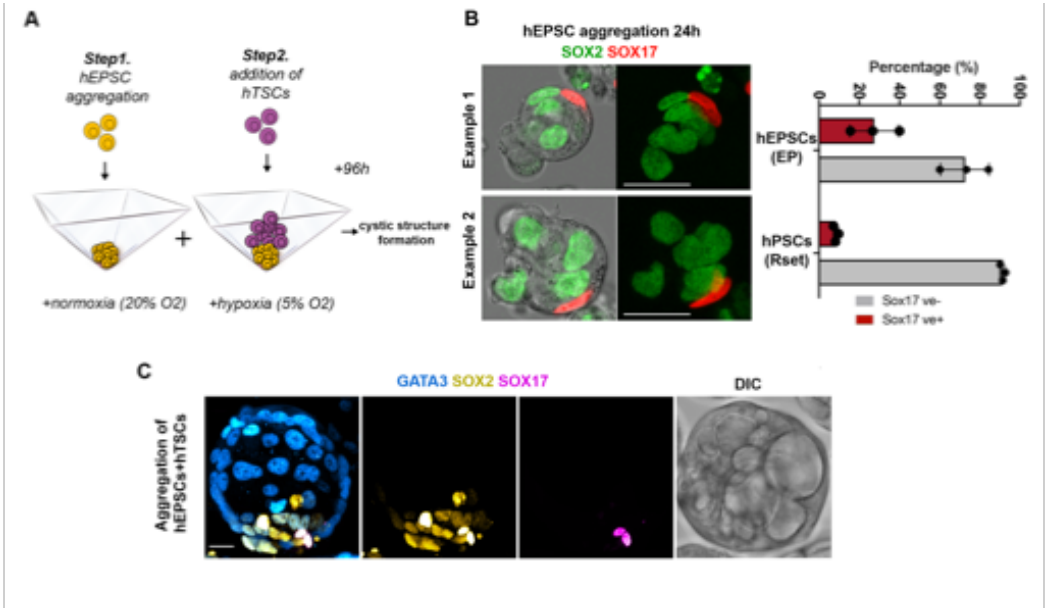


Figure 3.12: Co-assembly of hEPSCs and hTSCs. A. 2-step protocol for generating blastocyst-like structures from EPSCs combined with hTSCs. **B.** Aggregates co- expressing endogenous SOX2 (green) and SOX17 (red) after the first step (24h) of aggregation. SOX2 indicates naive pluripotent EPI-like cells; SOX17 indicates extra-embryonic HYPO-like cell formation. Quantification on the right shows structures scored as positive for hypoblast-like cell in conventional naïve Rset condition and EP condition. n = 3 experiments, 100 aggregates per group. Error bars show S.D. **C.** A representative structure generated from EPSCs combined with hTSCs stained for GATA3, SOX2 and SOX17. DIC image reveals the failure for cavitation. Representative of at least 10 independent experiments. All scale bars in the figure indicate 20 um.

3.4 Discussion

Cultured pluripotent stem cells show a dynamic spectrum of pluripotency states, reflecting stages in pre-to-post-implantation development in vivo, principally the transition from the naïve to the primed state[45,143,144]. This dynamic spectrum of pluripotency states exhibits distinct molecular and functional properties that affect the differentiation potential of cells and their ability to contribute to chimera formation[135,145]. Previous reports have demonstrated the ability of mouse and human EPSCs to contribute to chimeras[108] and show integration into both embryonic and extra-embryonic parts, a property not seen with other PSCs[135,146]. Thus, in the current study, we aimed to leverage this reported bi-potency of hEPSCs by developing a strategy enabling us to reconstitute the spatiotemporal lineage differentiation and self-organization of human early development. We assessed the multi-differentiation potential of hEPSCs to model early embryonic cell fate in a 3D culture platform. The results we present show that the hEPSC-based system we have established here allows hEPSCs to self-organize into structures that bear some resemblance to human blastocysts and peri-implantation stage embryos in morphology and, to a certain extent, also in lineage specification. However, we also show that their cell-lineage composition is imperfect, and these cells mainly adopt an intermediate transcriptional state.

Our knowledge of how to capture stem cell pluripotency in vitro is rapidly evolving and recent reports have challenged the previously proposed bi-potency of EPSCs[89,145]. Consistent with these reports, our previous study suggested that mouse EPSCs could not generate TE-like progeny in mouse blastoids, at least in our hands[88]. It has recently been reported that both mouse and human EPSCs are more similar to early post-implantation EPI cells and that the differentiation of pluripotent cells in either the naive or primed state may proceed toward different fate trajectories, suggesting that cells in different states might respond differently to stimuli[89,135,145]. Consistent with these observations, we find that hEPSCs mainly express genes that are specific to the late stage epiblast. We also observe that hEPSCs can adopt different morphologies under routine culture with some cells displaying features characteristic of pluripotent cells in the naïve state and others in the primed state, in different

ratios after each passage. This possibly reflects the range of developmental potencies associated with the hEPSCs we demonstrate here. However, while cells of the structures developing in our study display many EPI-like and HYPO-like markers and some TE-like markers, a full range of lineage markers is not expressed, and many of these markers are expressed at levels different from the *bone fide* blastocyst lineages. This is in line with recent findings emphasizing the restricted potency of mouse EPSCs[89]. Together, our results suggest important limitations in the molecular and epigenetic plasticity of hEPSC-derived structures leading to weak activation of important genes such as *GATA3*, *SOX2*, and *SOX17* resulting in inefficient lineage specification.

In our study to generate mouse blastoids, we had been able to correct some of the deficiencies of mouse EPSCs by co-culturing them with TSCs[88]. However, our attempts using a human TSC-line were not similarly encouraging. Given that these hTSCs are reported to be most similar to villous cytotrophoblasts[91], it is likely that these hTSCs may be more conducive to post- rather than pre-implantation development.

Recently alternative methods to generate human blastocyst-like structures have been described using naïve or induced pluripotent human stem cells[132,134,147]. In these reports, the resulting structures recapitulate the overall morphology of the blastocyst, with an inner cell mass and blastocoel cavity, similar to the structures presented here. Formation of structures with proper architecture, with a cavity and inner cell mass, ranges from 9.4 to 12.8% for human naïve blastocyst-like structures[134] and 5.8–18% for structures generated from iPSCs[132]. Additionally, the relationship between the efficiency for generation of proper structure morphology and correct segregation of all three lineages remains unclear and several gene programs do not appear to be shared between these models. While iPSC-derived blastoids seem to have the morphological and transcriptional organization, questions have been raised regarding the identity of their TE-like cells, as they appeared more similar to reported amnion-like cells[142]. This finding, in combination with our own, raises interesting disconnects between promising morphologies and cell behaviors in stem cell-derived models of human embryogenesis and transcriptional cell identities. This work also shows the ability

to generate morphologically similar structures with drastically different gene expression patterns, highlighting the uncoupling of morphology and gene expression in these models. These are yet to be explored and overcome in future work and emphasizes the need for stringent and comprehensive analyses to better understand the functionality of these models. Nevertheless, these recent studies, together with our own, will inform efforts to enhance the efficiency of both correct morphogenesis and robust lineage segregation in human embryo models.

In summary, our findings demonstrate that hEPSCs are not the equivalent of totipotent blastomeres and they are only partially able to specify embryonic cell progeny. This may reflect distinct molecular trajectories and an intermediate state adopted by these cells that lead to the generation of the improperly differentiated cells observed in this study. Nevertheless, these cells are able to generate multicellular structures showing some of the key morphological features and aspects of patterning similar to natural early human embryos. Thus, the system we present here may offer an alternative route with the potential to be harnessed into a fully functional embryo-like platform in vitro. We anticipate that, despite these shortcomings, this system together with others recently described, has the potential to lead to a variety of future applications that will be pivotal in unravelling many of the enigmas of human developmental regulation.

3.5 Materials, methods, and data availability

3.5.1 Methods

Ethics statement

Stem cell-derived multicellular structures described in this study show no evidence of germ line patterning, thus they do not have human organismal form or potential. Additionally, all experiments were terminated by no later than day 8 *in vitro*. Our research was subject to review and approval from the Human Embryo and Stem Cell (HESC) Committee of California Institute of Technology, in compliance with the ISSCR 2016 guidelines. The

human embryo work at California Institute of Technology was approved by the California Institute of Technology Committee for the Protection of Human Subjects (Institutional Review Board number 19-0948). Funding was obtained through Open Philanthropy Project fund at the Silicon Valley Community Foundation. Human embryos at the blastocyst stage were obtained from the University of Southern California (USC) through the preexisting USC Institutional Review Board-approved Biospecimen Repository for Reproductive Research (HS-15-00859) after appropriate approval was obtained unanimously from the Biorepository Ethics Committee.

Human embryo thawing

The human embryos at the blastocyst stage were warmed using Embryo Thaw Media Kit following the manufacturer's instructions (Fujifilm Irvine Scientific, Cat. No. 90124). The day before warming, Continuous Single Culture-NX Complete medium (Fujifilm Irvine Scientific, Catalog No: 90168) was equilibrated overnight at 37 °C + 5% CO2. On the day of warming (day 1), the straw that contains the embryo was defrosted at room temperature for 30 s and immersed in prewarmed (37 °C) water for 1 min until ice melted. The embryo was then transferred into T-1 (5 min), T-2 (5 min), T-3 (10 min) solutions for slow warming and finally into Multipurpose Handling Medium (MHM, Fujifilm Irvine Scientific, Cat. No. 90163) for recovery. All these incubation steps were done using 4 well plates (Nunc) and 1 ml per solution. Warmed embryos were finally incubated in drops of preequilibrated Continuous Single Culture-NX Complete medium under mineral oil (9305, Irvine Scientific). Culture conditions are the following: 37 °C 21% O2 and 5% CO2. Embryos were incubated for a total of 24 h until used for further RNA-sequencing protocols.

Human cell lines

The hPSC lines utilized in this study include: RUES2-GLR (kindly provided by Ali Brivanlou, The Rockefeller University, US), and ESI017 (kindly provided by Michael Elowitz, California Institute of Technology, US). Human TSCs (TS^{CT}) were kindly provided by Hiroaki Okae and Takahiro Arima (Tohoku University Graduate School of Medicine,

Japan). Each of these cell lines was tested negative for mycoplasma contamination, which was monitored on a bi-monthly basis (MycoScope™ PCR Mycoplasma Detection Kit, Genlantis).

Cell culture

All hEPSC lines were maintained under 20% O_2 and 5% CO_2 at 37 °C conditions on irradiated CF1 mouse embryonic fibroblasts (MEF) feeder cells. hEPSCs were grown using 'human Expanded Potential' (hEP) medium consisting of DMEM/F12 (Thermo Fisher Scientific, 11320-033), Neurobasal-A (Thermo Fisher Scientific, 21103-049), N2 supplement (Thermo Fisher Scientific, 17502-048), B27 supplement (Thermo Fisher Scientific, 12587-010), 1% GlutaMAX (Thermo Fisher Scientific, 35050-061), 1% nonessential amino acids (Thermo Fisher Scientific, 11140-050), 0.1 mM b-mercaptoethanol (Thermo Fisher Scientific, 31350-010), penicillin-streptomycin (Thermo FisherScientific, 15140–122) and 5% knockout serum replacement (KSR, Thermo Fisher Scientific, A3181502). LCDMYI supplementation was added as indicated at the following concentrations: 10 ng ml^{-1} recombinant human LIF (L, 10 ng ml^{-1}; Peprotech, 300-05), CHIR99021 (C, 1 mM; Stem Cell Technologies), (S)-(+)-Dimethindenemaleate (D, 1 mM; Tocris, 1425) and Minocycline hydrochloride (M, 2 mM; Santa Cruz Biotechnology, sc-203339), 1. All hEPSCs were used before reaching P70 and cell cultures were examined by eye to monitor for spontaneous differentiation of colonies into mesenchymal-like cells.

hTSCs were cultured on 6-well plates pre-coated with 5 mg/ml Col IV at 37 C for at least one hour, as previously described in Okae et al.. Cells were grown in 'human Trophoblast stem cell' (hTS) medium consisting of DMEM/F12 supplemented with 0.1 mM b-mercaptoethanol, 0.2% FBS, 0.5% Penicillin-Streptomycin, 0.3% BSA (A8806-5G, Sigma-Aldrich), 1% ITS-X supplement (51500-056, Thermo Fisher Scientific), 1.5 mg/ml L-ascorbic acid (A4403, Sigma-Aldrich), 50 ng/ml EGF (62253-63-8, Sigma-Aldrich), 2 mM CHIR99021, 0.5 mM, A83-01 (72024, Stemcell Technologies), 1 mM SB431542 (Stem Cell Technologies), 0.8 mM VPA (Sigma-Aldrich) and 5 mM Y27632

Preparing and plating cell suspensions for "AggreWell" aggregation experiments

AggreWell 400 format plates were prepared following the manufacturer's protocol. Briefly, wells were rinsed with the rinsing solution (Stem Cell Technologies), centrifuged for 5 min at 2000g and incubated at room temperature in the tissue culture hood for 20 min. The wells were then washed with 2 ml of 1× PBS. After PBS removal, 500 ml of final culture medium (*IVF-hEP-hTS*, see below) was added to each well and the plate placed at 37 C and 5% CO_2 until ready to use.

Generation of multicellular aggregates in 3D

To begin, hEPSCs were dissociated to single cells by incubation with Accutase (07920, Stem Cell Technologies) at 37 °C for 3 min. Cells were collected and pelleted by centrifugation for 4 min at 300 g and resuspended in hEP-LCDMYI medium (described above). This cell suspension was pre-incubated at 37 °C in an atmosphere of 5% CO_2 on gelatinized tissue-culture-grade plates for 30 min to remove inactive MEFs.

Post incubation on gelatin plates, cells were counted using a haemocytometer and a total of 7200 hEPSCs was added to 1 mL of media composed of 50% *IVF* media (Continuous Single Culture-NX Complete (CSCM-NXC)) (90168, FUJIFILM), 25% *hEP* media, and 25% *hTS* media. This media was also supplemented with CHIR99021 (2uM), Y27632 (5uM), BMP4 (20 ng/mL), FGF2 (40 ng/mL), and A83-01 (2uM). Cell suspensions were added dropwise to the Aggrewells. All wells without cells were filled with 1 mL PBS to humidify the local atmosphere to minimize evaporation. The AggreWell plate (24-well, 1200 Aggrewell format) was then centrifuged for 3 min at 100 g, and placed at 37 °C under hypoxic conditions (5% CO_2 and 5% O_2). After 48 h, media was removed from wells and replaced with fresh culture media as described above, although FGF2 concentration was lowered to 20 ng/mL and A83-01 was omitted. Cells were left to grow for an additional 48-72 h until proper morphology was observed, at which point structures were fixed for immunostaining or transferred to IVC media (see below) for mimicking development beyond implantation stages.

Criteria for selecting multicellular aggregates structures

Following completion of any given aggregation experiment (from day 4 to 6), all cystic structures those clearly displaying a cavity were included in further analyses. Non-cavitated structures were excluded from downstream analyses.

Co-culture of hEPSCs with hTSCs

To perform two-step aggregation experiments, hEPSCs were first seeded as described above hEP-LCDMYI media. After a 24 h period of aggregation, hTSC colonies were dissociated to single cells, and counted using a haemocytometer. For aggregation experiments, 50% *hEP* media and 50% *hTS* media is used as described above. A total of 16,800 hTSCs were added per well (24-well, 1200 Aggrewell format) and the plate was placed at 37 °C, 5% CO_2 and 5% O_2.

In Vitro Culture (IVC) of hEPSC-derived structures

To prepare plate for in vitro culture, 150 µL of modified *IVC1* (*mIVC1*) media was added to each well of a 96-well ultra-low attachment U-shaped plate (7007, Costar). *mIVC1* media consisted of the following: Advanced DMEM/F12 (12634-010; Thermo Fischer Scientific; Waltham, US) supplemented with 20% (vol/vol) heat-inactivated FBS (16141079, Thermo Fisher Scientific), 2 mM GlutaMAX, penicillin (25 units/ml)/Streptomycin (25 µg/ml), 1X ITS-X (10 mg/L insulin, 5.5 mg/L transferrin, 0.0067 mg/L sodium selenite, 2 mg/L etholamine; 51500-056; Thermo Fisher Scientific; Waltham, US), 8 nM β-estradiol (E8875; Sigma-Aldrich; St. Louis, US), 200 ng/ml progesterone (P0130; Sigma-Aldrich; St. Louis, US), 25 µM *N*-acetyl-L-cysteine (A7250; Sigma-Aldrich; St. Louis, US), 17 nm IGF1, 20 ng/mL FGF2 (Gibco), FGF4 (25 ng/mL; R&D Systems, 5846-F4) and heparin (1 mg ml^{-1}; Sigma, H3149).

Immunofluorescence staining

Stem cell-derived structures were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, 15710) for 20 min at room temperature, and then washed twice in PBT [phosphate-buffered saline (PBS) plus 0.05% Tween-20]. Structures were permeabilized for 30 min at room temperature in PBS containing 0.3% Triton-X-100 and 0.1% glycine. Primary antibody incubation was performed overnight at 4 °C in blocking buffer [PBS containing 10% fetal bovine serum (FBS), 1% Tween-20]. The following day, embryos were washed twice in PBT, then incubated overnight at 4 °C with secondary antibody (1:500) in blocking buffer. Structures were washed twice in PBT buffer and then transferred to PBT drops in oil-filled optical plates before confocal imaging. The antibodies used are given in Supplementary Table 1.

For human embryo images shown in Fig. 2e, embryos were fixed in IVIRMA Valencia, washed twice in a PBS solution containing 0.1% Tween-20 (Sigma, cat. no. P9416) and immediately placed into a 0.5 ml PCR tube within an oil-PBS-oil interphase. Tubes were stored at 4 °C were shipped to the University of Cambridge for immunofluorescence.

Image data acquisition, processing, and quantification

Fluorescence images were acquired with an inverted Leica SP8 confocal microscope (Leica Microsystems), using a Leica Fluotar VISIR 0.95 NA 25x objective. Fluorophores were excited with a 405-nm diode laser (DAPI), a 488-nm argon laser (GFP), a 543-nm HeNe laser (Alexa Fluor-543/555) and a 633-nm HeNe laser (Alexa Fluor-633/647). Images were acquired with 0.5–1.2 mm z-separation. Raw data were processed using open-source image analysis software Fiji Image J (version: 2.0.1) open access software and assembled in Photoshop CC 2019 (Adobe). Digital quantifications and immunofluorescence signal intensity graphs were obtained using Fiji software47.

Apical enrichment analysis: F-actin and PARD6 polarization were measured in a single focal plane, by taking the middle plane of the aggregate. A freehand line of the width of 0.5μm

was drawn along the cell-contact free surface (apical domain), or cell-contact (basal) area of the cell, signal intensity was obtained via the Region of Interest (ROI) function of Fiji. The apical/basal signal intensity ratio is calculated as: I(apical)/I(basal). A cell is defined as polarized when the ratio between the apical membrane and the cytoplasm signal intensity exceeds 1.5.

GATA3 expression analysis: the nucleus of each cells is masked using the Region of Interest (ROI) tool of Fiji. The average signal intensity of the ROI is calculated and a cell is defined as GATA3 positive when the nucleus to cytoplasm signal intensity exceeds 1.5.

siRNA-mediated knock-down in hEPSC-derived aggregates

Transfections of siRNA were performed using Lipofectamine RNAi MAX (13778075, Thermo Fisher Scientific) according to the manufacturer's instructions. Upon seeding hEPSCs into AggreWells (as described above), Lipofectamine and siRNA (Qiagen, Hs_PLCB1_4, SI00115521; Qiagen, Hs_PLCB1_6, SI02781184; Qiagen, negative control siRNA, 1022076) against target genes with Opti-MEM (31985070, Gibco) is mixed and the mixture of either control siRNA or PLCB1 siRNA were evenly added into each well. Cell aggregates at 48 h were collected to analyse the gene expression by qRT-PCR.

Bulk qRT-PCR analysis

Total RNA was extracted with using Arcturus PicoPure™ RNA Isolation Kit (12204-01, Applied Biosystems) as per manufacturer's instructions. QRT-PCR was performed with the Power SYBR Green RNA-to-CT 1-Step Kit (Life Technologies) and a Step One Plus Real-time PCR machine (Applied Biosystems). The amounts of mRNA were measured with SYBR Green PCR Master Mix (Ambion). Relative levels of transcript expression were assessed by the ΔΔCt method, with Gapdh as an endogenous control. For qPCR primers used, see Supplementary Table 2.

Single cell isolation of in vitro *cultured human embryos and hEPSC-derived aggregates*

Human blastocyst ($n = 6$) were exposed to Tryple Express Select ×10 (ThermoFisher A1217701) for 15 min in 37 °C, and subsequently dissected with glass capillaries of different diameters.

hEPSC-derived structures were cultured until Day 5 (see Generation of multicellular aggregates in 3D) or Day 6 (see In Vitro Culture of hEP-structures beyond implantation). We then selected structures based on the morphological criteria of having a cavity and acentric compact inner cell mass with glass capillaries. Roughly 50 structures were collected for Day 5 and Day 6 and each condition was performed in triplicate, for a total of ~150 structures per condition. hEP-structures were then first exposed to Tryple Express Select ×10 (ThermoFisher A1217701) for 15 mins in 37 °C, and subsequently dissected with glass capillaries of different diameters.

Single cell isolation of hEPSCs in 2D culture

hEPSCs were washed once with PBS and dissociated with Accutase (07920, Stem Cell Technologies) at 37 °C for 3 min. Cells were collected and pelleted by centrifugation for 4 min at 300 g and resuspended in hEP-LCDMYI medium (described above). This cell suspension was pre-incubated at 37 °C in an atmosphere of 5% CO_2 on gelatinized tissue-culture-grade plates for 30 min to remove inactive MEFs.

Single-cell mRNA-sequencing

For single-cell sequencing, we used lipid-modified oligonucleotides (LMOs) to multiplex multiple samples into a single droplet microfluidics run. Dissociated cells from Day 5 hEP-derived structures, Day 6 hEP-derived structures, and 2D hEPSCs were labelled with sample-specific lipid-modified oligos (LMOs). Samples were washed to remove any leftover LMOs and then pooled into a multiplexed cell suspension that was run on a single lane of a 10× Genomics chip, using v3.0 reagents. Cells from the natural embryo were not multiplexed but were run concurrently on a parallel lane within the chip. The single-cell sequencing library was prepared as per the manufacturer's instructions and sequenced on an Illumina Hiseq

4000 at a minimum coverage of 20,000 PE reads per cell (read 1: 28 bp, i7: 8 bp, read 2: 91 bp). Sample-specific LMO tags were separately amplified according to Ding et al. and sequenced at a read depth of 2,016 reads per cell. Sample demultiplexing was performed using an in-house demultiplexing pipeline that discovers sample-specific thresholds, scores each cell as being positive or negative for a specific tag, and retains only singly-labelled cells.

Single cell RNA-seq data analysis

Single-cell RNA-sequencing was performed using the 10× Genomics Chromium system. Reads were aligned against GRCh38. Further downstream analyses were performed in Python using the Scanpy toolkit (version 1.7.2) and Anndata (version 0.7.5). No cells were filtered for mitochondrial or ribosomal content. Initial analysis including normalization, scaling, identification of highly variable genes, and clustering was performed as described in the kalisto | bustools tutorial "Introduction to single-cell RNA-seq II: getting started with analysis". Single-cell data was further visualized using the UMAP dimensionality reduction, as determined by the sc.tl.umap function in Scanpy. Lineages were defined using sc.tl.score_genes function in scanpy and the gene list used for this function were taken from Supplementary Table 12 of Liu et al. (type: ALL-TE, ALL-EPI, ALL-PE), and are included as Supp. Data 1. A score for each lineage was given to every cell, and the "lineage" designation for each cell was determined by the highest score of the 3. Any cells with scores below 0.08 were denoted as "undefined". All violin plots were made using sc.pl.violin function in Scanpy.

Unsupervised clustering of gene expression programs

To compare global gene expression patterns across datasets, we constructed large-scale heatmaps for each dataset clustered using unsupervised methods (orthogonal non-negative matrix factorization for gene programs and hierarchical clustering for cells). First, we integrated multiple datasets together by keeping only the intersection of genes found across all datasets (n = 18379 genes). We then used the PopAlign framework to filter highly variable genes, and normalize the datasets. Normalization was done by dividing each transcript count

by the sum total within the cell, scaling by 10,000, adding a + 1 pseudocount, and then logging. To find gene programs, we randomly sampled 5,000 cells across the integrated dataset, and then ran orthogonal non-negative matrix factorization find a set of 16 feature vectors. We discovered the top genes within each feature, and used this gene list to organize the gene expression heatmaps for each dataset (popalign.plot_top_gene_features). Cells within each dataset were clustered using scipy hierarchical clustering (scipy.cluster.hierarchy) using correlation distance, and linkage = complete. Gene programs were then manually reorganized into groups which were Universal across all datasets, or more specifically enriched in the Natural Embryo, D5, or D6 stem cell-derived embryo-like structures. We then ran a gene set enrichment analysis based on the hypergeometric test (popalign.enrichment_analysis) on these gene groups using Reactome-signalling gene sets to identify specific signaling pathways that are up-regulated.

Statistics and reproducibility

Statistical tests were performed on GraphPad Prism 8.4.3 software. Figure legends indicate the statistical tests used and number of independent experiments performed in each analysis. All error bars defined in the legends. Unless otherwise noted, each experiment was performed at least two times. Statistical significance: $p < 0.05$ was considered statistically significant (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

3.5.2 Data availability

The scRNA-seq data for 2D hEPSCs, hEP-structures, and natural human blastocyst generated in this study have been deposited in the GEO database under accession code GSE178326 [GEO]. Published iBlastoid and StemBlastoid datasets used in this study were obtained from Liu *et al.* and Yu *et al.* under accession numbers GSE156596 and GSE150578, respectively. Source data are provided with this paper.

Code availability

The code generated in this study is provided at https://github.com/vjorgensen/hEP-structures_MZG.

Table 3.1: Antibodies for immunostaining in hEP-structures

Targeted Protein	Species	Dilution	Catalog No.	Vendor
PARD6	Rabbit	1:200	sc-166405	Santa Cruz Biotechnology
E-Cadherin (Clone 36)	Mouse	1:200	610182	BD Biosciences
OCT3/4	Rabbit	1:200	sc-9081	Santa Cruz Biotechnology
SOX2	Rabbit	1:200	D6D9	Cell Signaling Technologies
GATA3	Goat	1:200	AF2605	R&D Systems
PODXL	Mouse	1:20	MAB1658	R&D Systems
SOX17	Goat	1:200	AF1924	R&D Systems
FOXA2	Rabbit	1:200	D56D6	Cell Signaling Technologies
GATA6	Goat	1:200	AF1700	R&D Systems
Alexa Fluor® 488 Phalloidin (F-ACTIN)	N/A	1:500	A12379	Thermo Fisher Scientific
TFAP2C	Goat	1:200	AF5059	R&D Systems
KRT18	Mouse	1:200	Ab668	Abcam
Anti-GFP	Rat	1:1000	04404-84	Nacalai Tesque

Table 3.2: Primers for RT-qPCR in hEP-structures

Gene	FW Primer	RV Primer
GATA3	AAGGCATCCAGACCAGAAACCG	AGCATCGAGCAGGGCTCTAACC
PLAC8	GTTTCACCATCTTGGTCAGG	CTGTAATTCCAGCACCTTGG
CDX2	CCGAACAGGGACTTGTTTAGAG	CTCTGGCTTGGATGTTACACAG
KRT8	ACCCTCAACAACAAGTTTGCCTCC	TCCACTTGGTCTCCAGCATCTTGT
KRT18	ACACAGTCTGCTGAGGTTGGAG	TGCTCCATCTGTAGGGCGTAG
NANOG	ACACTGGCTGAATCCTTCCTCTCC	CGCTGATTAGGCTCCAACCATACTC
POU51F	TCTCGCCCCCTCCAGGT	GCCCCACTCCAACCTGG
KLF4	CCCACATGAAGCGACTTCCC	CAGGTCCAGGAGATCGTTGAA
SOX17	GAGCCAAGGGCGAGTCCCGTA	CCTTCCACGACTTGCCCAGCAT
PDGFRA	CTCCCTGGCTGTTCTGATCG	TGCCAACCCTGTTCCAAAGT
PLCB1	GGAAGCGGCAAAAAGAAGCTC	CGTCGTCGTCACTTTCCGT

Chapter 4

GATA4-INDUCBILE BLASTOIDS DEMONSTRATE POST-IMPLANTATION REMODELING *IN VITRO*

4.1. Abstract

4.2. Introduction

Mammalian embryogenesis commences upon fertilization of the oocyte to form the totipotent zygote, a single cell that has the potential to give rise to all embryonic and extra-embryonic lineages. Upon fusion of egg and sperm, the zygote will undergo rapid cleavage divisions to increase cell number without altering the total volume of the embryo. By embryonic day 3.5 (E3.5), the first lineage segregation events have occurred to form the cavitated structure known as the blastocyst. The blastocyst is initially composed of two distinct cell types: the bipotent inner cell mass (ICM) and the outer extra-embryonic trophectoderm (TE)[26,71]. A day later at E4.5, the ICM will undergo a second lineage segregation event to form the extra-embryonic primitive endoderm (PE) and the embryonic epiblast (EPI)[29].

At E4.75, all three lineages have been specified and the embryo is ready to implant. The outer trophectoderm will diverge into two major subtypes: the polar TE, which will give rise to the extra-embryonic ectoderm (ExE), and eventually form the placenta and chorion; and the mural trophectoderm, which is responsible for invasion into the maternal uterine tissue[39]. The PE lineage will also bifurcate into two subsequent lineages: the parietal endoderm (PaE), which will migrate along the mural trophectoderm and facilitate formation of the Reichert's membrane[148,149]; and the visceral endoderm, which will encompass the ExE and EPI compartments of the post-implantation egg-cylinder.

In recent years, several attempts have been made to form "blastoids", blastocyst-like structures generated solely from stem cells[86–88,150]. The first model developed by Rivron *et al*. in 2018 combined embryonic stem cells (ESCs) and trophoblast stem cells (TSCs) to

generate structures mimicking E3.5 blastocysts. While these blastoids nicely recapitulated the early blastocyst in morphology and lineage segregation of the TE and ICM, they failed to robustly develop the primitive endoderm lineage. To expand upon this model, systems using mouse expanded potential stem cells (EPSCs) alone[87] or in conjunction with mouse TSCs[88] have also been developed. EPSCs were believed to have extended pluripotency potential and thus have the ability to diverge into embryonic and extra-embryonic lineages. Nevertheless, while these models show increased formation of the PE lineage, further development beyond implantation is limited both *in vitro* and *in vivo.* Sozen *et al.* demonstrate that their EPS-blastoids can transition to post-implantation egg-cylinder morphology with proper allocation of the EPI, ExE, and VE lineages *in vitro,* however efficiency remains low. Additionally, development beyond ~E5.5 does not occur.

Several post-implantation stem cell-derived embryo models also exist[4,80–85]. These models can perform many events characteristic of post-implantation development including anterior-posterior axis patterning[80,81], primitive streak formation[151], gastrulation[81,83–85,151], and even organogenesis and neurulation[83–85]. Development beyond these stages, however, remains impossible as current methods for *ex utero* culture, even for natural embryos, is capped at E8.5[5]; this is likely due to insufficient nutrients being supplied to the embryo, as well as the absence of mechanical and chemical cues from the uterine tissues. Moreover, these structures must be cultured *in vitro,* as they do not maintain necessary tissue types required for implantation into the uterus of the mother. Blastoids, however, have the theoretical potential to be transplanted into host mice and develop further *in vivo*. Such experiments would not only expose blastoids to the natural signaling environment of the uterus, but would also give the blastoids access to the maternal nutrient supply allowing them to develop further. Transplantation of blastoids into pseudopregnant mice has been attempted; in each instance, however, blastoids are able to induce decidual formation, yet fail to develop any further[86–88,150].

A potential explanation limiting development of blastoids either *in vitro* or *in vivo* is due to insufficient formation of the PE lineage. Although ESCs are derived from the ICM, it has

been shown that they do not maintain the ability to give rise to extra-embryonic lineages[76]. Additionally, recent reports analyzing the potency of EPSCs demonstrate that they do not recapitulate the pre-implantation morula, but are rather more reminiscent of the post-implantation E5.5 epiblast[89]; such cells would also be limited in their differentiation potential and thus fail to give rise to a bona fide PE. It has been shown that independent overexpression of two transcription factors, Gata4 and Gata6, in ESCs is sufficient to drive cells towards the endoderm fate[79]. Moreover, these Gata4-inducible ESCs have been utilized for the formation of post-implantation embryoids and demonstrate the ability to faithfully differentiate into VE, a PE derivative[83–85,151]. Nevertheless, the question remains as to whether or not these cells can replicate pre-implantation development and the earlier stage PE.

To address this, we have aimed to establish an improved protocol for blastocyst-like structures that combines three distinct cell types: 1) TSCs, to generate the TE-like lineage; 2) ESCs, to generate the EPI-like lineage; and 3) tetO-Gata4 ESCs which transiently express *Gata4* in response to doxycycline (dox), to generate the PE-like lineage. Here we show that Gata4-inducible blastoids, hereafter termed G4i-balstoids, robustly mimic development of E4.5 blastocysts in size, shape, and lineage allocation. Moreover, G4i-blastoids demonstrate the ability to transition to post-implantation stages *in vitro*, as evidenced by an egg-cylinder morphology and indication of axis patterning up until primitive streak formation.

4.3. Results

4.3.1. Self-assembly of mouse ESCs, TSCs, and Gata4-inducible ESCs to form blastocyst-like structures

Studying the natural mouse embryo has greatly elucidated several aspects of early mammalian development and represents a powerful system for better understanding early embryogenesis. Unfortunately, ethical and technical restrictions limit the utility of such research. The emergence of robust stem cell-derived embryo models could greatly alleviate these constraints, but first requires accurate recapitulation of both embryonic and extra-embryonic lineages. As neither ESCs or EPSCs robustly capture the PE fate, we

hypothesized that tetO-Gata4 ESCs (G4i-ESCs) would enhance formation of this lineage and thus increase the developmental potential of blastocyst-like structures. To answer this, we combine together three distinct cell types, ESCs, G4i-ESCs and TSCs, in inverted pyramid-shaped microwells via a two-step seeding process and defined media changes (Fig. 4A). We first combine ESCs and G4i-ESCs in minimal media, termed FCLY, which consists of the following: standard feeder cell media (FC), the cytokine leukemia inhibitory factor (LIF) to maintain pluripotency, and the ROCK inhibitor Y27632 to promote self-renewal. After approximately 6-8 hours, the two ESC populations have coalesced together to form a single aggregate reminiscent of the compacted ICM (Fig. 4.1A, B). Once this ICM-like aggregate has formed, TSCs are seeded into the microwells and media is changed to TX medium as previously described in earlier works[86,88,152]. Upon addition of TSCs, we also switch from high-oxygen normoxia (20% O_2) to low-oxygen hypoxia (5% O_2) conditions; this is because embryos naturally develop in low-oxygen conditions, and hypoxia-inducible factors have been implicated in maintaining stemness and facilitating proper development of the placenta[153]. Previous models have shown that hypoxic conditions are critical for enhanced cavitation of these structures, yet mechanistic explanation for this phenomenon remains unexplored[88]. Within 24 h, the TSCs have fully encompassed the ICM-like ESC compartment forming the outer trophectoderm-like lineage (Fig. 4.1A, B). Once this has occurred, we introduce doxycycline (dox) to the media in order to selectively induce *Gata4* expression in our G4i-ESC population (Fig 4.1A); in as little as 12 h we see "salt and pepper" expression of Gata4 exclusively in the ESC compartment (Fig 4.1B). After 72 h in TX media, cells self-organize to form cavitated blastocyst-like structures that mimic the embryo in size, shape, and morphology (Fig. 4.1C, D, G), although we do see higher variation of these features in our stem cell-derived structures than in natural embryos. Additionally, immunostaining of the key lineage markers Oct4, Cdx2, and Gata4 demonstrates initiation of EPI-, TE-, and PE-like lineages, respectively. Not only are these markers present, we also saw correct localization of these populations as we would expect in the natural embryo with EPI- and PE-like cells forming an inner compartment surrounded by an outer TE-like layer (Fig. 4.1D). Furthermore, Gata4-positive cells, denoting the PE-like lineage, demonstrate successful sorting away from EPI-like cells to line the inner compartment along the

blastocoel cavity, a critical step needed for later stage development (Fig. 4.1D). Our model also shows higher frequency of PE-like lineage formation in comparison to previous models, with ~82% of cavitated structures containing this lineage in comparison to ~62% and ~10% frequency in EPSC-blastoids and ESC-blastoids, respectively (Fig. 4.1F). In certain instances, our stem cell-derived structures contain a larger number of cells per lineage than is expected in the natural embryo at this stage. Nevertheless, the overall lineage ratios are comparable, albeit with higher variation (Fig. 4.1E). This is important as it has been shown that increased size (i.e. cell number) of natural embryos does not hinder further development. Size has, however, been shown to affect the mechanism by which the pro-amniotic cavity will form[154], an important caveat to consider when assessing later stage development of our structures. Lastly, we validated that TSCs grown in hypoxia greatly enhanced cavity formation within our structures, demonstrating 80% efficiency compared to only 5% efficiency for TSCs grown in normoxia (Fig. 4.2A, B). Given the morphological and lineal similarities to natural embryos, we will refer to these structures as Gata4-inducible blastoids (G4i-blastoids) hence forward.

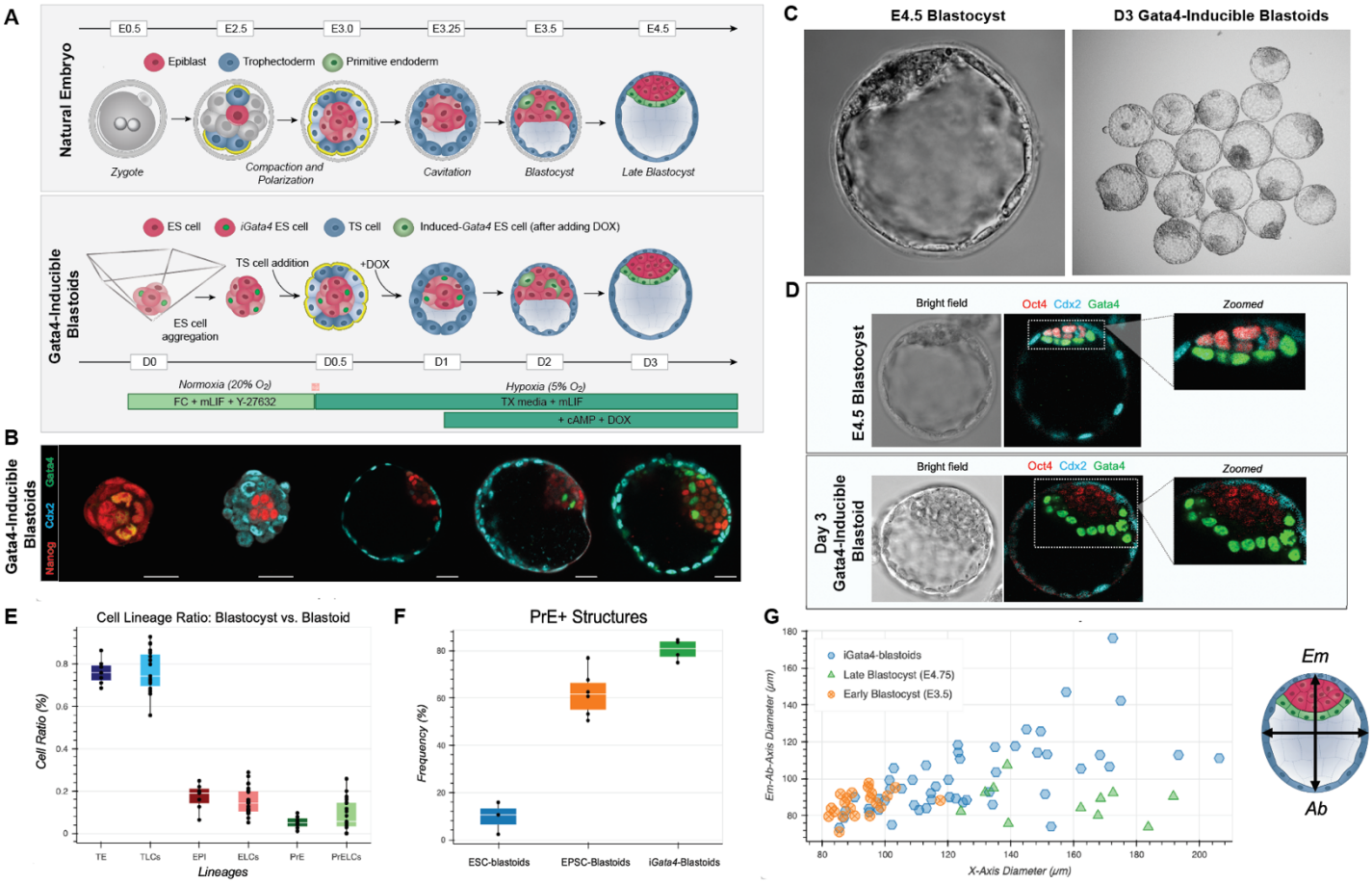


Figure 4.1: Combining ESCs, TSCs, and Gata4-inducible ESCs to form blastocyst-like structures.
A. Top: Schematic overview for natural mouse pre-implantation development from zygote to blastocyst. **Bottom:** Scheme for 3D aggregation protocol in Aggrewells for blastocyst-like structures using ESCs, Gata4-inducible ESCs, and TSCs. **B.** Representative examples of blastocyst-like structures at different timepoints of aggregation showing expression of key markers for each of the three linages: Nanog for epiblast, red; Cdx2 for trophectoderm, cyan; Gata4 for primitive endoderm, green. **C. Left:** Representative brightfield image of E4.5 natural mouse blastocyst. **Right:** Representative brightfield images of Day 3 (D3) Gata4-inducible blastoids. **D.** Brightfield images of either E4.5 natural blastocyst (top, left) or D3 Gata4-inducible blastoids (bottom, left). Blastocyst and Gata4-inducible blastoids are also immunostained for key lineage markers: Oct4 for epiblast, red; Cdx2 for trophectoderm, cyan; Gata4 for primitive endoderm, green. **E. Left:** Efficiency of primitive endoderm formation in blastocyst-like structures generated using ESCs, EPSCs, or ESCs and Gata4-inducible ESCs (123 ES-blastoids, n = 3; 233 EPSC-blastoids, n = 6; 125 Induced blastoids, n = 4). Present primitive endoderm cells were indicated by immunostaining of either Gata4, Gata6, or Sox17 in structures containing blastocyst-like morphology (i.e. visible cavity and ICM-like compartment). **F:** Cell lineage ratios in natural blastocysts compared to Gata4-inducible blastoids. Natural embryo lineages consist of trophectoderm (TE, dark blue), epiblast (EPI, red), and primitive endoderm (PrE, dark green). Gata4-inducible blastoid lineages consist of trophectoderm-like cells (TLCs, light blue), epiblast-like cells (ELCs, pink), and primitive endoderm-like cells (PrELCs, light green). For natural embryos, n = 7; for Gata4-inducible blastoids n = 23. **G:** Quantification showing the length of blastoids (blue hexagon, n = 49) compared to early blastocysts (E3.5, orange circle, n = 21) and late blastocysts (E4.5, green triangle, n = 11). Length was measured in μm along the embryonic (Em) to abembryonic (Ab) pole.

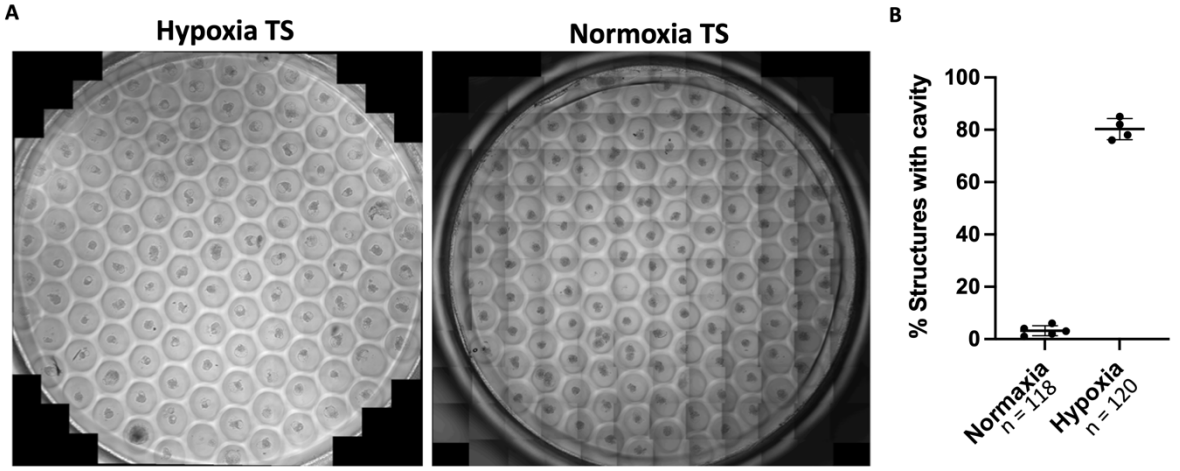


Figure 4.2: Hypoxic conditions show increased cavitation efficiency of blastoids than normoxic conditions
A. Gata4-inducible blastoids were generated in either low (5% O^2, hypoxic) or normal (20% O^2, normoxic) conditions. Bright field images of structures generated for hypoxic conditions (left) or normoxic conditions (right) are shown. **B.** Quantification of structures with cavity formation for each of the two conditions. n = 118 structures for normoxia, n = 120 structures for hypoxia.

4.3.2. Gata4-inducible ESCs solely contribute to primitive endoderm lineage

We next wanted to ensure that our G4i-ESCs were indeed generating the PE-like lineage, as opposed to spontaneous differentiation of wildtype ESCs. To assess this, we developed an H2B:RFP/tetO-Gata4 ESC line; these cells constitutively express histone H2B:RFP along with the ability to transiently express Gata4 in the presence of dox (Fig. 4.3A). As expected, RFP co-localizes with key PE markers such as Gata4, Gata6, and Sox17 in our G4i-blastoids, with 94.3% of cells analyzed showing RFP/PE-marker co-expression (Fig. 4.3C, D); this signifies that the majority of cells contributing to our PE-like lineage do indeed originate from our Gata4i-ESC population. We also find that induction of Gata4 in G4i-ESCs is a robust method for inducing the PE fate, as only 4.1% of RFP+ cells failed to induce any of the three lineage markers (Fig 4.3D). It is also important to note that Gata4/Gata6/Sox17 expression is rarely found in RFP-negative cells, occurring in only 1.6% of cells analyzed, indicating that our PE-like lineage is not generated from spontaneous differentiation of wildtype ESCs. In addition to proper expression of the designated markers, RFP-positive cells nicely segregate from the EPI-like compartment showing proper sorting adjacent to the blastocoel cavity (Fig. 4.3C).

To further validate that our dox-induced cell population was contributing to this lineage, we generated structures either with or without doxycycline and evaluated lineage formation accordingly. To ensure that *Gata4*-inducible cells were indeed present in our structures and loss of the PE lineage was not a result of seeding issues, we again utilized our H2B:RFP/tetO-Gata4 ESC line. As expected, blastoid structures formed in the presence of doxycycline upregulated expression of Gata4 within 12 hours, and showed proper segregation of these cells by 48 hours post-induction (Fig. 4.3E, top). Conversely, blastoids that were not treated with doxycycline showed little to no Gata4 expression, and nearly complete loss of the PE-like lineage in these structures despite containing RFP/Gata4-inducible cells (Fig 4.3E, bottom). To evaluate this induction in more detail, we performed a timecourse monitoring five distinct timepoints: t = 24 h, pre-dox induction; t = 48 h, with and without dox induction; and t = 72 h with and without dox induction (Fig. 4.4). At the initial timepoint, prior to any addition of dox, we see that the majority of cells (90%) are RFP+/Gata4-. We do also see a small population of cells (7%) that are RFP+/Gata4+; this perhaps indicates that our tetO-Gata4 cassette presents some leakiness. Finally, a minority of cells (3%) are RFP-/Gata4+, indicating that our G4i-ESCs may have lost H2B:RFP expression, or our wildtype population has begun expressing Gata4. At the t = 48 h and t = 72 timepoints without induction of dox, we see that the majority of cells (77% and 83%, respectively) remain RFP+/Gata4-, indicating a failure to induce Gata4 and the subsequent PE-lineage (Fig. 4.4). Alternatively, at t = 48 h and t = 72 h in the dox-induced structures, we see a significant increase in RFP+/Gata4+ cells (89% and 89%, respectively) (Fig. 4.4). We do see that 11% of cells in both dox-induced conditions remain RFP+/Gata4-, signifying that a portion of cells do fail to upregulate Gata4 expression, even in the presence of doxycycline (Fig. 4.4).

We additionally wanted to evaluate the optimal timing of dox-induction in our structures. To do this we performed a time course where we induced G4i-ESCs at initial seeding (t = -6 h), TSC seeding (t = 0 h), 12 h after TSC seeding (t = 12 h), and finally 24 h after seeding TSCs (t = 24 h). We then allowed structures to grow for 72 h, and evaluated cavitation efficiency. In doing so, we found that dox-induction prior to t = 24 h greatly affects cavity formation in blastoid structures, with the first two time points demonstrating less than ~40% cavitation

efficiency (Fig. 4.3B). The slightly later timepoint, at t = 12 h, showed partial rescue with ~57% of structures cavitating (Fig. 4.3B). We ultimately concluded that dox-induction at t = 24 h was the optimal condition, as we reached our maximum cavitation efficiency (~71% of total structures) at this timepoint (Fig. 4.3B).

We have demonstrated that our Gata4i-ESCs successfully initiate a PE-like lineage as indicated by the expression of the Gata4, Gata6, and Sox17. However, in addition to these key transcription factors, we would also expect to see secretion of the anti-adhesive sialomucin Podocalyxin (Podxl) along the PE-blastocoel cavity interface[53,54]. Indeed, we do see upregulation of Podxl along our RFP-positive cells adjacent to the blastocoel cavity, as well as a concomitant decrease in adhesion proteins such as E-Cadherin and F-actin at this boundary. Together this suggests that the PE-like lineage shows additional similarities to what would be expected in the natural embryo, and perhaps signifies a PE lineage preparing to migrate along the inner cavity for further differentiation into parietal and visceral endoderm. Analysis of later stage G4i-blastoids will be required to further validate this progression.

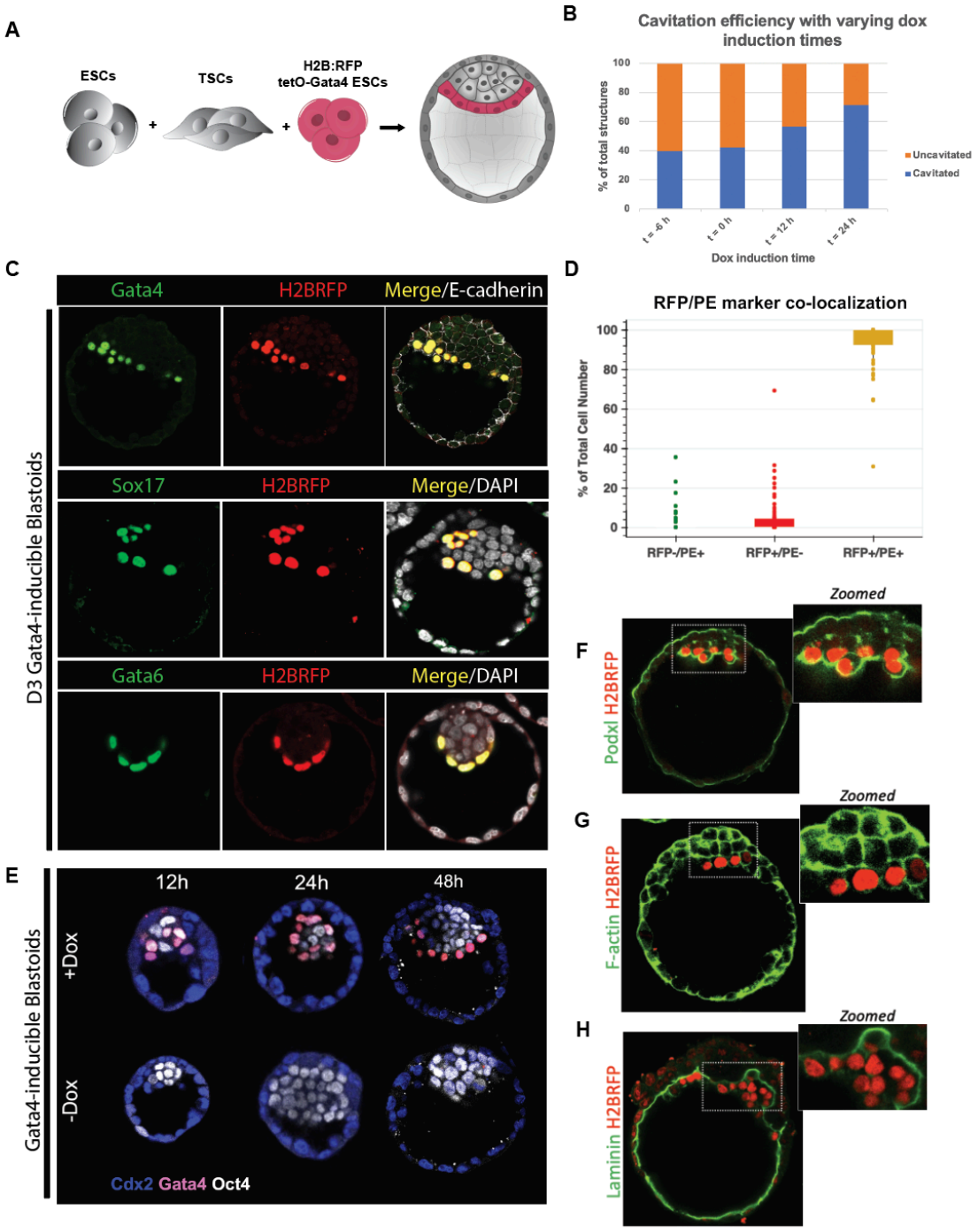


Figure 4.3: Characterization of primitive endoderm lineage in Gata4-inducible blastoids
A. Schematic of generating blastoids with tetO-Gata4/H2B:RFP ESCs, ESCs, and TSCs. **B.** Quantification of cavitation efficiency given varying induction times. Time t = -6h signifies addition of ESCs, whereas t = 0h denotes addition of TSCs. n = 1867 structures, 3 replicates. **C.** Representative examples of blastoids generated with tetO-Gata4/H2B:RFP ESCs showing co-localization of RFP and primitive endoderm (PE) markers, Gata4, Gata6, or Sox17.**D.** Quantification of RFP/PE marker colocalization. PE was identified via immunostaining of either Gata4, Gata6, or Sox17. Structures were classified in 3 categories: RFP-/PE+, RFP+/PE-, and RFP+/PE+. 243 structures, n = 4 experiments. **E.** Representative blastoids collected at certain timepoints cultured with or without doxycycline (dox) induction of tetO-Gata4 construct. **F.** Immunostaining of Podocalyxin (Podxl) secretion in blastoids. **G.** Assembly of F-actin filaments in blastoids. **H.** Establishment of Laminin basement membrane in Gata4-inducible blastoids along the blastocoel cavity.

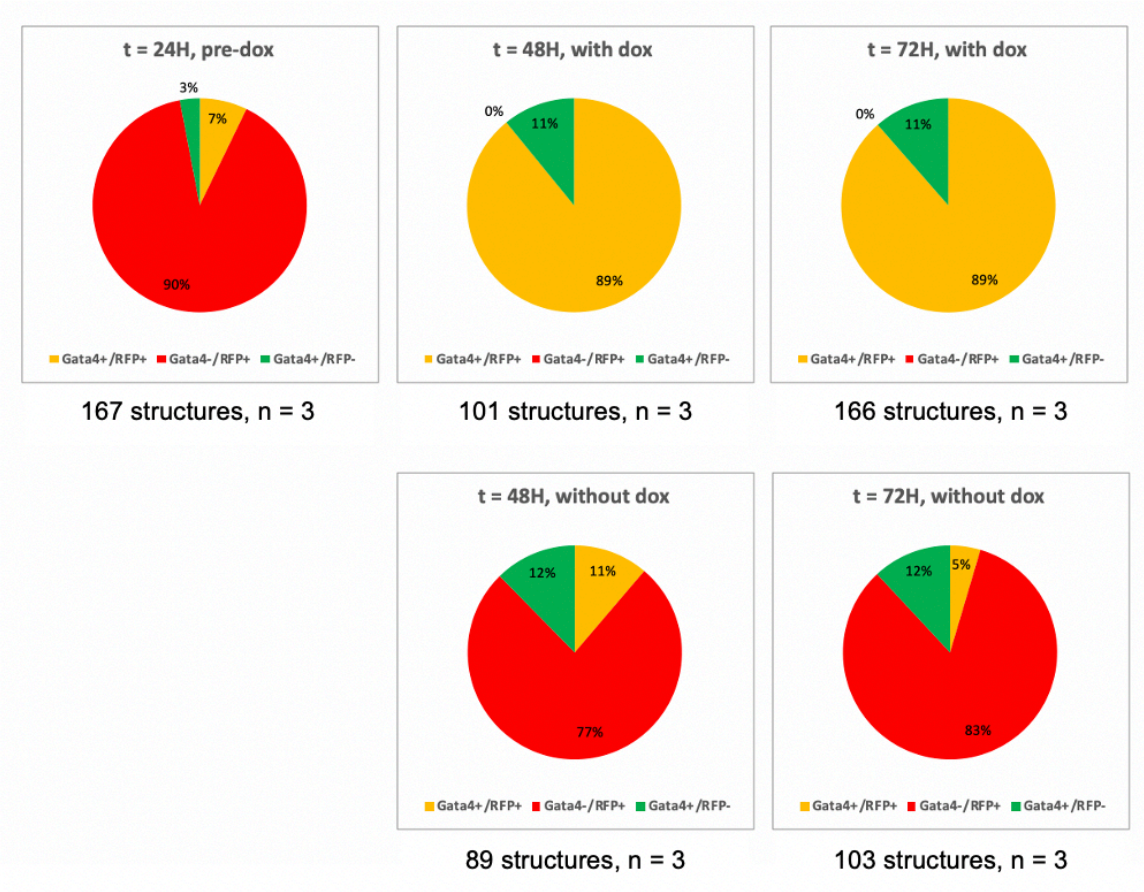


Figure 4.4: Quantification of Gata4-inducible blastoids with or without doxycycline induction
Quantification of Gata4-inducible blastoids generated with or without doxycycline induction and collected at 3 distinct timepoints: t = 24hrs after seeding (prior to doxycycline induction); t = 48hrs after seeding (24hrs post-dox induction in "with dox" condition); and t = 72hrs (48hrs post-dox induction in "with dox" condition). The portion of structures that were Gata4+/RFP+ are shown in yellow, Gata4-/RFP+ are shown in red, Gata4+/RFP- are shown in black, and Gata4-/RFP- are shown in grey. For t = 24H, pre-dox, n = 3, 167 structures. For t = 48H, with dox, n = 3, 101 structures. For t = 72H, with dox, n = 3, 166 structures. For t = 48H, without dox, n = 3, 89 structures. For t = 72H, without dox, n = 3, 103 structures.

4.3.3. G4i-blastoids undergo post-implantation remodeling *in vitro*

We next wanted to assess whether or not G4i-blastoids demonstrate enhanced developmental capacity and could transition to post-implantation stages. Using previously established post-implantation *in vitro* culture methods[155], we found that G4i-blastoids could indeed generate the elongated egg-cylinder-like structures reminiscent of post-implantation morphology. Additionally, these structures demonstrated formation of a lumenized EPI-like compartment,

an abutting ExE-like compartment, and both of these lineages were surrounded by a VE-like outer monolayer. Excitingly, we also found that post-implantation G4i-structures were able to upregulate Brachyury (T) on one side of the embryo, indicating another major hallmark of post-implantation development. While this demonstrates the ability for G4i-blastoids to undergo significant post-implantation remodeling and possible initiation of axis patterning, efficiency of this transition is very low (1 in 200). One potential explanation for this low efficiency may be due to insufficient formation of our polar TE/ExE lineages, as by immunostaining we see that this lineage in particular appears deficient.

We also performed experiments to assess the *in vivo* developmental potential of our G4i-blastoids via transplantation assays into pseudopregnant mice. Similar to previous blastoid models[86–88,150], we found that our blastoids could initiate decidualization, but failed to progress further *in vivo*. Previous models were believed to fail at implantation due to improper formation of the primitive endoderm lineage, and subsequent inability to form the Reichert's membrane. In our model, however, the Gata4-inducible ESCs form a much more robust primitive endoderm lineage, and yet still fail to implant. An alternative explanation for why these blastoids cannot implant is perhaps due to improper formation and later differentiation of the trophectoderm lineage into polar and mural TE. A key event triggering the divergence of these two lineages is receptor binding of Fgf4 ligand produced by the epiblast compartment. The TE cells within close proximity of this signal will generate the polar TE; a more pluripotent population that will later give rise to the ExE compartment, ectoplacental cone, and placenta. Alternatively, TE cells on the abembryonic pole, which are separated from the Fgf4-producing epiblast by the blastocoel cavity, will generate the mural trophectoderm. These cells will lose pluripotency markers, triggering differentiation into trophoblast giant cells, and invade the uterus to commence implantation. In our blastoids culture media, however, we introduce endogenous Fgf4, a major component of TSC culture media. It is possible that introduction of this factor creates aberrant signaling to our putative mural TE population and thus limits its differentiation and ability to induce implantation. Removal of this component at any point, however, greatly hinders survival of our TSC population and hinders blastoid formation. Thus, alternative culture conditions that exclude

Fgf4 yet maintain TSC survival will perhaps be necessary to trigger implantation of these structures. To further address this topic, we have performed scRNA-seq of our blastoids and natural embryos to better understand which lineages are lacking in our model.

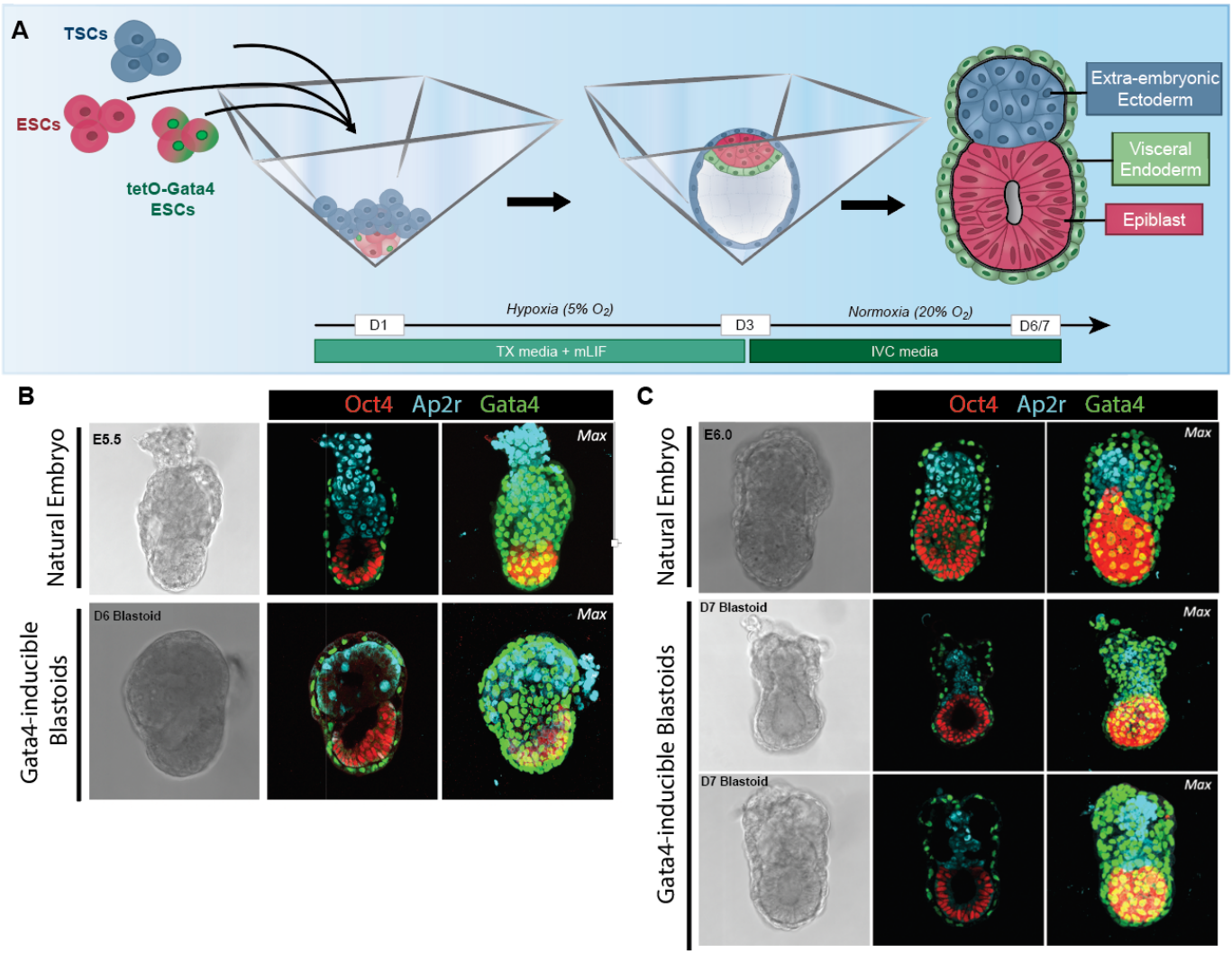
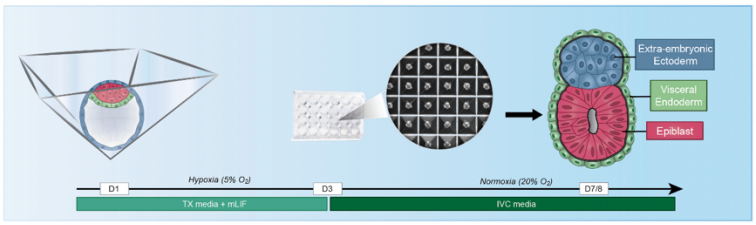
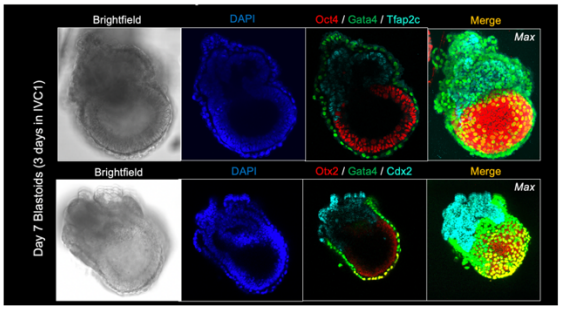
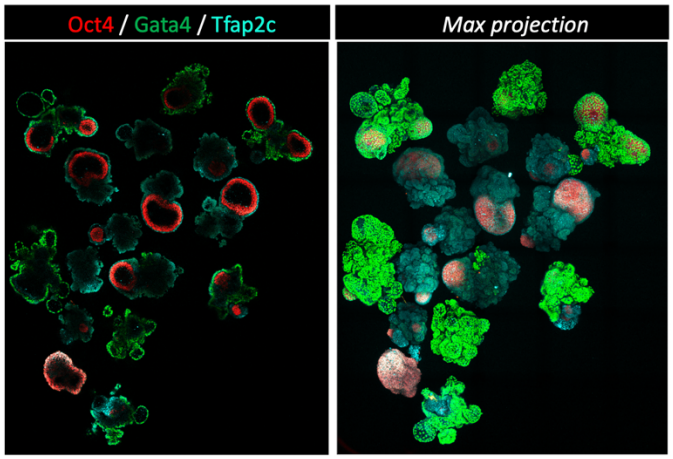


Figure 4.5: Gata4-inducible blastoids transition to post-implantation morphology *in vitro*.
A. Schematic of the protocol used to transition blastoids to post-implantation morphology. **B.** Natural E5.5 mouse embryo **(top)** or D6 blastoid **(bottom)**, each cultured *in vitro* for 48h in IVC media. Immunostaining shows extra-embryonic ectoderm-like compartment (cyan, Ap2r), epiblast-like compartment (red, Oct4), and visceral endoderm-like lineage (green, Gata4). **C.** Natural E6.0 mouse embryo **(top)** or D6 blastoid **(bottom)**, each cultured *in vitro* for 72h in IVC media. Immunostaining shows extra-embryonic ectoderm-like compartment (cyan, Ap2r), epiblast-like compartment (red, Oct4), and visceral endoderm-like lineage (green, Gata4).







Characteristic	# of Structures
Contains all 3 lineages, well-organized	3 / 40
Contains all 3 lineages, disorganized	11 / 40
Missing VE lineage	20 / 40
Missing EPI compartment	5 / 40
Missing ExE compartment	10 / 40
Multiple EPI lumens	3 / 40
Disorganized VE	12 / 40
Brachyury + Structures	2/10

Figure 4.6: Alternative method for post-implantation culturing Gata4-inducible blastoids *in vitro*.
A. Schematic for culturing Gata4-inducible blastoids in Aggrewells. **B.** Representative confocal images of structures cultured to post-implantation stages in Aggrewells. Epiblast is marked by Oct4 in red; Visceral endoderm is marked by Gata4 in green; Extra-embryonic ectoderm is marked by Tfap2c in cyan. **C.** Two representative images of structures cultured to post-implantation stages in Aggrewells with proper lineage segregation. **D.** Table of quantifications for various

Transfer experiment number	Number of Deciduae (Blastocyst)	Number of Deciduae (Blastoids)	Number of embryos (Blastocyst)	Number of embryos (Blastoids)
1	7	0	4	0
~~2~~	~~2~~	~~1~~	~~0~~	~~0~~
~~3~~	~~1~~	~~1~~	~~0~~	~~0~~
4	4	3	1	0
~~5~~	~~0~~	~~0~~	~~0~~	~~0~~
~~6~~	~~1~~	~~1~~	~~0~~	~~0~~
~~7~~	~~0~~	~~2~~	~~0~~	~~0~~
~~8~~	~~0~~	~~0~~	~~0~~	~~0~~
9	4	0	1	0
10 (UC Irvine)	0 / 10	0 / 10	0 / 10	0 / 10
11 (UC Irvine)	10 / 10	1 / 10	10 / 10	0 / 10
12 (UC Irvine)	9 / 10	0 / 10	9 / 10	0 / 10
13 (UC Irvine)	7 / 10	0 / 10	7 / 10	0 / 10
14 (UC Irvine)	7 / 10	1 / 10	7 / 10	0 / 10
15 (UC Irvine)	10 / 10	2 / 10	10 / 10	0 / 10
16 (UC Irvine)	4 / 10	2 / 10	4 / 10	0 / 10
17 (UC Irvine)	9 / 10	0 / 10	9 / 10	0 / 10

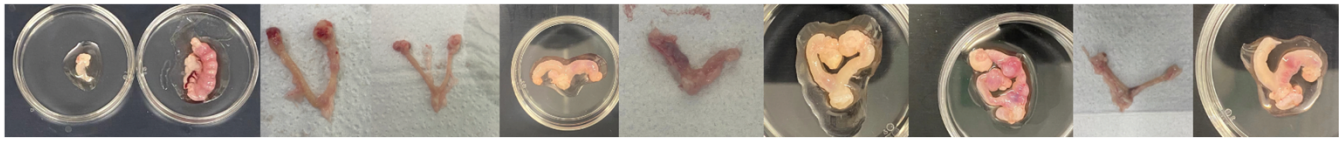


Figure 4.7: Gata4-inducible blastoids cannot transition to post-implantation morphology *in vivo*.
A. Table outlining results of 11 transfer experiments in which both natural blastocysts and Gata4-inducible blastoids were transferred into separate uterine horns of pseudopregnant mice. Experiments done with UC Irvine Transgenic facility are indicated. Experiments that have been crossed out represent failed transfers, as neither natural blastocysts nor blastoids were able to transplant. **B.** Representative images of dissected uterine horns.

4.3.4. scRNAseq of G4i-Blastoids

To further assess efficacy of our model, we performed single cell RNA-seq analysis on our G4i-blastoids and natural embryos at distinct timepoints. For natural embryos, we collected samples at E3.5, E4.5, and E4.75, in order to capture the full transition of each lineage through the pre-implantation stages. To better understand our G4i-blastoid system and how culturing conditions impact these models, we collected three distinct samples: Day 2 G4i-blastoids grown in hypoxia, Day 3 G4i-blastoids grown in hypoxia, and Day 3 G4i-blastoids grown in normoxia. In addition to our own samples, I have also incorporated data from Sozen et al., including and additional E4.5 natural embryo timepoint, EP-blastoids (blastoids generated from EPSCs), and 2iLif-Blastoids (blastoids generated from wildtype ESCs). Initial analysis reveals that both the natural embryos as well as the stem cell-derived blastoid samples segregate into three distinct cell clusters, and each of these clusters can be attributed to one of the three key lineages: epiblast, primitive endoderm, and trophectoderm (Fig. 4.8C). In order to classify cells into distinct lineages, I first generated a list of the top 100 differentially expressed genes in each lineage using an independent dataset published previously (Nakamura et al., 2016). I then performed module scoring for each score for each lineage (i.e. each cell receives 3 scores, one for each lineage). Whichever lineage score ranks highest denotes the lineage in the given cell. Finally, if all scores of a particular cell are below a certain threshold (as determined by natural embryo scoring), cells were deemed "undefined".

Analyzing the EPI-labeled cells, we see high upregulation of key pluripotency markers *Pou5f1*, *Sox2*, and *Nanog* as we would expect (Fig. 4.9A, B). Additionally, we see high expression of *Gata4* in our primitive endoderm lineage, the cell population derived from our *Gata4*-inducible ESCs; furthermore, we see co-expression of other PE markers *Gata6*, P*dgfra,* and *Sox17* in these cells as well (Fig. 4.9A, B). Lastly, we see expression of key markers *Cdx2, Gata3,* and *Tfap2c* in our TE-labeled cells, denoting the trophectoderm-like population (Fig. 4.9A, B). Interestingly, we also see moderate expression of *Tfap2c* in the epiblast-like lineage of our stem cell-derived model; this perhaps suggests a key divergence

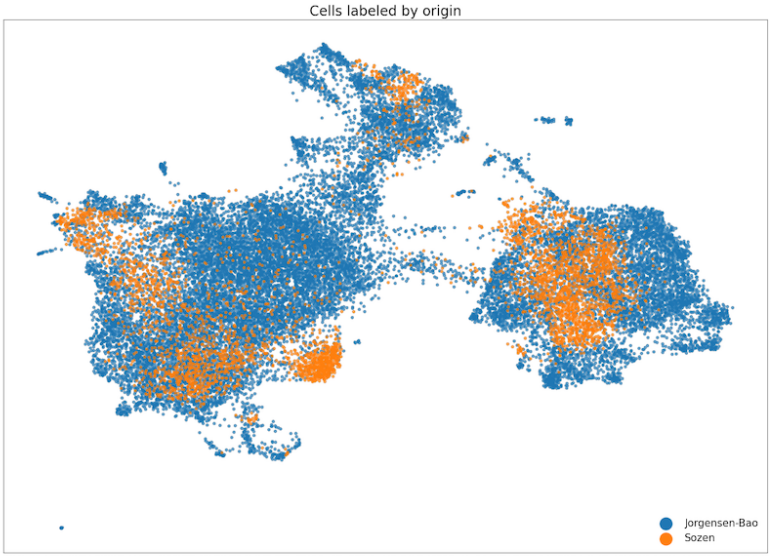
between EPI differentiation in the natural embryo compared to our *in vitro* models (Fig. 4.9A, B).

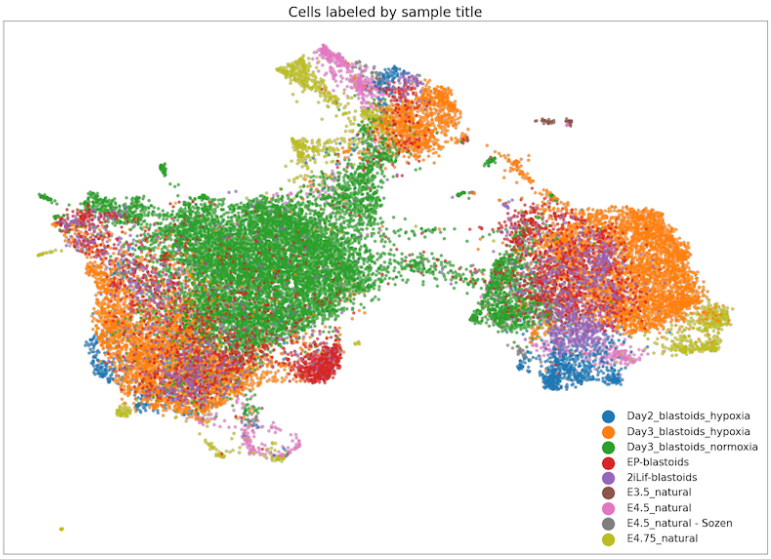
In order to better understand the development and characterization of our G4i-blastoids, I further subsetted the data by lineage to better understand where our model compares and diverges to the natural embryo. At first glance of the EPI-only UMAP clustering, two clear trends emerge. The first is that a subset of the EPI-like cells originating from Day 3 Blastoids grown in normoxia scatter away from the main EPI-cluster (Fig. 4.10A). This perhaps indicates a unique population of cells originates in this condition, and implies that normoxic versus hypoxic conditions could alter the trajectory of EPI cells. Secondly, we see that the E3.5 sample clusters independently from the other samples (Fig. 4.10A); this is not entirely surprising on account that E3.5 EPI may not have fully differentiated from an ICM-like state, further evidenced by the residual expression of Gata6 in this population (Fig. 4.9A, Fig.4.10C). Looking at the Pearson correlation of our G4i-blastoid EPI compared to that of the natural embryo, we see that all three conditions are most closely related to the later E4.5 and E4.75 timepoints (Fig. 4.10B). This correlation with later timepoints is also consistent with the observation that the stem cell-derived populations show higher expression of primed pluripotency markers, indicating that they are further along in their development (Fig. 4.9C). The EP- and 2iLif-Blastoids show closest similarity to the E4.5 natural embryos from the same dataset, yet modest divergence from our own natural timepoints; this is perhaps indicative of residual batch effects despite performing corrections. Incorporating additional independent natural timepoints will be interesting to further assess correlation of this lineage.

We next sought to further characterize the PE lineage. Similar to what we observed with the EPI lineage, we again see scattering of cells from the Day 3 Blastoids grown in normoxia, again implying a unique cell fate trajectory dependent on oxygen conditions (Fig. 4.11A). Unsurprisingly, we see upregulation of several collagen and laminin markers (*Col4a1, Col4a2, Lama1, Lamc1*) in our PE-labeled cells, with consistent expression across all samples suggesting upregulation of these pathways are robustly conserved in both the *in vivo* and *in vitro* setting. Notably, the divergent cells from Day 3 G4i Normoxia blastoids show

decreased expression of these extra-cellular matrix genes, and hints at one major discrepancy in this population. Lastly, the Pearson correlation shows that the PE lineage in all G4i-blastoids is again most reminiscent of E4.5 and E4.75 PE (Fig. 4.11A).

Lastly, we aimed to investigate the TE lineage. According to the Pearson correlation, G4i-blastoids, regardless of timepoint or condition, showed highest similarity to E4.75 TE (Fig. 4.12A); this is surprising on account that G4i-blastoids fail to implant yet seem to be progressing developmentally. One hypothesis for why G4i-blastoids fail to implant despite this advanced trajectory is that we are unable to properly form the mural TE lineage, which is responsible for invasion into the maternal uterine tissue. By looking at markers for global TE, polar TE, and mural TE in our G4i-blastoids, it is clear that both global TE and polar TE programs are present in the TE-like population of all conditions (Fig. 4.12C, D). The mural TE, however, is almost entirely absent from all samples, aside from a small portion of cells at the E4.75 timepoint. The polar TE, due to its proximity to the EPI lineage, receives high levels of Fgf4 signaling which allows for maintenance of pluripotency required for downstream differentiation into subsequent TE derivatives such as the ExE and ectoplacental cone. The mural TE, however, separated by the blastocoel cavity, does not receive these signals and is allowed to differentiate into the invasive cells that penetrate the uterine wall. Given that G4i-blastoids are cultured in media containing exogenous Fgf4 and LIF, two factors that promote a stem cell-like state, it is likely that cells are unable to proceed to a mural TE fate as differentiation is aberrantly blocked. Removal of either of these factors during culture, however, leads to poor maintenance of the TE lineage, and severely hinders cavity formation. It will therefore be important to assess alternative strategies to maintain the TE population while simultaneously allowing subsequent differentiation.





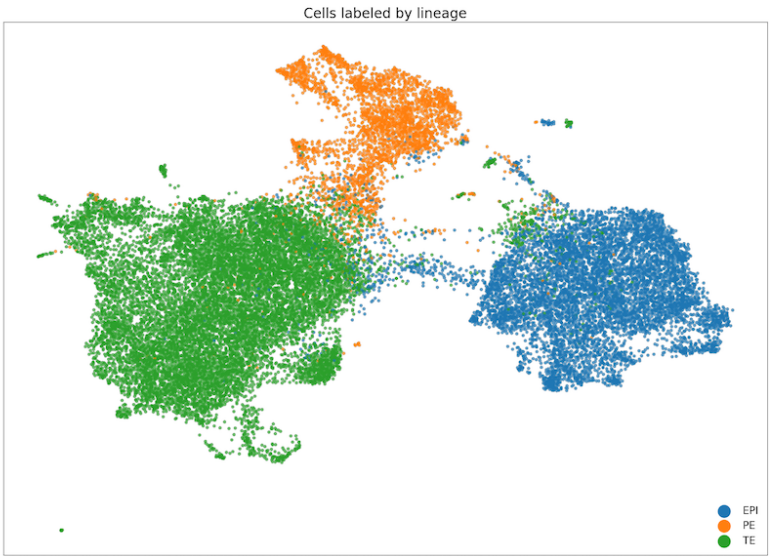
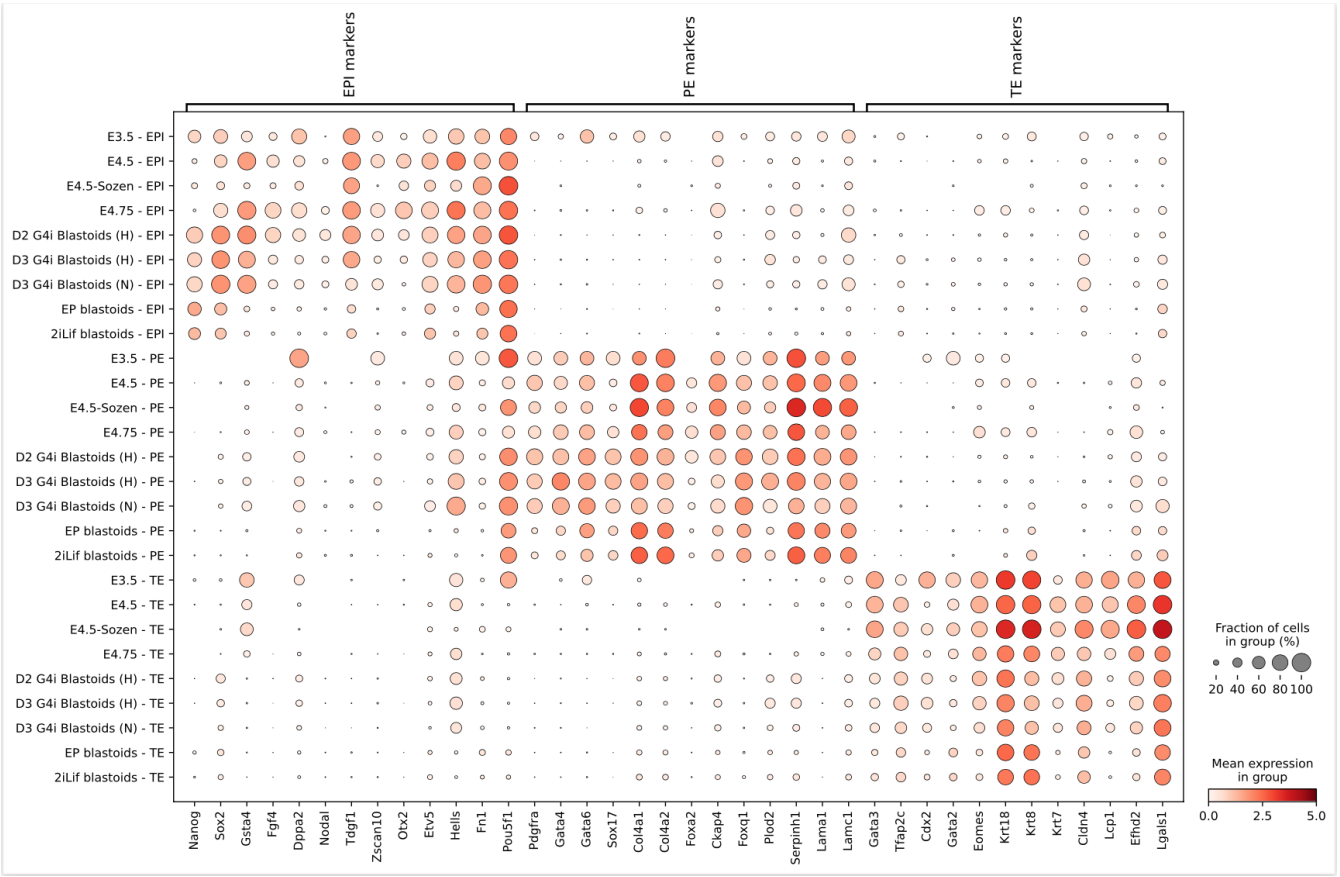


Figure 4.8: Clustering of natural and stem cell-derived embryos via uniform manifold approximation and projection (UMAP).
A. Uniform manifold approximation and projection (UMAP) labeled by cell origin, either from Jorgensen-Bao (blue; 23,400 cells) or Sozen et al. (orange; 4,798 cells) **B.** UMAP showing breakdown of cells by sample title: Day 2 G4i-Blastoids – Hypoxia (blue; 1,445 cells); Day 3 G4i-Blastoids – Hypoxia (orange; 8,784 cells); Day 3 G4i-Blastoids – Normoxia (green; 10,676 cells), EP-Blastoids (red; 3,177 cells); 2iLif-Blastoids (purple; 1,539 cells); E3.5 Natural Embryos (brown; 65 cells); E4.5 Natural Embryos (pink; 959 cells); E4.5 Natural Embryos from Sozen et al. (grey; 82 cells); E4.75 Natural Embryos (yellow; 1,471 cells). **C.** UMAP showing lineage scoring of all cells into epiblast-like cells (EPI; blue; 8,890 cells), primitive endoderm-like cells (PE; orange; 3,543 cells), trophectoderm-like cells (TE; green; 15,855 cells). Cells were annotated based on expression of 100 genes per lineage generated from Nakamura et al.



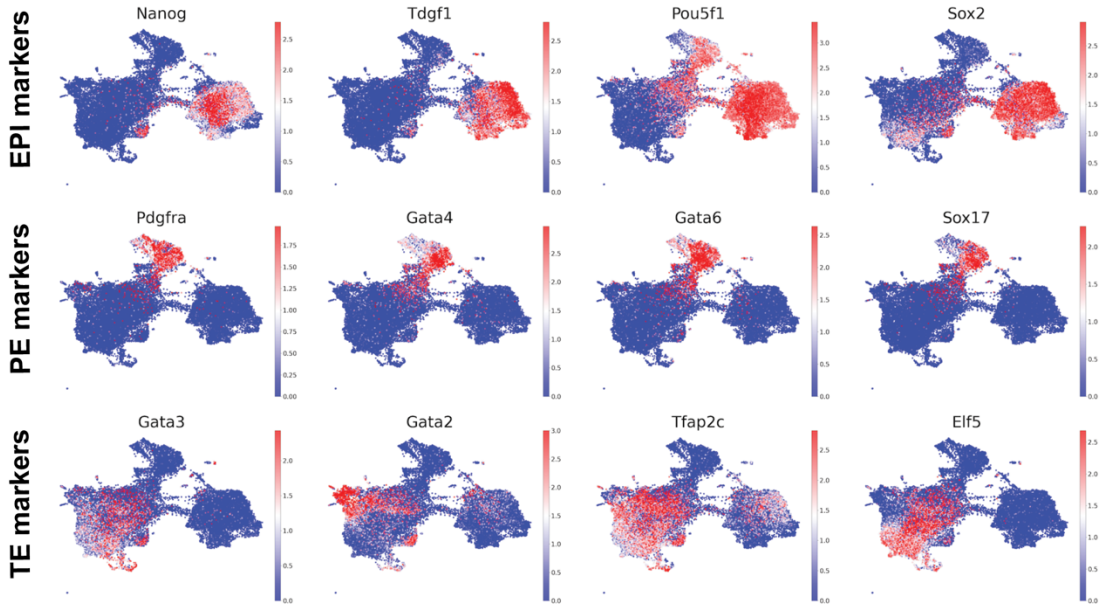


Figure 4.9: Each of the three embryo lineages are present in both natural and stem cell-derived embryo samples.
A. Expression of select key markers for the epiblast (EPI), primitive endoderm (PE) and trophectoderm (TE) lineages.
B. Heatmaps showing expression of canonical lineage markers for EPI, PE, and TE lineages within each of the UMAP clusters.

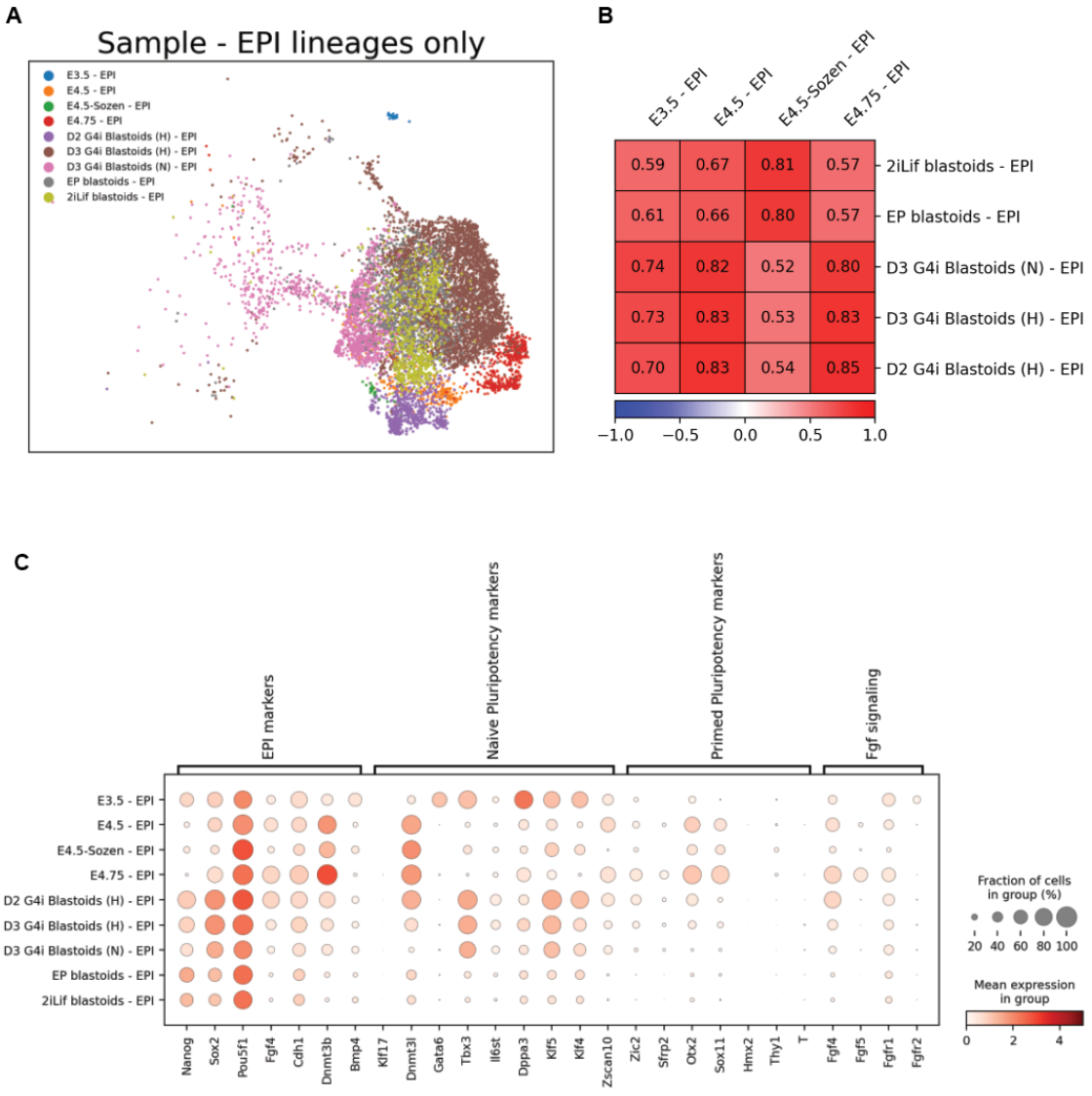


Figure 4.10: Characterization of epiblast lineage in stem cell-derived embryo models.
A. Pearson correlation matrix showing linear correlation between natural embryo trophectoderm (TE) populations in comparison to TE-like lineage of stem cell derived embryo models. **B.** UMAP showing clustering of TE-labeled cells according to source. **C.** Module scoring for either Polar TE (left) or Mural TE (right). Scores were generated using a list of 100 genes per lineage. **D.** Expression of a subset global TE markers, polar TE markers, and mural TE markers within each sample TE lineage.

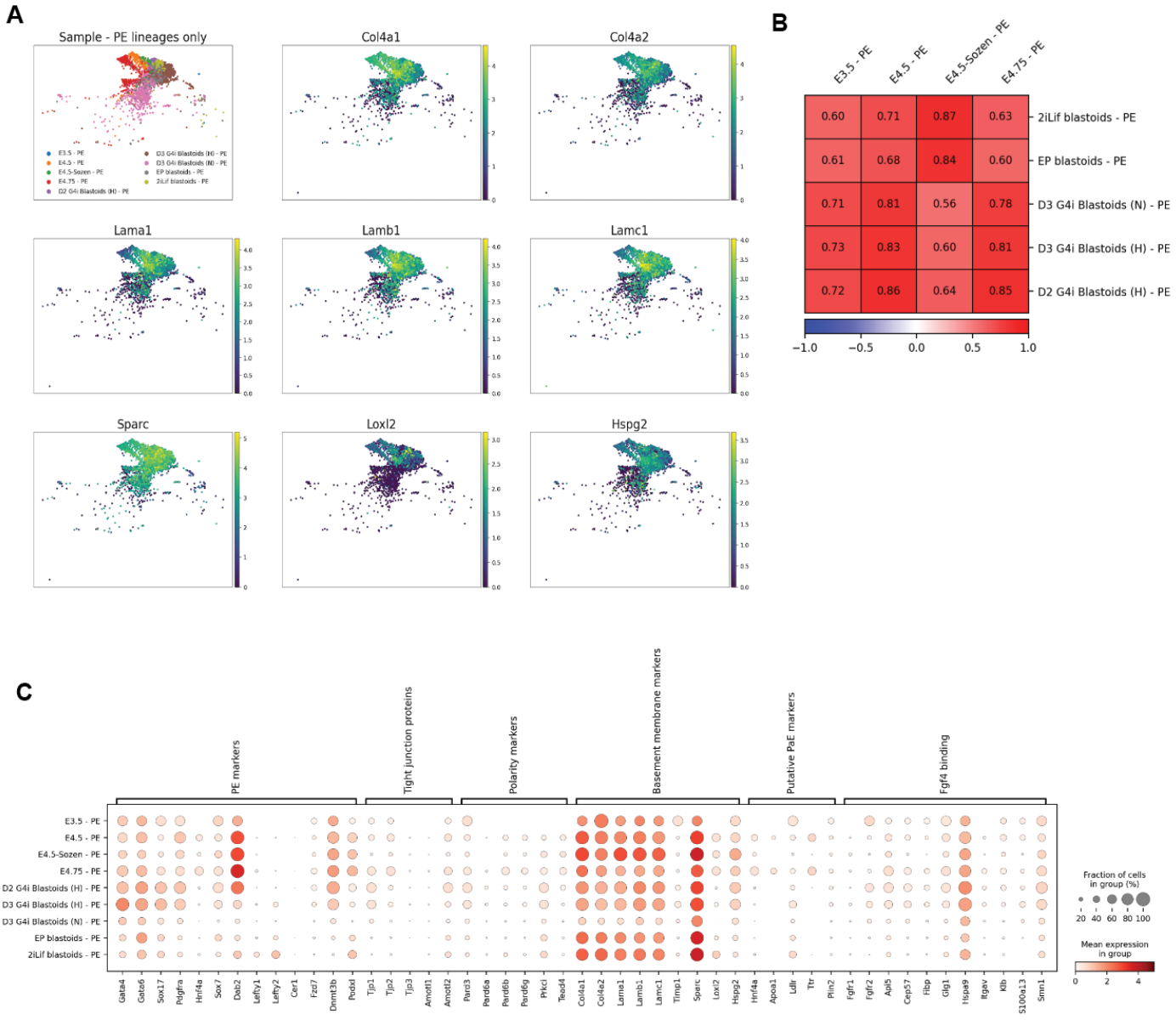


Figure 4.11: Characterization of primitive endoderm lineage in stem cell-derived embryo models.
A. In the top left, UMAP showing clustering of primitive endoderm-labeled cells according to source. **B.** Pearson correlation matrix showing linear correlation between natural embryo primitive endoderm (PE) populations in comparison to PE-like lineage of stem cell derived embryo models. **C.** Expression of several gene classes important to the PE lineage including: global PE markers, tight junction proteins, polarity markers, basement membrane markers, putative parietal endoderm (PaE) markers, and genes involved in Fgf4 binding.

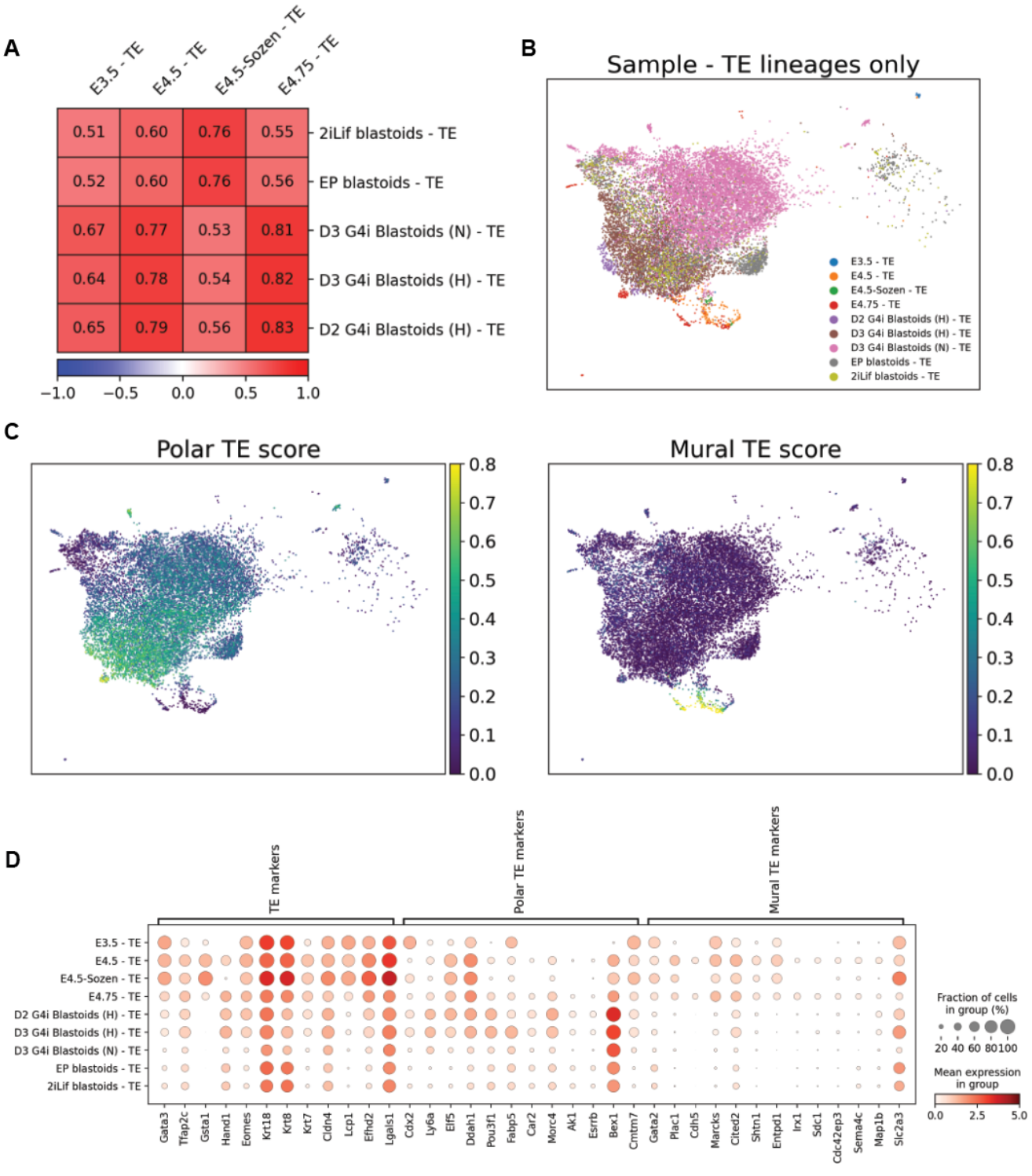


Figure 4.12: Characterization of trophectoderm lineage in stem cell-derived embryo models.
A. Pearson correlation matrix showing linear correlation between natural embryo trophectoderm (TE) populations in comparison to TE-like lineage of stem cell derived embryo models. **B.** UMAP showing clustering of TE-labeled cells according to source. **C.** Module scoring for either Polar TE (left) or Mural TE (right). Scores were generated using a list of 100 genes per lineage. **D.** Expression of a subset global TE markers, polar TE markers, and mural TE markers within each sample TE lineage.

4.4. Conclusion

In recent years, several pre-implantation stem cell-derived embryo models have emerged showing the ability to mimic the natural embryo in size, morphology and initiation of lineage formation[86–88]. With the ability to generate and monitor hundreds of structures in an easily controllable manner, these models are an exciting avenue to explore the mechanisms underpinning early development *in vitro*. Nevertheless, these models show limited cavitation efficiency, and imperfect lineage formation, most notably in the primitive endoderm. Additionally, these blastoids show minimal capacity to transition to post-implantation morphology *in vitro*.

In the system we describe here, we expand on previous models to generate blastoids by combining together three different stem cells: ESCs, TSCs, and tetO-Gata4 ESCs. This model not only demonstrates a significant increase in cavitation efficiency, but also the development of a more robust primitive endoderm lineage as seen via immunostaining of key markers as well as scRNAseq analysis. Similar to previous models, we show that our G4i-blastoids are capable of transitioning from pre- to post-implantation stages *in vitro*, yet fail to implant *in vivo* when transferred into pseudopregnant mice.

Based on scRNAseq analysis comparing G4i-blastoids to natural embryos, we are able to robustly generate EPI-, PE-, and TE-like lineages that are reminiscent of the E4.5/E4.75 embryo. While G4i-blastoids do seem to generate a polar TE-like population, they fail to develop any mural TE-like populations. This is likely in large part due to culture conditions maintaining a stem-like state, blocking further differentiation of downstream lineages. As Fgf4 and LIF appear necessary to maintain the overall TE population, it would be interesting as to whether or not a minimal concentration of these factors could be reached that allow for both cell survival and differentiation. Alternatively, a method that would allow directed/gradient levels of Fgf4 could also be sufficient for establishing the dichotomy of the polar and mural TE fates, and would more closely mimic events of natural development. Given that Day 2 G4i-Blastoids already show strong correlation to E4.75 embryos in all lineages, it is possible that transitioning G4i-blastoids from TX + LIF media to IVC at an

earlier timepoint may also help to trigger differentiation of the mural lineage; this may, however, consequently hinder maintenance of the polar TE fate, and thus may require additional factors to support this lineage. Furthermore, at the time of implantation, the mural TE elongates and undergoes key transcriptional and morphological changes; a process that does not appear to occur in our G4i-blastoid system. It is possible that mechanical cues within the uterine crypts is necessary to induce these changes, and therefore developing an *in vitro* system that mimics these mechanical forces may further help progression of this lineage.

With regards to hypoxic versus normoxic culture, there still remains open questions as to how this condition alters developmental progression. From our analysis with G4i-blatoids, it is clear that hypoxia shows drastic improvement with blastoid efficiency, and in particular seems critical for cavity formation within our structures. Additionally, it seems as though G4i-blastoids grown in normoxia show populations of cells that are still transitioning to distinct lineage states yet have not fully differentiated. Perhaps normoxic conditions promote a more stem-like state, and thus lineages are unable to differentiate further. We also see that cavitation is drastically hindered in normoxic conditions, which may also affect signaling gradients required for downstream differentiation. As natural pre-implantation embryos are grown in normoxic conditions, it remains unclear as to why stem cell derived embryos show much higher dependence for oxygen deprivation with regards to cavitation and lineage differentiation. Further analysis of these timepoints will be necessary to further elucidate this phenomenon.

In summary, our findings demonstrate that use of tetO-Gata4 ESCs for the formation of blastoids greatly improves on previous models both in efficiency and lineage differentiation. Nevertheless, there appears to be distinct divergences between our G4i-blstoids and the natural embryo that ultimately leads to a halt in developmental potential. Despite this, these G4i-blastoids do generate structures demonstrating clear differentiation of all three lineages expected of pre-implantation development, and many of the morphological features expected of the natural mouse embryo. As such, we believe this system offers exciting potential for use as an embryo-like model to be harnessed for high-throughput analysis of early

developmental events. Additionally, we anticipate that further comparison of our model to the natural embryo, and consequent modification, will lead to pivotal breakthroughs permitting the implantation of these stem cell-derived structures. Altogether, this system represents an exciting avenue for future applications, and demonstrates capability for disentangling the mysteries governing early developmental stages.

4.5. Materials, methods, and data availability

4.5.1 Methods

Cell culture

ESCs were cultured at 37°C in normoxic conditions (20% O_2, 5% CO_2) while TSCs were cultured at 37°C in hypoxic conditions (5% O_2). Cell lines were passaged after reaching 80% confluency. All cell lines were routinely tested once a month for mycoplasma using the ABM PCR Mycoplasma Detection Kit (G238, ABM).

ESCs: Wildtype ESCs and tetO-Gata4 ESCs were cultured using tissue culture-grade plates coated with 0.1% gelatin. For routine culture, cells were plated in N2B27 medium supplemented with 2iLIF (1μM MEK/ERK pathway inhibitor PD0325901 (72184, STEMCELL Technologies), 3μM GSK3 inhibitor CHIR99021 (72054, STEMCELL Technologies) and 10ng ml−1 Mouse recombinant LIF (78056, STEMCELL Technologies). N2B27 medium is composes of a 1:1 mix of DMEM/F12 (21331-020, Thermo Fisher Scientific) and Neurobasal A (10888-022, Thermo Fisher Scientific) supplemented with 1% v/v B-27 Supplement (17504044, Thermo Fisher Scientific), 0.5% v/v N-2 Supplement (17502048, Thermo Fisher Scientific), 100 μM 2-Mercaptoethanol (31350010, Thermo Fisher Scientific), 1% penicillin–streptomycin (15140122, Thermo Fisher Scientific) and 1% GlutaMAX (35050-061, Thermo Fisher Scientific).

TSCs: TSCs were cultured on tissue culture-grade plates with irradiated C57bl/6 Mouse Embryonic Fibroblasts (MEFs) (A34960, Thermo Scientific) in TSF4H medium. TSF4H medium is composed of Megacell™ RPMI-1640 Medium (M3817, Sigma-Aldrich)

containing 20% inactivated Fetal Bovine Serum (iFBS) (35-010-CV, Corning), 1% GlutaMAX (35050-061, Thermo Fisher Scientific), 1% Sodium Pyruvate (11360070, Thermo Fisher Scientific), and 1% penicillin-streptomycin, supplemented with 25 ng/mL FGF4 (5846-F4, R&D Systems) and 1 μg ml−1 heparin (H3149, Sigma).

Preparing and Plating Cell Suspensions for "AggreWell" Aggregation Experiments

Aggrewell™ 400 plates (34415, STEMCELL Technologies) were prepared following the manufacturer's protocol. Briefly, wells were rinsed with the anti-adherence rinsing solution (07010, STEMCELL Technologies), centrifuged for 5 min at 2000g and incubated at room temperature in the tissue culture hood for 20 minutes. The wells were then washed twice with 1 ml of 1X PBS. After PBS removal, 500 μl of filtered FCLY medium was added to each well and the plate was centrifuged for 5 min at 2000g and then placed at 37°C and 5% CO2 until ready to use. FCLY is composed of Feeder Cell (FC) media supplemented with 10ng ml−1 Mouse recombinant LIF and 2 nM ROCK inhibitor Y27632 (72304, STEMCELL Technologies). FC is comprised of Dulbecco's modified Eagle's medium (DMEM) (11995040, Gibco) with 15% iFBS, 1% Glutamax, 0.1 mM 2-mercaptoethanol, 1% nonessential amino acids (11140-050, Thermo Fisher Scientific), 1% sodium pyruvate, and 1% penicillin-streptomycin.

Generation of Blastocyst-like Structures in Aggrewell™ 400 plates

ESC and tetO-Gata4 ESCs were dissociated to single cells by incubation with 0.05% trypsin-EDTA (25300054, Thermo Scientific) at 37°C for 3 min. Cells were pelleted by centrifugation for 4 min at 1000 rpm and resuspended in FC media. Cells were counted using a hemocytometer and a total of 6000 ESCs and 4800 tetO-Gata4 ESCs were added per well in FCLY medium. Cell suspensions were added dropwise to the Aggrewells to a total volume of 1 mL FCLY per well. Empty wells were filled with PBS to humidify the local environment and minimize evaporation from neighboring wells. The Aggrewell plate was then centrifuged for 3 min at 100g and then placed at 37°C, in normoxic conditions (5% CO2 or 20% O2).

Once ESCs have aggregated to form a singular amorphous aggregate (~6-8 hours), TSCs were prepared for seeding. TSC colonies were dissociated to single cells using 0.05% Trypsin EDTA for 8-12 minutes, centrifuged, and resuspended in TSF4H. In order to remove inactive MEFs, the cell suspension was plated on plates coated wiuth 0.1% gelatin and placed in the 37°C incubator for 30 minutes. Following MEF depletion, cells were counted using a hemocytometer and a total of 14,400 TSCs were added per well. The ESC media (FCLY) was removed and TSCs were added dropwise in TX media + LIF prepared as follows: DMEM/F12 without HEPES (21041-025, Thermo Fisher Scientific), 64 mg l-1 l-ascorbic acid-2-phosphate magnesium (A8960, Sigma-Aldrich), 14 mg l-1 sodium selenite (S5261, Sigma-Aldrich), 19.4 mg l-1 insulin (I9278, Sigma-Aldrich), 543 mg/l NaHCO3 (S5761, Sigma-Aldrich), 10.7 mg l-1 holo-transferrin (T4132, Sigma-Aldrich), 25 ng ml-1 human recombinant FGF4 (5846-F4, R&D systems), 2 ng ml-1 human recombinant TGF-ß1 (100-21, PeproTech), 1 mg ml-1 heparin (H3149, Sigma-Aldrich), 1% GlutaMAX, 1% penicillin/streptomycin, 20 μM Y27632 (72304, Stemcell technologies), 3 μM CHIR99021, 1 mM 8-Bromo-cAMP (73602, STEMCELL Technologies), 30 ng ml−1 IL-11 (200-11, PeproTech), and 10ng ml−1 mouse recombinant LIF. Upon TSC seeding, the plate was placed at 37°C, in hypoxic conditions (5% O_2). The tetO-Gata4 ESCs were induced 24 hours after TSC seeding via supplementation of 1 μg/mL doxycycline (D9891-5G, Sigma-Aldrich).

Criteria for selecting multicellular aggregates structures

Upon completion of any given aggregation experiment (from day 3 to 4), all structures from a given Aggrewell well were collected. Blastoids exhibiting a robust cavity, clear ICM-like inner compartment, and lacking any aberrant growths were selected for downstream analyses. Non-cavitated structures were excluded from further characterization.

IVC Culture of Natural Embryos and Blastocyst-like Structures

Natural embryos and blastoids were cultured on on ibidi-u plates in IVC1 media as previously described (Bedzhov *et al.,* 2014). In brief, mural trophectoderm was cut away from the embryos or blastoids and the remaining structures was placed in IVC1 media. Ibidi-u plates were then placed in 37C incubator in normoxic conditions (20% O_2 or 5% CO_2). Media was changed daily. IVC1 consisted of the following: Advanced DMEM/F12 (12634-010, Thermo Fischer Scientific) supplemented with 20% (vol/vol) heat-inactivated FBS, 1% GlutaMAX, penicillin (25 units/ml)/Streptomycin (25 μg/ml), 1X ITS-X (10 mg/L insulin, 5.5 mg/L transferrin, 0.0067 mg/L sodium selenite, 2 mg/L etholamine; 51500-056, Thermo Fisher Scientific), 8 nM β-estradiol (E8875, Sigma-Aldrich), 200 ng/ml progesterone (P0130, Sigma-Aldrich), 25 μM N-acetyl-L-cysteine (A7250, Sigma-Aldrich).

Immunofluorescence Staining

Natural embryos and stem cell-derived structures were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, 15710) for 20 min at room temperature, washed twice in PBT [phosphate-buffered saline (PBS) plus 0.05% Tween-20] and permeabilized for 30 min at room temperature in 0.3% Triton-X-100, 0.1% glycine. Primary antibody incubation was performed overnight at 4°C in blocking buffer [PBS containing 10% fetal bovine serum (FBS), 1% Tween-20]. The following day, embryos were washed twice in PBT, then incubated overnight with secondary antibody (1:500) in blocking buffer at 4°C. On day 3, embryos were washed twice in PBT transferred to PBT drops in oil-filled optical plates before confocal imaging.

Mouse husbandry

All animal procedures were approved by the Institute Animal Care and Use Committee (IACUC), protocol number 1772, at the California Institute of Technology (Caltech) and performed in accordance with NIH guidelines. Mice were housed in ventilated microisolator cages in a temperature and humidity-controlled environment in an automated light-dark cycle room, with 11-hour dark–13-hour light period.

Mouse embryo collection

E3.5/E4.5: For pre-implantation embryo experiments 4- week-old B6SJLF1/J (The Jackson Lab Strain No:100012) females were injected with 10 IU of pregnant mare's serum gonadotrophin (PMSG) and, 48 hours later, with 10 IU of human chorionic gonadotrophin (hCG). Super-ovulated mice were then mated with B6CBAF1/J (The Jackson Lab Strain No:100011) males. Females were checked for coital plugs the next day. Embryos were recovered at E3.5 by flushing the uterine horns with M2 medium. Embryos were recovered at E4.5 by flushing followed by scraping of the inside of the uterine horns with dissection tweezers.

E4.75: 7-week-old B6SJLF1/J (The Jackson Lab Strain No:100012) females were mated with B6CBAF1/J (The Jackson Lab Strain No:100011) males. Embryos ranging from E4.75-E5.0 were dissected out from the uteri in warmed M2 media (Sigma Catalog No.M7167) and excess desidual tissues were manually removed. Embryos were kept in M2 media within 2 hours in an 37C incubator before the dissociation.

Sample Collection and dissociation for Single Cell RNA Sequencing (scRNA-Seq)

Natural embryo recovery and dissociation: Natural embryos were recovered as described above. Using a 96-well plate, embryos were washed twice in DEPC treated PBS, and then moved to dissociation mix: Ebioscience Accumax Cell Aggregate Dissociation Medium (00-4666-56, Thermo Scientific) supplemented with 1:1000 DNase to remove aberrant DNA released from lysed cells. Embryos were left in dissociation media at room temperature for 25-60 minutes with trituration by mouth pipetting and glass capillaries every 10 minutes. The dissociation reaction was quenched using equal parts FC media. Cells were passed through a 20 μM filter and spun down at 450 rcf. Cells were then resuspended in 0.04% BSA-PBS according to the 10X Chromium Next GEM single cell 3' Kit v3.1 user guide.

Blastoid recovery and dissociation: Blastoids were treated nearly identical to natural embryos, however took slightly longer to dissociate. In brief, blastoids with proper blastocyst-like morphology (structures containing a cavity, visible inner cell mass, and outer epithelial-like layer) were selected by mouth pipetting. These structures were washed twice in DEPC-PBS before moving them to Accumax:DNAse. Blastoids were treated for 60 minutes with trituration every 10 minutes. The dissociation reaction was quenched using equal parts FC media. Cells were passed through a 20 μM filter and spun down at 450 rcf. Cells were then resuspended in 0.04% BSA-PBS according to the 10X Chromium Next GEM single cell 3' Kit v3.1 user guide.

Chapter 5

CONCLUSION AND FUTURE DIRECTIONS

5.1. Limitations of natural embryo research and the power of embryoids

The study of natural mouse embryogenesis has elucidated many aspects of mammalian development, from lineage segregation, to axis patterning, to gastrulation, to organogenesis, and more. Mouse systems have also been monumental in understanding our own development and have paved the way for many of the exploratory studies done on human embryos to date. Nevertheless, countless questions remain surrounding aspects of early development, in both mouse and human, and this is for several reasons. First, and most importantly, access to natural embryos is incredibly difficult, especially in a human context. In addition to limited access, many ethical boundaries are put in place to ensure practices remain as humane and principled as possible. One such example is the "14-day rule" which states that human embryos cannot be cultured beyond 14 days, or beyond formation of the primitive streak. While even limited access to embryos can provide a wealth of knowledge to the field, researchers are severely lacking in statistical power, reproducibility, and easy manipulability. Another major caveat in the field is that *ex utero* culture methods are not perfect. Certainly, incredible advances have been made, and much knowledge has been gained with the emergence of this technology. Even so, future work to improve these conditions will be necessary not only to increase reproducibility, but also to reliably develop embryos to later stages. Unfortunately, the experiments required to optimize these conditions would require significant numbers of embryos, which is again technically and ethically restrained.

Given these hurdles, stem cell-derived embryoids represent exciting tools to greatly alleviate many of the aforementioned burdens. Firstly, the derivation and sustained cultivation of stem cells greatly minimizes the need for embryos as such stem cells can give rise to infinite embryoid structures. Introducing genetic mutations into stem cell lines is also significantly

easier than producing knockout or overexpression mouse lines. This is in part because such experiments in mice are technically challenging, but also because many of these adaptations are lethal. By developing these perturbations in cell lines, we could also perform lineage specific alterations or transient knockdowns/over-expressions; in other words, we greatly improve the manipulability of these systems. Additionally, genetic alteration of human embryos is incredibly controversial; using human stem cells and human embryoid systems, many mechanistic experiments looking at targeted gene perturbations could be achieved, avoiding the need for human embryos. Another important benefit of stem cell derived embryoids is the ability to generate hundreds if not thousands of these structures at a time—something that is just not possible in a natural context. This ability could lead to high-throughput screens testing the effects of certain chemicals/molecules on early development. This is an important innovation as pregnant women are, in large part, excluded from clinical trials. While this is of course understandable, it severely limits our knowledge of medical treatments available to pregnant women and could lead to life-threatening situations for both mother and child. High-throughput analysis would also greatly increase the statistical power of these studies, solidifying experimental results and observations.

Altogether, it is clear that embryoid systems could revolutionize early embryo research. That being said, I believe that several important improvements must happen to these systems in order to achieve the goals listed above. I will outline these in the coming sections. Although several stem cell-derived embryo models have emerged within the last decade, I will focus solely on human and mouse blastoids, which form the bulk of my PhD research.

5.1. Human Blastoids

At the onset of my PhD, there was not a single published protocol for human blastoids; in the last 2 years, however, six distinct protocols have been developed[123,132,134,147,156,157]. In 2021, the first protocols published[123,132,134] demonstrated the ability for human pluripotent stem cells (hPSCs) to generate structures that mimic natural human blastocysts in size, morphology, and initiation of lineage segregation, but with relatively low efficiency. Moreover, there is considerable debate as to whether or not the lineages generated in these

system are truly recapitulating those of the natural blastocyst[123,142]. In our own EP-structures, we found that while certain lineages could be initiated, as evidenced by scRNAseq and immunostaining, full commitment to these lineages was not achieved. In particular we saw overcommitment to the hypoblast-like lineage, and minimal commitment to the trophectoderm-like lineage. Several explanations could be responsible for this observation. First, EPSCs have clearly been shown to mimic late stage epiblast as opposed to the originally believed pre-implantation morula, meaning extra-embryonic lineage potential would be very low in these cells. Furthermore, the fact that we can see upregulation of certain lineage markers by immunostaining, which is less convincingly observed in the scRNAseq data, indicates that there might be incongruencies with mRNA production and protein translation. Similar failure to commit to lineages were observed in other systems as well. In the so-called "iBlastoids", they found two "intermediate cell types" which show minimal resemblance to lineages of the natural embryo at this stage[132,142]. Further analysis went on to show that the putative trophectoderm lineage in these structures bore more resemblance to the amnion, rather than the trophectoderm[142].

In the models to follow, several key improvements were made to these systems that not only enhanced lineage efficacy, but also significantly improved structure formation efficiency. Firstly, it remains clear that "naïve pluripotent stem cells" generated from either the 5iLA or PXGL protocols (as opposed to EPSCs or iPSCs) demonstrate much higher reprogramming ability to faithfully capture pre-implantation lineages. One publication also showed that introduction of ERK, ROCK, and NODAL inhibitors greatly improved structure formation and generated higher transcriptional fidelity of lineages compared to the natural embryo[147]. Inhibition of these ERK and NODAL pathways have been implicated in trophectoderm differentiation; as many of the previous models show deficient trophectoderm formation, it is unsurprising that addition of these molecules leads to higher cavitation and structure formation. Finally, Kagawa *et al.* expanded on this optimized system further by introducing a HIPPO pathway inhibitor and LIF to their culture media. HIPPO inhibition, downstream of apico-basal polarity formation, triggers differentiation towards the trophectoderm fate in natural embryos[23,24]. Appropriately, apical polarity and subsequent HIPPO inhibition was

also required for trophectoderm lineage speciation in blastoids and demonstrates that this model can accurately recapitulate key lineage decisions compared to the natural blastocyst.

While each of these models are incredible, demonstrating the remarkable ability for human PSCs to self-organize into blastocyst-like structures, several improvements must be made before such techniques can be widely adopted or used for high-throughput screens. For example, it has become routine to use naïve induction protocols on human ESCs as the starting cell populations for these experiments. Such conversion protocols remain very difficult to reproduce and show significant cell line dependencies. Before we can use these cells to reliably generate embryoid structures, we must first better understand the transition between naïve and primed pluripotency, and perhaps more importantly the converse, primed to naïve reprogramming. It will also be interesting to understand what role genetic background, gender, and epigenetic modifications play on reprogramming/differentiation. To begin to investigate these questions, however, derivation of additional human ESC lines will be necessary, as there are currently very few lines established. This is certainly no easy endeavor but will be a critical step towards better understanding the pluripotent state in humans and ultimately better culturing conditions for the future. Another important and exciting area worthy of further investigation is continued comparison between the various models and the natural embryo. It amazes me that each of these protocols can use different starting cell populations, different media conditions, different protocols, and yet still generate structures that are morphologically very similar. Using these models, we could begin to understand the core principles of self-organization necessary to form the blastocyst by identifying the "common denominator" pathways present in each of these systems. As these models develop, it will be important to continue to compare these systems to the natural embryo, not only to see where they are similar, but also where they are different. Finally, nearly all of these models demonstrate the ability to initiate post-implantation remodeling, yet these results remain deficient and incomplete. By performing the aforementioned studies, there is hope that these protocols can be improved upon and one day blastoids could indeed be able to reliably transition from pre- to post-implantation stages. Once this is accomplished, these structures could be used for countless applications. Perhaps the most exciting

application would be pre- to post-implantation culture assays that would elucidate mechanisms of the most precarious stage of development: implantation.

Nevertheless, while these human models hold incredible potential, there is a major caveat. When developing *in vitro* models, it is important to have a natural *in vivo* embryo to compare them to. Under current guidelines, and with current technical limitations, having post-implantation human embryos to compare these models to will be difficult, if not impossible. For this reason, the generation of pre-implantation stem cell models in mouse with post-implantation capabilities in just one alternative to overcome this issue.

5.2. Mouse Blastoids

While the field of human stem cell-derived embryos seems to be expanding and evolving at an exponential rate, mouse models remain of critical importance. This is particularly true for the case of studying implantation. Firstly, studying human implantation will continue to be difficult both for ethical and technical limitations. Although it is possible to perform *in vitro* assays that investigate human blastoid implantation, such models have limited *in vivo* examples. Without the natural counterpart, it becomes difficult to truly assess the efficacy of the stem cell models. A key advantage of mouse blastoids is the ability to perform *in vivo* transfer experiments and recapitulate events of development within the womb of the mother. Hypothetically, such experiments could lead to the development of a stem cell derived embryos that are capable of developing beyond E8.5, a feat that as of now remains impossible *ex utero* for both natural and stem cell-derived embryos. Nevertheless, since the first mouse blastoid model was published in 2019, minimal progress has been made toward development beyond pre-implantation stages. The major advantages that have arisen with newer generation models predominantly focuses on increased cavitation efficiency, and significant improvements to lineage formation. Why current models fail to implant remains unclear, and further investigation will be necessary to elucidate this process.

Another inherent benefit to mouse blastoid models is the ability to more easily perform genetic manipulations and testing. For example, once blastoid models become robust and

successfully transition beyond implantation, one could perform overexpression/knock-out experiments, potentially even screens, to see how such perturbations affect development. These experiments are theoretically possible in human blastoids as well, but as these models advance, I anticipate more severe regulation will be put into place. Moreover, mouse models again permit natural *in vivo* experiments to be run in parallel in order to validate results seen *in vitro.*

Though the potential of this field is immense, major limitations must be overcome. The first of these being reproducibility. Although mouse blastoids can be generated at high efficiency, with representation of all three lineages, these efficiencies vary greatly depending on the cell lines used. The reason behind such cell line dependencies, however, are unknown. One possibility is that each derivation produces cell lines of slightly different developmental potential/pluripotency state. Additionally, this likely corresponds to different epigenetic states, which enables or prohibits downstream development. Moving forward, multiomic characterization of starting cell populations and assays that assess their subsequent blastoid formation efficiency would be useful to better understand the ground state that is permissive of embryoid formation. This would not only lead to the generation of more robust models, but would give insight into the various pluripotent states that characterize early development.

A second limitation that hinders the progress of these models is the inability to perform high throughput screening on 3D models. To overcome this, development of new culturing methods that permit live imaging and tracking is necessary. One first step would be the transition away from plastic pyramidal Aggrewells, to perhaps agarose or PEG culture wells. While these methods do exist, efficiency in comparison to Aggrewells is lacking. Furthermore, successful generation of blastoids is tied to having a precise ratio of each cell type, yet current seeding methods relies on random Poisson distribution. By employing a method that more reliably seeds a precise number of each cell type, one could likely improve the efficiency of these models significantly. For example, employment of microfluidic devices could perhaps be one way to overcome this dilemma.

Taken together, mouse blastoids demonstrate remarkable potential to elucidate many events of early development. The versatility and easy manipulability of these models would allow for high-throughput analyses, pending technological advances. Importantly, unlike human blastoids, these models are less limited by ethical restriction and can be easily compared to natural embryos to ensure efficacy. While advancements have been made over the last few years, the ability for blastoids to seamlessly transition from pre- to post-implantation development will remain the a necessary breakthrough for the relevance of these models for human health.

5.3. Closing thoughts

Embryogenesis is a fascinating yet complex process. While current studies in the natural embryo have revolutionized our understanding of early development, major technical and ethical limitations hinder progress forward. With the emergence and optimization of stem cell-derived embryo models, researchers can finally overcome these hurdles and dive even deeper into the mechanisms controlling early cell fate decisions. Lastly, pre-implantation blastoid models are a particularly exciting system and could be influential tools for understanding implantation and the critical fail-mechanisms leading to pregnancy arrest.

BIBLIOGRAPHY

1.Macklon, N. S., Geraedts, J. P. M. & Fauser, B. C. J. M. Conception to ongoing pregnancy: the ‘black box’ of early pregnancy loss. *Human Reproduction Update* **8**, 333–343 (2002).

2.Larsen, E. C., Christiansen, O. B., Kolte, A. M. & Macklon, N. New insights into mechanisms behind miscarriage. *BMC Med* **11**, 154 (2013).

3.Morris, S. A. *et al.* Dynamics of anterior–posterior axis formation in the developing mouse embryo. *Nat Commun* **3**, 673 (2012).

4.Bedzhov, I. & Zernicka-Goetz, M. Self-Organizing Properties of Mouse Pluripotent Cells Initiate Morphogenesis upon Implantation. *Cell* **156**, 1032–1044 (2014).

5.Aguilera-Castrejon, A. *et al.* Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis. *Nature* **593**, 119–124 (2021).

6.Shahbazi, M. N. *et al.* Self-organization of the human embryo in the absence of maternal tissues. *Nat Cell Biol* **18**, 700–708 (2016).

7.Xiang, L. *et al.* A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* **577**, 537–542 (2020).

8.Deglincerti, A. *et al.* Self-organization of the in vitro attached human embryo. *Nature* **533**, 251–254 (2016).

9.Zhou, L. & Dean, J. Reprogramming the genome to totipotency in mouse embryos. *Trends Cell Biol* **25**, 82–91 (2015).

10.Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033–3042 (2009).

11.Schulz, K. N. & Harrison, M. M. Mechanisms regulating zygotic genome activation. *Nat Rev Genet* **20**, 221–234 (2019).

12.Braude, P., Bolton, V. & Moore, S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* **332**, 459–461 (1988).

13.Piotrowska, K., Wianny, F., Pedersen, R. A. & Zernicka-Goetz, M. Blastomeres arising from the first cleavage division have distinguishable fates in normal mouse development. *Development* **128**, 3739–3748 (2001).

14.Piotrowska, K. & Zernicka-Goetz, M. Role for sperm in spatial patterning of the early mouse embryo. *Nature* **409**, 517–521 (2001).

15.Gardner, R. L. Specification of embryonic axes begins before cleavage in normal mouse development. *Development* **128**, 839–847 (2001).

16.Rosàs-Canyelles, E., Modzelewski, A. J., Geldert, A., He, L. & Herr, A. E. Assessing heterogeneity among single embryos and single blastomeres using open microfluidic design. *Sci Adv* **6**, eaay1751 (2020).

17.Torres-Padilla, M.-E., Parfitt, D.-E., Kouzarides, T. & Zernicka-Goetz, M. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**, 214–218 (2007).

18.Goolam, M. *et al.* Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell* **165**, 61–74 (2016).

19.Chen, Q., Shi, J., Tao, Y. & Zernicka-Goetz, M. Tracing the origin of heterogeneity and symmetry breaking in the early mammalian embryo. *Nat Commun* **9**, 1819 (2018).

20.Piras, V., Tomita, M. & Selvarajoo, K. Transcriptome-wide Variability in Single Embryonic Development Cells. *Sci Rep* **4**, 7137 (2014).

21.Fleming, T. P., Cannon, P. M. & Pickering, S. J. The cytoskeleton, endocytosis and cell polarity in the mouse preimplantation embryo. *Developmental Biology* **113**, 406–419 (1986).

22.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).

23.Zhu, M. *et al.* Human embryo polarization requires PLC signaling to mediate trophectoderm specification. *eLife* **10**, e65068 (2021).

24.Gerri, C. *et al.* Initiation of a conserved trophectoderm program in human, cow and mouse embryos. *Nature* **587**, 443–447 (2020).

25.Watanabe, Y. *et al.* Notch and Hippo signaling converge on Strawberry Notch 1 (Sbno1) to synergistically activate Cdx2 during specification of the trophectoderm. *Sci Rep* **7**, 46135 (2017).

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

27.White, M. D., Zenker, J., Bissiere, S. & Plachta, N. Instructions for Assembling the Early Mammalian Embryo. *Developmental Cell* **45**, 667–679 (2018).

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

29.Chazaud, C., Yamanaka, Y., Pawson, T. & Rossant, J. Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway. *Developmental Cell* **10**, 615–624 (2006).

30.Wamaitha, S. E. & Niakan, K. K. Human Pre-gastrulation Development. *Curr Top Dev Biol* **128**, 295–338 (2018).

31.Morris, S. A. *et al.* Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. *Proc Natl Acad Sci U S A* **107**, 6364–6369 (2010).

32.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).

33.Roode, M. *et al.* Human hypoblast formation is not dependent on FGF signalling. *Developmental Biology* **361**, 358–363 (2012).

34.Boroviak, T. *et al.* Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. *Development* **145**, dev167833 (2018).

35.Gardner, R. L., Papaioannou, V. E. & Barton, S. C. Origin of the ectoplacental cone and secondary giant cells in mouse blastocysts reconstituted from isolated trophoblast and inner cell mass. *J Embryol Exp Morphol* **30**, 561–572 (1973).

36.Frias-Aldeguer, J. *et al.* Embryonic signals perpetuate polar-like trophoblast stem cells and pattern the blastocyst axis. 510362 Preprint at https://doi.org/10.1101/510362 (2020).

37.Christodoulou, N. *et al.* Morphogenesis of extra-embryonic tissues directs the remodelling of the mouse embryo at implantation. *Nat Commun* **10**, 3557 (2019).

38.Suzuki, D., Sasaki, K., Kumamoto, S., Tanaka, K. & Ogawa, H. Dynamic Changes of Gene Expression in Mouse Mural Trophectoderm Regulated by Cdx2 During Implantation. *Front Cell Dev Biol* **10**, 945241 (2022).

39.Gardner, R. L. Flow of cells from polar to mural trophectoderm is polarized in the mouse blastocyst. *Human Reproduction* **15**, 694–701 (2000).

40.Aplin, J. D. & Ruane, P. T. Embryo–epithelium interactions during implantation at a glance. *Journal of Cell Science* **130**, 15–22 (2017).

41.Lindenberg, S., Hyttel, P., Sjøgren, A. & Greve, T. A comparative study of attachment of human, bovine and mouse blastocysts to uterine epithelial monolayer. *Hum Reprod* **4**, 446–456 (1989).

42.Su, R.-W. & Fazleabas, A. T. Implantation and Establishment of Pregnancy in Human and Nonhuman Primates. *Adv Anat Embryol Cell Biol* **216**, 189–213 (2015).

43.Carter, A. M., Enders, A. C. & Pijnenborg, R. The role of invasive trophoblast in implantation and placentation of primates. *Philos Trans R Soc Lond B Biol Sci* **370**, 20140070 (2015).

44.Woods, L., Perez-Garcia, V. & Hemberger, M. Regulation of Placental Development and Its Impact on Fetal Growth—New Insights From Mouse Models. *Frontiers in Endocrinology* **9**, (2018).

45.Nichols, J. & Smith, A. Naive and Primed Pluripotent States. *Cell Stem Cell* **4**, 487–492 (2009).

46.Bradley, A., Evans, M., Kaufman, M. H. & Robertson, E. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255–256 (1984).

47.Heard, E. Recent advances in X-chromosome inactivation. *Current Opinion in Cell Biology* **16**, 247–255 (2004).

48.Guo, G. *et al.* Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063–1069 (2009).

49.Mello, J. C. M. de *et al.* Random X Inactivation and Extensive Mosaicism in Human Placenta Revealed by Analysis of Allele-Specific Gene Expression along the X Chromosome. *PLOS ONE* **5**, e10947 (2010).

50.Petropoulos, S. *et al.* Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165**, 1012–1026 (2016).

51.Deng, X., Berletch, J. B., Nguyen, D. K. & Disteche, C. M. X chromosome regulation: diverse patterns in development, tissues and disease. *Nat Rev Genet* **15**, 367–378 (2014).

52.Okamoto, I. *et al.* Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* **472**, 370–374 (2011).

53.Bedzhov, I. & Zernicka-Goetz, M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032–1044 (2014).

54.Shahbazi, M. N. *et al.* Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* **552**, 239–243 (2017).

55.Luckett, W. P. The development of primordial and definitive amniotic cavities in early rhesus monkey and human embryos. *American Journal of Anatomy* **144**, 149–167 (1975).

56.Yang, R. *et al.* Amnion signals are essential for mesoderm formation in primates. *Nat Commun* **12**, 5126 (2021).

57.Sasaki, K. *et al.* The Germ Cell Fate of Cynomolgus Monkeys Is Specified in the Nascent Amnion. *Dev Cell* **39**, 169–185 (2016).

58.Matsuo, I., Kimura-Yoshida, C. & Ueda, Y. Developmental and mechanical roles of Reichert's membrane in mouse embryos. *Philosophical Transactions of the Royal Society B: Biological Sciences* **377**, 20210257 (2022).

59.Kwon, G. S., Viotti, M. & Hadjantonakis, A.-K. The Endoderm of the Mouse Embryo Arises by Dynamic Widespread Intercalation of Embryonic and Extraembryonic Lineages. *Developmental Cell* **15**, 509–520 (2008).

60.Viotti, M., Foley, A. C. & Hadjantonakis, A.-K. Gutsy moves in mice: cellular and molecular dynamics of endoderm morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**, 20130547 (2014).

61.Wu, Q. *et al.* Sall4 Interacts with Nanog and Co-occupies Nanog Genomic Sites in Embryonic Stem Cells*. *Journal of Biological Chemistry* **281**, 24090–24094 (2006).

62.Okumura-Nakanishi, S., Saito, M., Niwa, H. & Ishikawa, F. Oct-3/4 and Sox2 Regulate Oct-3/4 Gene in Embryonic Stem Cells*. *Journal of Biological Chemistry* **280**, 5307–5317 (2005).

63.Chew, J.-L. *et al.* Reciprocal Transcriptional Regulation of Pou5f1 and Sox2 via the Oct4/Sox2 Complex in Embryonic Stem Cells. *Molecular and Cellular Biology* **25**, 6031–6046 (2005).

64.Sun, L. T. *et al.* Nanog co-regulated by Nodal/Smad2 and Oct4 is required for pluripotency in developing mouse epiblast. *Developmental Biology* **392**, 182–192 (2014).

65.Arnold, S. J. & Robertson, E. J. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol* **10**, 91–103 (2009).

66.Srinivas, S., Rodriguez, T., Clements, M., Smith, J. C. & Beddington, R. S. P. Active cell migration drives the unilateral movements of the anterior visceral endoderm. *Development* **131**, 1157–1164 (2004).

67.Beddington, R. S. P. & Robertson, E. J. Axis Development and Early Asymmetry in Mammals. *Cell* **96**, 195–209 (1999).

68.Lu, C. C., Brennan, J. & Robertson, E. J. From fertilization to gastrulation: axis formation in the mouse embryo. *Current Opinion in Genetics & Development* **11**, 384–392 (2001).

69.Takaoka, K. *et al.* The mouse embryo autonomously acquires anterior-posterior polarity at implantation. *Dev Cell* **10**, 451–459 (2006).

70.Torres-Padilla, M.-E. *et al.* The anterior visceral endoderm of the mouse embryo is established from both preimplantation precursor cells and by de novo gene expression after implantation. *Dev Biol* **309**, 97–112 (2007).

71.Rossant, J. & Tam, P. P. L. Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701–713 (2009).
72.Molè, M. A. *et al.* A single cell characterisation of human embryogenesis identifies pluripotency transitions and putative anterior hypoblast centre. *Nat Commun* **12**, 3679 (2021).
73.Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
74.Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**, 7634–7638 (1981).
75.Silva, J. *et al.* Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol* **6**, e253 (2008).
76.Ying, Q.-L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
77.Tanaka, S., Kunath, T., Hadjantonakis, A. K., Nagy, A. & Rossant, J. Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282**, 2072–2075 (1998).
78.Niakan, K. K., Schrode, N., Cho, L. T. Y. & Hadjantonakis, A.-K. Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. *Nat Protoc* **8**, 1028–1041 (2013).
79.Fujikura, J. *et al.* Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* **16**, 784–789 (2002).
80.Harrison, S. E., Sozen, B., Christodoulou, N., Kyprianou, C. & Zernicka-Goetz, M. Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. *Science* **356**, eaal1810 (2017).
81.Sozen, B. *et al.* Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. *Nat Cell Biol* **20**, 979–989 (2018).
82.Amadei, G. *et al.* Inducible Stem-Cell-Derived Embryos Capture Mouse Morphogenetic Events In Vitro. *Developmental Cell* **56**, 366-382.e9 (2021).
83.Amadei, G. *et al.* Embryo model completes gastrulation to neurulation and organogenesis. *Nature* **610**, 143–153 (2022).

84.Tarazi, S. *et al.* Post-gastrulation synthetic embryos generated ex utero from mouse naive ESCs. *Cell* **185**, 3290-3306.e25 (2022).

85.Lau, K. Y. C. *et al.* Mouse embryo model derived exclusively from embryonic stem cells undergoes neurulation and heart development. *Cell Stem Cell* **29**, 1445-1458.e8 (2022).

86.Rivron, N. C. *et al.* Blastocyst-like structures generated solely from stem cells. *Nature* **557**, 106–111 (2018).

87.Li, R. *et al.* Generation of Blastocyst-like Structures from Mouse Embryonic and Adult Cell Cultures. *Cell* **179**, 687-702.e18 (2019).

88.Sozen, B. *et al.* Self-Organization of Mouse Stem Cells into an Extended Potential Blastoid. *Developmental Cell* **51**, 698-712.e8 (2019).

89.Posfai, E. *et al.* Evaluating totipotency using criteria of increasing stringency. *Nat Cell Biol* **23**, 49–60 (2021).

90.Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).

91.Okae, H. *et al.* Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* **22**, 50-63.e6 (2018).

92.Mischler, A. *et al.* Two distinct trophectoderm lineage stem cells from human pluripotent stem cells. *Journal of Biological Chemistry* **296**, (2021).

93.Dong, C. *et al.* Derivation of trophoblast stem cells from naïve human pluripotent stem cells. *eLife* **9**, e52504 (2020).

94.Io, S. *et al.* Capturing Human Trophoblast Development with Naïve Pluripotent Stem Cells In Vitro. 2020.12.17.416800 Preprint at https://doi.org/10.1101/2020.12.17.416800 (2020).

95.Li, Y. *et al.* BMP4-directed trophoblast differentiation of human embryonic stem cells is mediated through a ΔNp63+ cytotrophoblast stem cell state. *Development* **140**, 3965–3976 (2013).

96.Guo, G. *et al.* Trophectoderm Potency is Retained Exclusively in Human Naïve Cells. 2020.02.04.933812 Preprint at https://doi.org/10.1101/2020.02.04.933812 (2020).

97.Wamaitha, S. E. *et al.* Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes Dev.* **29**, 1239–1255 (2015).

98.Séguin, C. A., Draper, J. S., Nagy, A. & Rossant, J. Establishment of Endoderm Progenitors by SOX Transcription Factor Expression in Human Embryonic Stem Cells. *Cell Stem Cell* **3**, 182–195 (2008).

99.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).

100.Hanna, J. *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proceedings of the National Academy of Sciences* **107**, 9222–9227 (2010).

101.Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286 (2013).

102.Chan, Y.-S. *et al.* Induction of a Human Pluripotent State with Distinct Regulatory Circuitry that Resembles Preimplantation Epiblast. *Cell Stem Cell* **13**, 663–675 (2013).

103.Ware, C. B. *et al.* Derivation of naïve human embryonic stem cells. *Proceedings of the National Academy of Sciences* **111**, 4484–4489 (2014).

104.Theunissen, T. W. *et al.* Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).

105.Qin, H. *et al.* YAP Induces Human Naive Pluripotency. *Cell Reports* **14**, 2301–2312 (2016).

106.Takashima, Y. *et al.* Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell* **158**, 1254–1269 (2014).

107.Guo, G. *et al.* Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports* **6**, 437–446 (2016).

108.Yang, Y. *et al.* Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. *Cell* **169**, 243-257.e25 (2017).

109.Guo, G. *et al.* Epigenetic resetting of human pluripotency. *Development* **144**, 2748–2763 (2017).
110.Gao, X. *et al.* Establishment of porcine and human expanded potential stem cells. *Nat Cell Biol* **21**, 687–699 (2019).
111.Bayerl, J. *et al.* Tripartite Inhibition of SRC-WNT-PKC Signalling Consolidates Human Naïve Pluripotency. 2020.05.23.112433 Preprint at https://doi.org/10.1101/2020.05.23.112433 (2020).
112.Liu, X. *et al.* Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat Methods* **14**, 1055–1062 (2017).
113.Allegrucci, C. & Young, L. E. Differences between human embryonic stem cell lines. *Human Reproduction Update* **13**, 103–120 (2007).
114.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).
115.Brons, I. G. M. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
116.Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
117.Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat Methods* **11**, 847–854 (2014).
118.Martyn, I., Kanno, T. Y., Ruzo, A., Siggia, E. D. & Brivanlou, A. H. Self-organization of a human organizer by combined Wnt and Nodal signalling. *Nature* **558**, 132–135 (2018).
119.X, X. *et al.* Mechanics-guided embryonic patterning of neuroectoderm tissue from human pluripotent stem cells. *Nature materials* **17**, (2018).
120.Haremaki, T. *et al.* Self-organizing neuruloids model developmental aspects of Huntington's disease in the ectodermal compartment. *Nat Biotechnol* **37**, 1198–1208 (2019).
121.Simunovic, M. *et al.* A 3D model of a human epiblast reveals BMP4-driven symmetry breaking. *Nat Cell Biol* **21**, 900–910 (2019).

122.Zheng, Y. *et al.* Controlled modelling of human epiblast and amnion development using stem cells. *Nature* **573**, 421–425 (2019).

123.Sozen, B. *et al.* Reconstructing aspects of human embryogenesis with pluripotent stem cells. *Nat Commun* **12**, 5550 (2021).

124.Shahbazi, M. N. & Zernicka-Goetz, M. Deconstructing and reconstructing the mouse and human early embryo. *Nat Cell Biol* **20**, 878–887 (2018).

125.Pera, M. F. Human embryo research and the 14-day rule. *Development* **144**, 1923–1925 (2017).

126.Beccari, L. *et al.* Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* **562**, 272–276 (2018).

127.Kime, C. *et al.* Induced 2C Expression and Implantation-Competent Blastocyst-like Cysts from Primed Pluripotent Stem Cells. *Stem Cell Reports* **13**, 485–498 (2019).

128.van den Brink, S. C. *et al.* Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* **141**, 4231–4242 (2014).

129.van den Brink, S. C. *et al.* Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* **582**, 405–409 (2020).

130.Veenvliet, J. V. *et al.* Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science* **370**, eaba4937 (2020).

131.Kuchroo, M., Godavarthi, A., Tong, A., Wolf, G. & Krishnaswamy, S. Multimodal Data Visualization and Denoising with Integrated Diffusion. Preprint at https://doi.org/10.48550/arXiv.2102.06757 (2022).

132.Liu, X. *et al.* Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* **591**, 627–632 (2021).

133.Moris, N. *et al.* An in vitro model of early anteroposterior organization during human development. *Nature* **582**, 410–415 (2020).

134.Yu, L. *et al.* Blastocyst-like structures generated from human pluripotent stem cells. *Nature* **591**, 620–626 (2021).

135.Tan, T. *et al.* Chimeric contribution of human extended pluripotent stem cells to monkey embryos ex vivo. *Cell* **184**, 2020-2032.e14 (2021).

136.Yang, J. *et al.* Establishment of mouse expanded potential stem cells. *Nature* **550**, 393–397 (2017).

137.Nastri, C. O. *et al.* Low versus atmospheric oxygen tension for embryo culture in assisted reproduction: a systematic review and meta-analysis. *Fertility and Sterility* **106**, 95-104.e17 (2016).

138.Richter, K. S., Harris, D. C., Daneshmand, S. T. & Shapiro, B. S. Quantitative grading of a human blastocyst: optimal inner cell mass size and shape. *Fertility and Sterility* **76**, 1157–1167 (2001).

139.Johnson, M. H. & Ziomek, C. A. The foundation of two distinct cell lineages within the mouse morula. *Cell* **24**, 71–80 (1981).

140.Blakeley, P. *et al.* Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3613 (2015).

141.Massey, J. *et al.* Synergy with TGFβ ligands switches WNT pathway dynamics from transient to sustained during human pluripotent cell differentiation. *Proceedings of the National Academy of Sciences* **116**, 4989–4998 (2019).

142.Zhao, C. *et al.* Reprogrammed blastoids contain amnion-like cells but not trophectoderm. 2021.05.07.442980 Preprint at https://doi.org/10.1101/2021.05.07.442980 (2021).

143.Riveiro, A. R. & Brickman, J. M. From pluripotency to totipotency: an experimentalist's guide to cellular potency. *Development* **147**, dev189845 (2020).

144.Theunissen, T. W. *et al.* Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* **19**, 502–515 (2016).

145.Guo, G. *et al.* Human naive epiblast cells possess unrestricted lineage potential. *Cell Stem Cell* **28**, 1040-1056.e6 (2021).

146.Yang, Y. *et al.* Derivation of Pluripotent Stem Cells with In Vivo Embryonic and Extraembryonic Potency. *Cell* **169**, 243-257.e25 (2017).

147.Yanagida, A. *et al.* Naive stem cell blastocyst model captures human embryo lineage segregation. *Cell Stem Cell* **28**, 1016-1022.e4 (2021).

148.Lehtonen, E., Lehto, V. P., Paasivuo, R. & Virtanen, I. Parietal and visceral endoderm differ in their expression of intermediate filaments. *EMBO J* **2**, 1023–1028 (1983).

149.Paca, A. *et al.* BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. *Developmental Biology* **361**, 90–102 (2012).

150.Seong, J. *et al.* Epiblast inducers capture mouse trophectoderm stem cells in vitro and pattern blastoids for implantation in utero. *Cell Stem Cell* **29**, 1102-1118.e8 (2022).

151.Amadei, G. *et al.* Inducible Stem-Cell-Derived Embryos Capture Mouse Morphogenetic Events In Vitro. *Dev Cell* **56**, 366-382.e9 (2021).

152.Kubaczka, C. *et al.* Derivation and maintenance of murine trophoblast stem cells under defined conditions. *Stem Cell Reports* **2**, 232–242 (2014).

153.Simon, M. C. & Keith, B. The role of oxygen availability in embryonic development and stem cell function. *Nature reviews. Molecular cell biology* **9**, 285 (2008).

154.Orietti, L. C. *et al.* Embryo size regulates the timing and mechanism of pluripotent tissue morphogenesis. *Stem Cell Reports* **16**, 1182–1196 (2021).

155.Bedzhov, I., Leung, C. Y., Bialecka, M. & Zernicka-Goetz, M. In vitro culture of mouse blastocysts beyond the implantation stages. *Nat Protoc* **9**, 2732–2739 (2014).

156.Fan, Y. *et al.* Generation of human blastocyst-like structures from pluripotent stem cells. *Cell Discov* **7**, 1–14 (2021).

157.Kagawa, H. *et al.* Human blastoids model blastocyst development and implantation. *Nature* **601**, 600–605 (2022).