Modelling Human Post-Implantation Development via Extra-Embryonic Niche Engineering

Joshua Hislop[1-3], Amir Alavi[4,5], Qi Song[4,5], Rayna Schoenberger[1,3], Kamyar Keshavarz F.[1-3], Ryan LeGraw[2,3], Jeremy Velazquez [2,3], Tahere Mokhtari[1-3], Mohammad Nasser Taheri[1-3], Matthew Rytel[2,3], Susana M Chuva de Sousa Lopes[8], Simon Watkins[6,7], Donna Stolz[6,7], Samira Kiani [1-3,9], Berna Sozen[10], Ziv Bar-Joseph[4,5], Mo R. Ebrahimkhani[1-3,9 *]

Affiliations:

1. Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA
2. Department of Pathology, Division of Experimental Pathology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15213, USA
3. Pittsburgh Liver Research Center, University of Pittsburgh, Pittsburgh, PA 15261, USA
4. Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
5. Machine Learning Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA.
6. Center for Biologic Imaging, University of Pittsburgh, Pittsburgh, PA, USA.
7. Department of Cell Biology and Molecular Physiology, University of Pittsburgh, Pittsburgh, PA, USA
8. Department of Anatomy and Embryology, Leiden University Medical Center, Einthovenweg, 2333 ZC Leiden, the Netherlands
9. McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA
10. Department of Genetics, Yale School of Medicine, Yale University, New Haven, CT, 06510, USA.

*Correspondence: mo.ebr@pitt.edu

Abstract

Implantation of the human embryo commences a critical developmental stage that comprises profound morphogenetic alteration of embryonic and extra-embryonic tissues, axis formation, and gastrulation events. Our mechanistic knowledge of this window of human life remains limited due to restricted access to *in vivo* samples for both technical and ethical reasons. Additionally, human stem cell models of early post-implantation development with both embryonic and extra-embryonic tissue morphogenesis are lacking. Here, we present iDiscoid, produced from human induced pluripotent stem cells via an engineered a synthetic gene circuit. iDiscoids exhibit reciprocal co-development of human embryonic tissue and engineered extra-embryonic niche in a model of human post-implantation. They exhibit unanticipated self-organization and tissue boundary formation that recapitulates yolk sac-like tissue specification with extra-embryonic

mesoderm and hematopoietic characteristics, the formation of bilaminar disc-like embryonic morphology, the development of an amniotic-like cavity, and acquisition of an anterior-like hypoblast pole and posterior-like axis. iDiscoids offer an easy-to-use, high-throughput, reproducible, and scalable platform to probe multifaceted aspects of human early post-implantation development. Thus, they have the potential to provide a tractable human model for drug testing, developmental toxicology, and disease modeling.

Main Text

Decoding mechanisms of human embryogenesis has tremendous biomedical impact, from treating congenital diseases and infertility, to engineering functional human organs. Immediately after implantation, the embryo and co-developing extra-embryonic tissues are profoundly remodeled and initiate morphological changes central to the success of pregnancy, including formation of the amniotic cavity and emergence of yolk sac hematopoiesis[1,2]. However, hidden within the uterine tissues, the key steps of early post-implantation development in humans largely remain beyond our reach, owing to limited access to this stage for both technical and ethical reasons[3,4]. The early post-implantation stage human embryo, with its flat bilaminar disc structure, is morphologically different from its mouse counterpart. As such it may demand new engineering approaches to be efficiently modeled. *In vitro* models of embryogenesis based on both mouse and human-derived cell types have emerged as instrumental platforms to probe human-specific developmental mechanisms[5-20]. However, they often suffer limitations, such as low efficiency and high technical complexity. Traditionally these studies use cocktails of growth factors at supra-physiological levels and encounter challenges finding common media that support diverse cellular fates. Additionally, human post-implantation embryo models with co-developing embryonic and extra-embryonic layers are still missing, and a full integrated human embryo-like model can encounter ethical challenges, which limits our ability to study integrated morphogenetic events of human embryo after implantation[3,21].

Here, we present a human stem cell-derived embryo model, dubbed "***iDiscoid***", that overcomes several of these limitations. The iDiscoid platform leverages an approach that genetically engineers a population of human induced pluripotent stem cells (hiPSCs) with inducible circuits. Upon activation of these genetic circuits, an extra-embryonic niche is formed that triggers self-organization of the adjacent wild-type hiPSC clusters. Using iDiscoids, we demonstrate development of human bilaminar disc-like structures, specification of anterior-like hypoblast cells, lumenogenesis, formation of amnion-like tissue, symmetry breaking leading to posterior-like axis specification, and yolk sac specification with hematopoietic characteristics. iDiscoids model human embryogenesis at an early stage of implantation when the embryo *in vivo* is at its most inaccessible state.

Engineering co-development of embryonic and extra-embryonic endoderm tissues

We engineered an hiPSC cell line with an inducible gene circuit that expresses human GATA6, a key transcription factor for extra-embryonic endodermal fate (henceforth referred to as iGATA6) **(Fig. 1A, Supplementary Fig. 1A)**. These iGATA6 cells express different levels of GATA6 upon the addition of a small molecule, Doxycycline (Dox), distinguishable by expression of the reporter enhanced green fluorescent protein (EGFP) **(Supplementary Fig. 1B)**. When iGATA6 and wild-type (WT) hiPSCs were co-assembled in 3D, EGFP$^+$ extra-embryonic endoderm cells segregated to the outer layer, similar to what occurs during implantation period *in vivo.* However, the overall organization of these aggregates did not recapitulate the structural format of bilaminar disc and yolk sac tissue. This approach yielded aggregates in which the entire EGFP$^+$

extra-embryonic endoderm is in close contact with wild type epiblast-like cells, missing the part of the yolk sac tissue (parietal) that is formed away from the epiblast cells. Additionally, in these 3D aggregates, we found that WT cells could not consistently induce symmetry breaking to amnion-like and epiblast-like domains **(Supplementary Fig. 1C-D)**. Moreover, *in vitro* reattachment of both synthetic and true human blastocysts to a dish to mimic implantation has shown organizational instability that results in tissue organization that cannot reflect the developmental stages approaching gastrulation[7,16,20,22-24]. Mimicking stable co-development of embryonic and extra-embryonic cell types through post-implantation embryonic stages from an initial 3D state is technically difficult to control.

To address these challenges, we aimed to develop an alternative strategy to generate structures off the cell culture plate via igniting 2D-to-3D self-organization of hiPSCs. The iGATA6-hiPSCs were mixed with WT hiPSCs, and the mixed population of cells were seeded onto standard culture plates at a defined ratio and cell density **(Fig. 1A)**. After treatment with Dox, the iGATA6 cells expectedly upregulate GATA6 and EGFP and exhibit loss of the pluripotency markers NANOG **(Supplementary Fig. 2A)**. During the first 48 hours, from an initial state of random distribution, induced iGATA6 cells and WT cells proliferate, and WT cells are organized into disc-shaped clusters confined by iGATA6 cells **(Supplementary Video 1, Fig. 1B, Supplementary Fig. 1G)**. Subsequently, the migration of the iGATA6 cells over the upper surface of these WT clusters occurs in conjunction with 3D growth within the WT disc **(Supplementary Fig. 2C, Supplementary Video 2)**. These events result in a membrane of extra-embryonic endoderm-like cells that express GATA4 and SOX17 but lack OCT4 over a multilayer WT structure **(Fig. 1A, 1C-D, Supplementary Fig. 2C-D)**. The process of upward iGATA6 migration is in parallel with the deposition of a laminin membrane surrounding the WT clusters **(Supplementary Fig. 3A-B)**. The laminin deposition by the iGATA6 cells can subsequently trigger cell polarization within cells of the WT cluster, which promotes the conversion of cellular rosettes into lumens **(Fig. 1C, Supplementary Fig. 3C, Supplementary Video 3, 4)**[25,26]. This is facilitated by expression of PODXL, an apically expressed antiadhesive surface protein implicated in embryonic lumen formation, and ZO-1, an apical tight junction protein **(Supplementary Fig. 3C, D)**[27]. This laminin deposition and lumen formation is not observed in WT-only culture in equivalent conditions **(Supplementary Fig. 3E)**.

We observed that within WT clusters for which an iGATA6 layer was not surrounding the cluster, lumen formation was not observed **(Fig. 1E)**. The number of clusters with no coverage by an iGATA6 layer progressively decreases by day 5 as the migration of iGATA6 cells moves forward **(Fig. 1E)**. Ultimately, following expansion of the lumen, the two cell monolayers of epiblast-like WT ($NANOG^+$) and endoderm-like $GATA6^+$ cells are positioned on either side of a laminin membrane **(Fig. 1A, 1C, Supplementary Fig. 3C)**, forming a cellular arrangement similar to the bilaminar disc in the early post-implantation human embryo **(Supplementary Fig. 1E)**[28]. Analysis of iDiscoid culture media supported high levels of secreted AFP and APOA1, two proteins known to be produced by primitive endoderm and yolk sac tissues **(Supplementary Fig. 3F)**.

In our initial analysis, we observed that for WT clusters at day 5 with higher area and often lower circularity, cavity formation occurred via multiple lumens **(Supplementary Fig. 4A, B)**. In order to minimize the occurrence of large WT clusters with multiple lumens and control the cluster size and number, we tested a range of seeding ratio and densities for iGATA6 and WT cells, with the objective to achieve cluster area most consistently within a size range corresponding to the E9-E17 human bilaminar disc **(Supplementary Fig. 4C, D)** (Carnegie #8004, #7700 in [29], Carnegie #7801 in [30]). To this end, we selected a seeding ratio of 81:5 iGATA6:WT at a density of 54k cells per cm^2 **(Supplementary Fig. 4C, D**, dotted boxes**)**. When we applied this optimized seeding density, we could detect 74% of iDiscoids in the expected physiological size range **(Supplementary Fig. 4E)**. Importantly we also observed the majority (70.9%) of iDiscoids in this size range produced a single lumen **(Supplementary Fig. 4E)**. Hence, by altering the initial parameters we could control iDiscoid size, number, circularity, and lumenogenesis.

Single cell RNA transcriptomics reveals embryonic and extra-embryonic fate trajectories

To investigate the cellular identities of the developed population, we performed single cell RNA sequencing (scRNA-seq) from day 0 to day 5 of iDiscoid cultures **(Supplementary Fig. 5-7)**. Performing an unsupervised comparison of the transcriptomes with the E16-19 human embryo[31], E6-E14 human embryo samples[24,32], and the E18-20 cynomolgus embryo[33] revealed the acquisition of distinct post-implantation cellular fates in identified iDiscoid subpopulations **(Fig. 1F, G, Supplementary Fig. 6A-C)**.

Beginning as early as 24 hours (D1) after induction, iDiscoids showed the emergence of cell populations (D1_G1, G2, G3) displaying transcriptomic similarity to human hypoblast, alongside a corresponding epiblast-like population (D1_W1) **(Supplementary Fig. 6A-C)**. By 36 hours (D1.5), iGATA6 cells become statistically similar to the human E12 hypoblast lineage in from Xiang et al., and this similarity strengthens through day 5 of culture **(Supplementary Fig. 6B)**[32]. Comparison with the E16-19 human and cynomolgus embryo shows significant similarity to yolk sac endoderm and hypoblast appearing at D1 and strengthening through day 5 ($p = 6.00\times10^{-54}$ to 2.84×10^{-93} similarity significance to YS Endoderm at D5) **(Fig 1F, Supplementary Fig. 6A, 6C)**.

On day 2, there is significant statistical similarity to yolk sac endoderm within the iGATA6 population (Clusters D2_G1, D2_G2, D2_G3, D2_G4; $p = 1.3\times10^{-17}$ to 9.5×10^{-69}), with the highest GATA6-expressing cells (Cluster D2-G5) diverging from this fate towards a more yolk sac mesoderm-like state ($p = 6.8\times10^{-11}$) **(Fig. 1F, Supplementary Fig. 6A-C)**. By day 4, we observed the strengthening of the identities of subpopulations within the iGATA6 and fate divergence within the WT compartment. The lower-GATA6 clusters (D4_G1 and D4_G2) showed increased expression of key yolk sac marker genes (*GATA4*, *PDGFRA*, *LAMA1*, *CUBN*, *AMN*, *NODAL*) and statistical similarity to human yolk sac endoderm **(Fig. 1G, Supplementary Fig. 5B, 6A-C,** $p_{G1} = 1.3\times10^{-92}$, $p_{G2} = 1.8\times10^{-63}$**)**[34]. A "medium" GATA6 population (D4_G3 and D4_G4) showed lower statistical similarity to either of these fates, but has a strong anterior hypoblast identity (homologous to mouse Anterior Visceral Endoderm, AVE), with specific upregulation of anterior markers (*LHX1*, *HHEX*, *FZD5*, *CER1*, *LEFTY1*) **(Supplementary Fig. 5B)**[31,35-37]. The highest-GATA6

clusters, D4_G5 and D4_G6, demonstrated statistical similarity to human yolk sac mesoderm **(Fig. 1G, Supplementary Fig. 6A,B**, p_{G5} = 1.7×10^{-22}, p_{G6} = 2.5×10^{-25}**)**.Taken together, this demonstrates a divergence in yolk sac-associated fates within iGATA6 cells that is correlated with different GATA6 expression levels, with yolk sac endoderm, anterior hypoblast, and yolk sac mesoderm-like cells stratifying between populations expressing low, medium, and high average levels of *GATA6* **(Supplementary Fig. 6** top row**)**. Recently observations in cynomolgus monkey by Zhai et al. suggest a similar stratification of GATA6 levels in a variety of tissues, notably between yolk sac endoderm (lower GATA6) and mesoderm (higher GATA6) populations[38].

Within the WT clusters, identified via a lack of *GATA6* or *EGFP* transcripts, we observe the divergence of three separate populations from the epiblast-like compartment. The first and largest cluster (D4_W1) maintained the similarity to human epiblast, with expression of the pluripotency markers *POU5F1* (OCT4), *SOX2*, and *NANOG* **(Supplementary Fig. 5B,** p_{W1} = 1.15×10^{-49}**)**. The second (D4_W2) WT cluster had lower statistical similarity to human epiblast (p_{W2} = 3.4×10^{-20} to D16-19 Epi), but stronger similarity to human primitive streak (p_{W2} = 3.6×10^{-95} to D16 PrS) **(Fig. 1G, Supplementary Fig. 5B, 6A)**. Our assessment of the third wild type cluster at this timepoint, (D4_W3) showed occurrence of a unique transcriptomic similarity to human and cynomolgus amnion (p_{W3} = 8.0×10^{-15} and 1.6×10^{-25}, respectively) **(Fig. 1G, Supplementary Fig. 5B, 6A, 6C)**[17,39,40]. We did not observe a similarity to the trophoblast lineages **(Supplementary Fig. 6)**.

When we integrated all clusters across different days, iGATA6 and WT lineages occupied distinct compartments within the projection space. However, we did observe spatial positioning of populations aligned with their time of sampling **(Supplementary Fig 7A)**. Populations sampled at a different time can show similar fates (i.e., D3 versus D5 iGATA6 yolk sac endoderm cells) **(Supplementary Fig 7B)**. The top upregulated genes show distinct patterns across different cell fates, with overlapping genes present across the similar fates sampled at different time points **(Supplementary Fig 7C)**.

Specification of amniotic ectoderm in iDiscoids

Primate amniogenesis is structurally and temporally distinct from the process in mouse, highlighting the need for human model systems to provide a window into understanding human-specific developmental dynamics[40]. *In vivo*, following lumenogenesis within the epiblast of the primate post-implantation embryo, the cells of the epiblast undergo dorsal-ventral (D-V) patterning to separate the epiblast from amniotic ectoderm[15,17,29]. Our single cell transcriptomic analysis in wild type cells showed a group of cells with transcriptomic similarity to human and cynomolgus amnion that emerges around day 3 **(Fig. 2A-B, Supplementary Fig. 6A, 7B)**. These cells express amnion markers, including *ISL1*, *TFAP2A* (AP-2**α**), and *GATA3*, maintaining expression of OCT4 while substantially lacking expression of *NANOG* **(Fig. 2A-B)**[17,39,40]. Our analysis via immunofluorescent staining corroborated this finding, demonstrating a layer of ISL1^{+}/AP-2**α**$^{+}$ in iDiscoids **(Fig. 2C-E, Supplementary Fig. 8A)**. We observe that these cells are spatially segregated within the cavitated sac-like structure, with the amnion-like lineage

224 positioning away from the bilaminar iGATA6/WT disc, forming a membrane against the tissue
225 culture dish. WT cells above within the bilaminar iGATA6/WT disc display expression of NANOG
226 **(Fig. 2C-E)**.
227
228 Further immunofluorescence staining for pSMAD1, pSMAD5, and pSMAD8/9 (BMP4
229 effectors) revealed BMP4 signal transduction in a ring pattern concentrated at the edges of the
230 iDiscoid **(Fig. 2F)**. scRNA-seq at day 4 also showed specific upregulation of BMP4 targets and
231 related genes within the amnion-like population **(Supplementary Fig. 8B)**. Recently, the active
232 role of primate amniotic tissue to promote BMP4 signaling has been suggested, acting
233 downstream of ISL1[40]. It was also reported that BMP4 can promote the formation of ISL1^{+}
234 amniotic tissue in a human embryo model[15,17,39]. Hence, we aimed to inhibit BMP signaling via
235 treatment with the inhibitor Noggin to investigate the relevance of amnion-like tissue formation
236 in iDiscoid. We observed effective suppression of SMAD1/5/8 phosphorylation **(Fig. 2G)**, and a
237 marked decrease in amnion fate acquisition assessed by ISL1-expression **(Fig. 2H-I).** However,
238 the expression of NANOG remained intact in epiblast-like cells **(Fig. 2H)**. To this end, we showed
239 98.8% of WT clusters express the amnion marker ISL1 around day 4 of iDiscoid development
240 that was consistent across experiments. We show that all (n = 87) but one of the observed WT
241 clusters (n = 88) have ISL1 expression above the expression levels in the noggin treatment group.
242
243 Collectively, we conclude that iDiscoids show robust amniogenesis with formation of a
244 dorsal-ventral (D-V) axis that is dependent on BMP4 signaling, supporting the utility of iDiscoids
245 for studying this key developmental event in humans.
246
247 **Development of anterior hypoblast poles and posterior axis in iDiscoids**
248
249 *In vivo* evidence shows a group of cells with anterior visceral endoderm characteristics in
250 human and non-human primates that have been termed anterior hypoblast[36]. In mouse, this
251 tissue has been shown to pattern the specification of the posterior axis via the expression of
252 signaling inhibitors including CER1. However, common *in vitro* stem cell models of the human
253 embryo (i.e., blastoids) have not yet captured this specification[7,16,20]. We asked whether iDiscoids
254 can exhibit features associated with anterior-posterior (A-P) axis formation **(Fig. 3)**. Our scRNA-
255 seq analysis of iGATA6 cells as early as day 2 revealed a cluster of cells that exhibit markers
256 associated with anterior hypoblast. These cells are positive for *CER1*, *LHX1*, *HHEX*, *LEFTY1*, and
257 *LEFTY2*, and continued to present these markers in the following days **(Fig. 3A)**[31,35-37].
258 Additionally, we also detected the emergence of a distinct cluster of cells at day 3 and 4
259 presenting *TBXT* (Brachyury), *MIXL1*, *GDF3*, and *EOMES*, key markers associated with
260 development of the posterior pole in the embryo **(Fig. 3B)**[41].
261
262 Immuno-staining identified domains of cells expressing the anterior hypoblast-related
263 proteins (CER1, LHX1, HHEX) in a subset of extra-embryonic cells surrounding the WT clusters
264 **(Fig. 3C)**. Out of 396 of iDiscoids assessed we identified 42.4±4.1% (169) discs with polar CER1-
265 expressing domains on day 4 **(Fig. 3C)**, exhibiting a similar efficiency to existing *ex vivo* human
266 blastocyst cultures[36]. In parallel, we showed development of TBXT^{+} domains in 46.4±9.0% (323

267 out of 746) of WT clusters on day 4 in iDiscoids, of which 30.1±6.3% (215) of clusters with these
268 domains displayed asymmetric (polar) expression **(Fig. 3D-E)**. These posterior domains are
269 dependent to local BMP4 signaling, as a small molecule, Noggin, could eliminate TBXT$^+$ poles
270 with no effect on development of CER1-expressing cells. This supports a role for BMP4 from
271 amnion in posterior axis specification in line with past studies within these tissues in iDiscoids
272 (**Supplementary Fig. 9)**[42].
273
274 We then analyzed iDiscoid cultures to evaluate A-P axis positionings. On day 4, we
275 observed that 60.4±5.3% (102 of 169) of iDiscoids with polar configuration in CER1-expressing
276 cells possessed a TBXT-expressing pole **(Fig. 3F)**. Within iDiscoids with polarity in both CER1
277 and TBXT we observed two distinct patterns **(Fig. 3F, G, Supplementary Fig. 10)**. 42.2±3.2%
278 exhibit anti-polar configuration, where the TBXT$^+$ pole occupies the opposite radial region as
279 the CER1$^+$ domain, similar to A-P axis positioning in mouse at E6.5 **(Fig. 3F-H)**. In a second set
280 of polar iDiscoids, we saw high expression of CER1 in the proximity of TBXT$^+$ domains **(Fig. 3F-**
281 **H, Supplementary Fig. 10A)**. This set of iDiscoids shows CER1$^+$ and TBXT$^+$ poles with a same
282 radial region (syn-polar), in contrast to past reported literature. We also noted expression of
283 CER1 in a subset of TBXT$^+$ cells in the posterior domains in some cases **(Supplementary Fig.**
284 **10A, B)** and subsequent scRNA-seq analysis showed the presence of a subpopulation of
285 *CER1*$^+$/*TBXT*$^+$ cells in the cluster that represents posterior identity **(Supplementary Fig. 10C)**.
286
287 Next, we aimed to determine whether CER1 expression adjacent to or within the posterior
288 region observed in iDiscoids can be identified during *in vivo* embryogenesis. To address this
289 question, we first examined the recent spatial scRNA-seq data from human and non-human
290 primates. In marmoset embryos[37], we noted non-polar *CER1* at CS5, followed by very slight
291 polarization away from the TBXT-expressing posterior at CS6 (subtly anti-polar-like)
292 **(Supplementary Fig. 11A)**. However, in human E16-19 (CS7)[31], we observed higher *CER1*
293 transcript expression in the same caudal embryonic region as the primitive streak
294 **(Supplementary Fig. 11B)**. Lastly, pseudospatial assessment of the E18-20 cynomolgus embryo
295 shows *CER1* and *TBXT* expression overlapping within the lower region of the embryonic disc
296 (hypoblast) and embryonic disc compartments (syn-polar-like) **(Supplementary Fig. 11C)**[43]. These
297 three datasets show a subset of cells or regions that display expression of *CER1* near or within
298 the posterior domain. To re-examine this phenotype in mouse, we assessed three independent
299 mouse scRNA-seq datasets spanning ~E6.5-E8 and could also detect a similar pattern of *Cer1*
300 expression within the anterior boundary of the primitive streak **(Supplementary Fig. 11D-F)**[44-46].
301
302 Taken together, our analyses of iDiscoids, in tandem with exploration of available *in vivo*
303 data, suggest that the observed expression and positioning of TBXT$^+$ and CER1$^+$ cells in iDiscoids
304 can be seen *in vivo* and reveal distinct developmental stages. Anti-polar positioning of CER1 and
305 TBXT domains reflect positioning of anterior hypoblast-like and posterior domains during early
306 primitive streak formation around E14; syn-polar positioning and co-expression of CER1 and
307 TBXT is reflective of a more advanced stages post-E14 where CER1 expression acts either to
308 prevent spreading of posterior domain[47] or represent early mesendoderm commitment derived
309 from a primitive streak[38]. We propose iDiscoids provide a model to further study A-P axis biology,

310 the formation of anterior hypoblast-like cells, the posterior pole and their respective cellular
311 progeny in a human context.
312
313 **Specification of yolk sac mesoderm-like cells with hematopoietic potential**
314
315 The yolk sac represents a multifunctional extra-embryonic tissue that provides nutritional
316 support for the embryo while also acting as a hub for early blood production that emerges
317 around E16-18 of human development and matures thereafter[2,48,49]. Multicellular models of
318 human yolk sac morphogenesis with subsets of mesoderm-derived cells and *in vivo*-like structural
319 characteristics are missing, imposing a challenge for the dissection molecular underpinnings of
320 the process. We detected progressive developmental maturation aligned with yolk sac cellular
321 fates in iGATA6 compartment of iDiscoids **(Fig. 1G, Supplementary Fig 3F, 6)**. We further asked
322 if iGATA6 cells in iDiscoids could mimic development of mesodermal subsets within the
323 developing yolk sac endoderm. Hypergeometric statistical analysis of the clusters with highest
324 expression of GATA6 on day 4 (D4_G6) and 5 (D5_G3) reveals significant similarity to human yolk
325 sac mesoderm, hematopoietic lineage progenitors, and hemogenic endothelial cells as
326 compared with the E16-19 human embryo **(Fig. 4A, Supplementary Fig. 7A)**[31]. Examination of
327 the D5_G3 cluster reveals higher expression of the yolk sac mesoderm marker cell surface marker
328 BST2, as well as of ECM proteins **(Supplementary Fig. 12A)**[49,50]. Subpopulations of cells within
329 this cluster also exhibit a higher average expression markers of human yolk sac mesoderm
330 (*CREB3L1*, *NR2F2*, *PLAGL1*, *ANXA1*, *NID2*)[31,40,50], markers of endothelial cells (*CD34*, *PECAM1*
331 (CD31), *TEK*, *ICAM1, ETV2*)[31,51,52], and markers of human hematopoiesis (*GATA1*, *GATA2*, *KLF1*,
332 *KLF2*, *ZEB1*, *ZEB2*, *GYPA* (CD235a), and *GYPB* (CD235b)) **(Fig. 4B, Supplementary Fig.**
333 **12B)**[31,49,53].
334
335 Hematopoietic processes emerge *in vivo* in the form of blood islands within yolk sac
336 tissues. In iDiscoids, immunofluorescence analysis initially reveals spindle shape cells positive for
337 TAL1 (scl), a key regulator of blood development[54]. Further analysis shows a subset of these cells
338 were also positive for CD34 and ERG, markers associated with hematopoietic and endothelial
339 fates or RUNX1, a master transcription factor of hematopoiesis **(Supplementary Fig. 13A-C)**[54].
340 Further analysis shows a subset of these cells were also positive for CD34 and ERG, markers
341 associated with hematopoietic and endothelial fates or RUNX1, a master transcription factor of
342 hematopoiesis **(Supplementary Fig. 13A-C)**[55,56]. Image analysis of z confocal sections shows ERG^+
343 cells consistently emerged underneath an iGATA6 layer **(Fig. 4C, Supplementary Fig. 13D)**, a
344 morphology that may mimic an initial *in vivo* localization similar to human yolk sac mesoderm,
345 where this tissue is positioned against ECM in the area underneath the yolk sac endoderm[55,56].
346 Observed $CD34^+$/$TAL1^+$ cells were also $EGFP^+$, indicating they develop from an iGATA6-derived
347 parental population, supporting a hypoblast origin **(Fig. 4C)**. Staining for other endothelial
348 markers KDR (VEGFR2) and CDH5 also confirmed similar localization in day 5 iDiscoids
349 **(Supplementary Fig. 13E)**. A colony-forming unit (CFU) assay initiated with $CD34^+$ cells followed
350 by flow cytometry analysis of the cultures, showed generation of $CD45^+$ cells, erythroid cells
351 ($CD71^+$, $CD235a^-$ (early stage) or $CD235a^+$ (late stage)) as well as $CD11b^+$ myeloid cells
352 **(Supplementary Fig. 13F)**.

353
354 Having observed early hematopoietic potential in day 5 iDiscoid, we followed iDiscoids
355 beyond day 5 to better examine the hematopoietic potential of the system. We observed
356 minimal expansion of the CD34+ area when the culture was maintained on mTeSR
357 **(Supplementary Fig. 3G)**. Therefore, to facilitate the differentiation of progenitor cells, we used
358 basal media (IMDM without Dox) after day 5. We observed the expansion of the CD34$^+$
359 endothelial-like cell area significantly between day 5 and day 12, from covering 1.2±0.46% of the
360 culture area on day 5 to covering 14.7±6.0% of the culture area at day 12, an expansion that was
361 also confirmed by flow cytometry analysis **(Supplementary Fig. 13G, H)**. Around this time, we
362 observed the emergence of condensed areas containing spherical cells within the yolk sac-like
363 compartment **(Fig. 4D)**. Immunofluorescence staining of these areas revealed the presence of a
364 CD31$^+$ endothelial layer surrounding spherical cells expressing Leukosialin (CD43), a marker of
365 hematopoietic progenitors **(Fig. 4E)**[53,57]. Further investigation of these spherical cells found that
366 many expressed the erythropoietic marker genes CD235a and hemoglobin **(Supplementary Fig.**
367 **14A, 15A)**[31]. Having observed yolk sac mesodermal and hemogenic progenitors of this cell type
368 primarily correlating within the GATA6-hi clusters in scRNA-seq, we further supplemented
369 iDiscoid cultures at the time of seeding with high-GATA6 expressing cells and observed
370 significantly increased generation of these CD43$^+$ hematopoietic foci **(Fig. 4F-G, Supplementary**
371 **Fig. 15B)**, demonstrating predictable and controllable fate augmentation at the time of seeding.
372
373 Further investigation of the cell types observed at day 12 revealed cells with markers of
374 yolk sac-derived myeloid lineages, including macrophages (CD33$^+$/CX$_3$CR1$^+$) and
375 megakaryocytes (CD41$^+$/CD42b$^+$) **(Supplementary Fig. 14B-C)**[58-60]. Structural image analysis of
376 CD43$^+$ foci show intricate 3D hierarchical organization, with CD34$^+$ endothelial vessels packed
377 between FOXA2$^+$ endodermal cells occupying the top compartment and Desmin$^+$ mesodermal
378 cells forming the basal layer **(Fig. 4H-J, Supplementary Fig. 16A-B)**. This is in line with the tissue
379 morphology reported *in vivo* in blood islands of the developing human yolk sac at around E19-
380 23 **(Fig. 4K)**[49]. Late time point analysis of day 21 iDiscoid cultures using scRNA-seq and flow
381 cytometry similarly supports emergence of populations aligned with those expected from the
382 human yolk sac[61]. This analysis shows identification of: *CDH5* (VE-cadherin)$^+$ /*CD34*$^+$ /*CD31*$^+$ /*KDR*
383 (VEGFR2)$^+$ /*CD93* (AA4.1)$^+$ /*NT5E* (CD73)$^-$ /*GYPA*$^-$ hemogenic endothelium-like cells; *GYPA*$^+$
384 /*GYPB*$^+$ /*TRDN* (CD71)$^+$ /*HBG1*$^+$ /*HBE1*$^+$ /*HBZ*$^+$ erythrocyte-like cells; *CD226*$^+$ /*CLEC1B*$^+$ /*PLEK*$^+$
385 /*GP1BA* (CD42b)$^+$ /*GATA1*$^+$ megakaryocyte-like cells; and *CD33*$^+$ /*CD68*$^+$ /*CD45*$^+$ /*CD53*$^+$
386 /*CX3CR1*$^+$ macrophage-like cells **(Supplementary Fig. 14D-F)**. Flow cytometry analysis of these
387 cultures also corroborated the development of a CD43$^+$/CD235a$^+$/CD71$^+$ erythrocyte-like
388 population, a CD43$^+$/CD33$^+$/CD31$^+$ macrophage-like population, and a CD43$^+$/CD42b$^+$
389 megakaryocyte-like population **(Supplementary Fig. 14G)**.
390
391 **Discussion**
392
393 iDiscoids, presented here, develop both embryonic and extra-embryonic lineages with
394 several features observed together for the first time in an *in vitro* stem cell-based human
395 developmental model. These features show close proximity to native post-implantation human

396 embryogenesis yet are uniquely observed together in an hiPSC-based system **(Supplementary**
397 **Fig. 17)**. The iDiscoid platform shows morphogenesis of bilaminar disc-like structures including
398 the formation of an amniotic cavity, the segregation of amnion from pluripotent epiblast-like
399 cells, the generation of anterior hypoblast-like cells and a posterior axis, as well as yolk sac
400 morphogenesis. Our extended follow up revealed morphologically relevant blood islands and
401 the specification of hematopoietic lineages **(Supplementary Fig. 17)**. In our follow up, WT clusters
402 exhibited characteristics mainly aligned with ectodermal fates, which is a subject for future
403 studies and optimization **(Supplementary Fig. 18)**. We also show the utility of the iDiscoid
404 platform to study developmental processes.
405
406 From an engineering angle, iDiscoid platform highlights the power of tissue niche
407 engineering and co-differentiation to promote morphogenetic events *in vitro*. It demonstrates
408 that a combination of genetic circuit-based tissue engineering with the formation of tissue-tissue
409 boundaries enables geometric confinement and subsequent self-organization in 3D. Use of
410 genetic circuits can also alleviate the need for complex regimes of growth factors to support
411 heterotypic tissues. The extra-embryonic differentiation of iDiscoids is pre-programmed into
412 undifferentiated hiPSCs via an inducible genetic switch and mixed with WT hiPSCs; hence, the
413 undifferentiated cell mix can be expanded, cryostored, and shipped for use on demand,
414 requiring only 2D culture plates, commercially available medium, and addition of a small
415 molecule, Dox **(Supplementary Fig. 19)**. The high-throughput format, efficient generation,
416 compatibility with live imaging, and an easy-to-implement protocol enables establishment and
417 utility across different labs. These characteristics, coupled with the potential to be produced from
418 hiPSCs with diverse genetic backgrounds **(Supplementary Fig. 20)**, facilitate the studies on early
419 stages of post-implantation, opening new routes for drug testing, developmental toxicology,
420 and tractable disease modeling with fewer ethical concerns in the human context.
421

422 ## References

423 1 Muller, F. & O'Rahilly, R. The primitive streak, the caudal eminence and related
424 structures in staged human embryos. *Cells Tissues Organs* **177**, 2-20,
425 doi:10.1159/000078423 (2004).
426 2 Palis, J. & Yoder, M. C. Yolk-sac hematopoiesis: the first blood cells of mouse and man.
427 *Exp Hematol* **29**, 927-936, doi:10.1016/s0301-472x(01)00669-5 (2001).
428 3 Clark, A. T. *et al.* Human embryo research, stem cell-derived embryo models and
429 in vitro gametogenesis: Considerations leading to the revised ISSCR guidelines. *Stem*
430 *Cell Reports* **16**, 1416-1424, doi:https://doi.org/10.1016/j.stemcr.2021.05.008 (2021).
431 4 Pera, M. F. Human embryo research and the 14-day rule. *Development* **144**, 1923-1925,
432 doi:10.1242/dev.151191 (2017).
433 5 Beccari, L. *et al.* Multi-axial self-organization properties of mouse embryonic stem cells
434 into gastruloids. *Nature* **562**, 272-276, doi:10.1038/s41586-018-0578-0 (2018).
435 6 Harrison, S. E., Sozen, B., Christodoulou, N., Kyprianou, C. & Zernicka-Goetz, M.
436 Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro.
437 *Science* **356**, doi:10.1126/science.aal1810 (2017).

438 7 Kagawa, H. *et al.* Human blastoids model blastocyst development and implantation.
439 *Nature* **601**, 600-605, doi:10.1038/s41586-021-04267-8 (2022).
440 8 Karzbrun, E. *et al.* Human neural tube morphogenesis in vitro by geometric constraints.
441 *Nature* **599**, 268-272, doi:10.1038/s41586-021-04026-9 (2021).
442 9 Li, R. *et al.* Generation of Blastocyst-like Structures from Mouse Embryonic and Adult
443 Cell Cultures. *Cell* **179**, 687-702 e618, doi:10.1016/j.cell.2019.09.029 (2019).
444 10 Simunovic, M. *et al.* A 3D model of a human epiblast reveals BMP4-driven symmetry
445 breaking. *Nat Cell Biol* **21**, 900-910, doi:10.1038/s41556-019-0349-7 (2019).
446 11 Sozen, B. *et al.* Self-Organization of Mouse Stem Cells into an Extended Potential
447 Blastoid. *Dev Cell* **51**, 698-712 e698, doi:10.1016/j.devcel.2019.11.014 (2019).
448 12 Veenvliet, J. V. *et al.* Mouse embryonic stem cells self-organize into trunk-like structures
449 with neural tube and somites. *Science* **370**, doi:10.1126/science.aba4937 (2020).
450 13 Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to
451 recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat*
452 *Methods* **11**, 847-854, doi:10.1038/nmeth.3016 (2014).
453 14 Yu, L. *et al.* Blastocyst-like structures generated from human pluripotent stem cells.
454 *Nature* **591**, 620-626, doi:10.1038/s41586-021-03356-y (2021).
455 15 Zheng, Y. *et al.* Controlled modelling of human epiblast and amnion development
456 using stem cells. *Nature* **573**, 421-425, doi:10.1038/s41586-019-1535-2 (2019).
457 16 Liu, X. *et al.* Modelling human blastocysts by reprogramming fibroblasts into iBlastoids.
458 *Nature* **591**, 627-632, doi:10.1038/s41586-021-03372-y (2021).
459 17 Shao, Y. *et al.* A pluripotent stem cell-based model for post-implantation human
460 amniotic sac development. *Nature Communications* **8**, 208, doi:10.1038/s41467-017-
461 00236-w (2017).
462 18 Mathew, B. *et al.* Mouse ICM Organoids Reveal Three-Dimensional Cell Fate Clustering.
463 *Biophysical Journal* **116**, 127-141, doi:https://doi.org/10.1016/j.bpj.2018.11.011 (2019).
464 19 Schröter, C., Rué, P., Mackenzie, J. P. & Martinez Arias, A. FGF/MAPK signaling sets the
465 switching threshold of a bistable circuit controlling cell fate decisions in embryonic stem
466 cells. *Development* **142**, 4205-4216, doi:10.1242/dev.127530 (2015).
467 20 Fan, Y. *et al.* Generation of human blastocyst-like structures from pluripotent stem cells.
468 *Cell Discovery* **7**, 81, doi:10.1038/s41421-021-00316-8 (2021).
469 21 Weatherbee, B. A. T., Cui, T. & Zernicka-Goetz, M. Modeling human embryo
470 development with embryonic and extra-embryonic stem cells. *Dev Biol* **474**, 91-99,
471 doi:10.1016/j.ydbio.2020.12.010 (2021).
472 22 Zhao, C. *et al.* Reprogrammed blastoids contain amnion-like cells but not
473 trophectoderm. *bioRxiv*, 2021.2005.2007.442980, doi:10.1101/2021.05.07.442980
474 (2021).
475 23 Deglincerti, A. *et al.* Self-organization of the in vitro attached human embryo. *Nature*
476 **533**, 251-254, doi:10.1038/nature17948 (2016).
477 24 Chhabra, S. & Warmflash, A. BMP-treated human embryonic stem cells transcriptionally
478 resemble amnion cells in the monkey embryo. *Biol Open* **10**, doi:10.1242/bio.058617
479 (2021).

480 25 Bedzhov, I. & Zernicka-Goetz, M. Self-organizing properties of mouse pluripotent cells
481 initiate morphogenesis upon implantation. *Cell* **156**, 1032-1044,
482 doi:10.1016/j.cell.2014.01.023 (2014).
483 26 Li, S., Edgar, D., Fassler, R., Wadsworth, W. & Yurchenco, P. D. The role of laminin in
484 embryonic cell polarization and tissue organization. *Dev Cell* **4**, 613-624,
485 doi:10.1016/s1534-5807(03)00128-x (2003).
486 27 Shahbazi, M. N. *et al.* Pluripotent state transitions coordinate morphogenesis in mouse
487 and human embryos. *Nature* **552**, 239-243, doi:10.1038/nature24675 (2017).
488 28 O'Rahilly, R., Müller, F. & Washington, C. I. o. *Developmental Stages in Human*
489 *Embryos: Including a Revision of Streeter's "Horizons" and a Survey of the Carnegie*
490 *Collection*. (Carnegie Institution of Washington, 1987).
491 29 Hertig, A. T. & Rock, J. Two human ova of the pre-villous stage, having a
492 developmental age of about 8 and 9 days respectively. *Contrib Embryol* **33**, 169-186
493 (1949).
494 30 Heuser, C. H., Rock, J. & Hertig, A. T. Two human embryos showing early stages of the
495 definitive yolk sac. *Contributions to embryology* **201**, 85-99 (1945).
496 31 Tyser, R. C. V. *et al.* Single-cell transcriptomic characterization of a gastrulating human
497 embryo. *Nature* **600**, 285-289, doi:10.1038/s41586-021-04158-y (2021).
498 32 Xiang, L. *et al.* A developmental landscape of 3D-cultured human pre-gastrulation
499 embryos. *Nature* **577**, 537-542, doi:10.1038/s41586-019-1875-y (2020).
500 33 Ma, H. *et al.* In vitro culture of cynomolgus monkey embryos beyond early gastrulation.
501 *Science* **366**, doi:10.1126/science.aax7890 (2019).
502 34 Cindrova-Davies, T. *et al.* RNA-seq reveals conservation of function among the yolk sacs
503 of human, mouse, and chicken. *Proc Natl Acad Sci U S A* **114**, E4753-E4761,
504 doi:10.1073/pnas.1702560114 (2017).
505 35 Costello, I. *et al.* Lhx1 functions together with Otx2, Foxa2, and Ldb1 to govern anterior
506 mesendoderm, node, and midline development. *Genes Dev* **29**, 2108-2122,
507 doi:10.1101/gad.268979.115 (2015).
508 36 Mole, M. A. *et al.* A single cell characterisation of human embryogenesis identifies
509 pluripotency transitions and putative anterior hypoblast centre. *Nat Commun* **12**, 3679,
510 doi:10.1038/s41467-021-23758-w (2021).
511 37 Bergmann, S. *et al.* Spatial profiling of early primate gastrulation in utero. *Nature* **609**,
512 136-143, doi:10.1038/s41586-022-04953-1 (2022).
513 38 Zhai, J. *et al.* Primate gastrulation and early organogenesis at single-cell resolution.
514 *Nature* **612**, 732-738, doi:10.1038/s41586-022-05526-y (2022).
515 39 Shao, Y. *et al.* Self-organized amniogenesis by human pluripotent stem cells in a
516 biomimetic implantation-like niche. *Nature Materials* **16**, 419-425,
517 doi:10.1038/nmat4829 (2017).
518 40 Yang, R. *et al.* Amnion signals are essential for mesoderm formation in primates. *Nat*
519 *Commun* **12**, 5126, doi:10.1038/s41467-021-25186-2 (2021).
520 41 Bardot, E. S. & Hadjantonakis, A.-K. Mouse gastrulation: Coordination of tissue
521 patterning, specification and diversification of cell fate. *Mechanisms of Development*
522 **163**, 103617, doi:https://doi.org/10.1016/j.mod.2020.103617 (2020).

523 42 Chhabra, S., Liu, L., Goh, R., Kong, X. & Warmflash, A. Dissecting the dynamics of
524 signaling events in the BMP, WNT, and NODAL cascade during self-organized fate
525 patterning in human gastruloids. *PLoS Biol* **17**, e3000498,
526 doi:10.1371/journal.pbio.3000498 (2019).
527 43 Cui, G. *et al.* Spatial molecular anatomy of germ layers in the gastrulating cynomolgus
528 monkey embryo. *Cell Reports* **40**, 111285,
529 doi:https://doi.org/10.1016/j.celrep.2022.111285 (2022).
530 44 Mittnenzweig, M. *et al.* A single-embryo, single-cell time-resolved model for mouse
531 gastrulation. *Cell* **184**, 2825-2842.e2822, doi:10.1016/j.cell.2021.04.004 (2021).
532 45 Peng, G. *et al.* Spatial Transcriptome for the Molecular Annotation of Lineage Fates and
533 Cell Identity in Mid-gastrula Mouse Embryo. *Dev Cell* **36**, 681-697,
534 doi:10.1016/j.devcel.2016.02.020 (2016).
535 46 Pijuan-Sala, B. *et al.* A single-cell molecular map of mouse gastrulation and early
536 organogenesis. *Nature* **566**, 490-495, doi:10.1038/s41586-019-0933-9 (2019).
537 47 Stuckey, D. W., Di Gregorio, A., Clements, M. & Rodriguez, T. A. Correct patterning of
538 the primitive streak requires the anterior visceral endoderm. *PLoS One* **6**, e17620,
539 doi:10.1371/journal.pone.0017620 (2011).
540 48 Nakazawa, F., Alev, C., Jakt, L. M. & Sheng, G. Yolk sac endoderm is the major source
541 of serum proteins and lipids and is involved in the regulation of vascular integrity in
542 early chick development. *Dev Dyn* **240**, 2002-2010, doi:10.1002/dvdy.22690 (2011).
543 49 Ross, C. & Boroviak, T. E. Origin and function of the yolk sac in primate embryogenesis.
544 *Nat Commun* **11**, 3760, doi:10.1038/s41467-020-17575-w (2020).
545 50 Pham, T. X. A. *et al.* Modeling human extraembryonic mesoderm cells using naive
546 pluripotent stem cells. *Cell Stem Cell* **29**, 1346-1365.e1310,
547 doi:https://doi.org/10.1016/j.stem.2022.08.001 (2022).
548 51 Bruveris, F. F. *et al.* Human yolk sac-like haematopoiesis generates RUNX1-, GFI1-
549 and/or GFI 1B-dependent blood and SOX17-positive endothelium. *Development* **147**,
550 doi:10.1242/dev.193037 (2020).
551 52 Koyano-Nakagawa, N. *et al.* Etv2 is expressed in the yolk sac hematopoietic and
552 endothelial progenitors and regulates Lmo2 gene expression. *Stem Cells* **30**, 1611-
553 1623, doi:10.1002/stem.1131 (2012).
554 53 Vanhee, S. *et al.* In vitro human embryonic stem cell hematopoiesis mimics MYB-
555 independent yolk sac hematopoiesis. *Haematologica* **100**, 157-166,
556 doi:10.3324/haematol.2014.112144 (2015).
557 54 Vagapova, E. R., Spirin, P. V., Lebedev, T. D. & Prassolov, V. S. The Role of TAL1 in
558 Hematopoiesis and Leukemogenesis. *Acta Naturae* **10**, 15-23 (2018).
559 55 Enders, A. C. & King, B. F. in *The Human Yolk Sac and Yolk Sac Tumors* (ed Francisco
560 F. Nogales) 33-47 (Springer Berlin Heidelberg, 1993).
561 56 Knoth, M. & Larsen, J. F. Ultrastructure of a human implantation site. *Acta Obstet*
562 *Gynecol Scand* **51**, 385-393, doi:10.3109/00016347209156879 (1972).
563 57 Vodyanik, M. A., Thomson, J. A. & Slukvin, II. Leukosialin (CD43) defines hematopoietic
564 progenitors in human embryonic stem cell differentiation cultures. *Blood* **108**, 2095-
565 2105, doi:10.1182/blood-2006-02-003327 (2006).

566 58 Stremmel, C. *et al.* Yolk sac macrophage progenitors traffic to the embryo during
567 defined stages of development. *Nature Communications* **9**, 75, doi:10.1038/s41467-
568 017-02492-2 (2018).
569 59 Kappelmayer, J. *et al.* Consecutive Appearance of Coagulation Factor XIII Subunit A in
570 Macrophages, Megakaryocytes, and Liver Cells During Early Human Development.
571 *Blood* **86**, 2191-2197,
572 doi:https://doi.org/10.1182/blood.V86.6.2191.bloodjournal8662191 (1995).
573 60 Miyawaki, K. *et al.* Identification of unipotent megakaryocyte progenitors in human
574 hematopoiesis. *Blood* **129**, 3332-3343, doi:10.1182/blood-2016-09-741611 (2017).
575 61 Wittamer, V. & Bertrand, J. Y. Yolk sac hematopoiesis: does it contribute to the adult
576 hematopoietic system? *Cellular and Molecular Life Sciences* **77**, 4081-4091,
577 doi:10.1007/s00018-020-03527-6 (2020).
578

579 ## Methods

580 ### Ethics regarding development of iDiscoids

581 This research was approved and performed under the oversight of the University of
582 Pittsburgh Human Stem Cell Research Oversight Committee. The work that has been reported
583 in this study followed 2016 Guidelines for Stem Cell Research and Clinical Translation as well as
584 2021 ISSCR Guidelines on Ethical Standards for Stem Cell Embryo Models[62,63]. To this end,
585 iDiscoids do not involve human embryonic stem cells and are generated from human iPSCs that
586 are derived from human somatic cells such as fibroblasts. iDiscoids are attached to a cell culture
587 dish, lacking an extra-embryonic trophectoderm tissue critical for full integrated embryo
588 development and implantation in the uterine cavity. The yolk sac cavity in iDiscoid is not closed,
589 and the tissue cannot be harvested for any implantation without substantial disruption of their
590 structures. $TBXT^+$ posterior domains were observed during the study but were not sustained
591 during the development of our system. These features collectively restrict the ability of this model
592 to undergo the full integrated development of a human embryo *in vitro* and/or implant to
593 support further development of a conceptus *in vivo*.
594

595 ### Cell Culture

596 All cells and tissues were cultured in a humidified incubator at 37°C and 5% CO_2. Our
597 hiPSC lines were cultivated under sterile conditions in mTeSR-1 (Stem Cell Technologies,
598 Vancouver), changed daily. Tissue culture plates were coated for 1 hour at room temperature
599 with BD ES-qualified Matrigel (BD Biosciences) diluted according to the manufacturer's
600 instructions in ice cold DMEM/F-12. Routine passaging was performed by incubating hiPSC
601 colonies for 5 minutes in Accutase (Sigma) at 37°C, collecting the suspension and adding 5mL
602 DMEM/F-12 medium containing 10 μM Y-27632, centrifuging at 300g for 5 minutes, and
603 resuspending in DMEM/F-12 supplemented with 10 μM Y-27632 for counting. Cells were seeded
604 at a cell density of 25,000 cells per cm^2 for routine maintenance.
605

606 ### GATA6-Engineered Cell Line Generation

607 rtTA-expressing PGP1 hiPSCs previously generated[64] were transfected using Lipofectamine
608 3000 (Thermo Fisher Scientific) with Super PiggyBac Transposase (System Biosciences) and the

609 PiggyBac transposon vector with human GATA6-2A-EGFP under control of the tetracycline
610 responsive element promoter. Transfected cells were selected by adding 0.5mg/mL puromycin
611 to the mTeSR1 maintenance medium. PGP9 iGATA6 hiPSCs were engineered as explained
612 previously[64]. For the generation of GATA6-hi cell line, the iGATA6 cell line was sorted via FACS
613 (one cell per each well of a 96 well plate) in mTeSR-1 supplemented with 10 μM Y-27632 and
614 2μM Thiazovivin. The media was replaced on the day after sorting with mTeSR-1 supplemented
615 with 10 μM Y-27632 and 2μM Thiazovivin. On day 3 after sort, 125μl of mTeSR-1 was added to
616 each well. The wells were monitored afterwards for colony formation. On day 6, rtTA-expressing
617 PGP1 hiPSCs previously generated[64] were transfected using Lipofectamine 3000 (Thermo Fisher
618 Scientific) with Super PiggyBac Transposase (System Biosciences) and the PiggyBac transposon
619 vector with hGATA6-2A-EGFP under control of the tetracycline responsive element promoter.
620 Transfected cells were selected by adding 0.5mg/mL puromycin to the mTeSR1 maintenance
621 medium. PGP9 iGATA6 hiPSCs were engineered as explained previously[64]. For the generation
622 of GATA6-hi cell line, the iGATA6 cell line was sorted via FACS (one cell per each well of a 96
623 well plate) in mTeSR-1 supplemented with 10 μM Y-27632 and 2μM Thiazovivin. The media was
624 replaced one the day after sorting with mTeSR-1 supplemented with 10 μM Y-27632 and 2μM
625 Thiazovivin. On day 3 after sort, 125μl of mTeSR-1 was added to each well. The wells were
626 monitored afterwards for colony formation. On day 6, the wells with a considerable colony were
627 passaged and the level of GATA6-EGFP was characterized by dox induction.
628
629 **Generation of iDiscoid**
630 The GATA6-engineered hiPSCs were seeded at a ratio of 81:5 with rtTA expressing PGP1
631 hiPSCs either containing or lacking an mKate reporter gene at a total density of 54,000 cells per
632 cm^2 in mTeSR-1 supplemented with 10 μM Y-27632. The next day, the medium was changed to
633 mTeSR-1 with 1 μg/mL doxycycline to induce expression of the GATA6 transgene, and this
634 medium used for daily replacement for up to 5 days. For experiments continued beyond 5 days,
635 medium was switched to IMDM on day 6, and half media changes were subsequently preformed
636 daily for up to 21 days. For the mTeSR condition in Supplementary Figure 13E, medium was
637 instead switched to mTeSR on day 6 and cultures were maintained for 12 days.
638
639 **iDiscoid Cryostorage**
640 Uninduced iDiscoid cells were incubated with Dispase at 70% confluence for 10 minutes,
641 or until visible lifting of colony edges. Cells were washed twice with DMEM/F-12, then colonies
642 were manually scraped from the plate. Colonies were centrifuged at 300g for 5 minutes, and
643 then were resuspended in Cryostor 10. Cells were cooled at -80°C for 24 hours prior to transfer
644 to liquid nitrogen for long-term storage. Cells were stored in liquid nitrogen for at least 24 hours
645 prior to defrosting.
646
647 **Signaling Pathway Inhibition**
648 BMP4 signaling was inhibited via application of 100ng/ml Noggin into the normal media
649 at day 2 of iDiscoid culture. This concentration of Noggin was supplemented into the media
650 changed on subsequent days.
651

652 **Staining on glass coverslips**

653 Cells were grown on Matrigel-coated 8mm diameter circular glass coverslips, on 14mm
654 coverslip bottom dishes (Mattek Corporation), or on 8 Well μ-Slide (ibidi). Cultures were fixed for
655 10 minutes in 4% paraformaldehyde (Electron Microscopy Sciences) at room temperature.
656 Coverslips were then washed three times with PBS followed by 15 minutes permeabilization with
657 0.2% Triton X-100 in PBS. Subsequently the coverslips were washed three times in wash buffer
658 (0.05% Tween-20 in PBS) for 5 minutes and blocked for 20 minutes in 200 ml wash buffer plus
659 5% normal donkey serum (Jackson ImmunoResearch Laboratories). The primary antibodies were
660 diluted in 5% normal donkey serum in PBS and incubated with the tissues 1 hour at room
661 temperature followed by three washes in wash buffer for 5 minutes each. The secondary
662 antibodies were diluted in 5% normal donkey serum in PBS and incubated with the tissues 1 hour
663 at room temperature followed by three washes in wash buffer for 5 minutes each. Afterwards,
664 the 8mm coverslips were mounted on microscopy glass slides using ProLong Glass Antifade (Life
665 Technologies), cured overnight at room temperature and then sealed with nail polish. Coverslips
666 in coverslip bottom dishes or μ-slides were stored in PBS at 4°C for 3D imaging.
667

668 **Image Acquisition and Processing**

669 Images were acquired using the EVOS M700 automated scanning microscope, Leica SP8
670 confocal microscope, Sartorius Incucyte Live Cell Imaging System, or Nikon A1 Confocal
671 microscope and processed using Fiji/ImageJ software (NIH)[65]. Any contrast adjustments were
672 made in individual channels and applied evenly across the whole image in that channel. Contrast
673 and color balance for color images was applied evenly across the whole image. 3D
674 reconstructions were generated using the Nikon A1 confocal microscope to generate z-stacks
675 spanning ~100um deep into the tissues and using Imaris (Bitplane) to construct a 3D volume
676 from the stacks. Time lapse videos were initially processed using Fiji/ImageJ, and annotations
677 were added using the Adobe Premiere software. For Supplementary Video 4, a video moving
678 through the Z-slices was initially recorded performing this action in the NIS Elements HC software
679 (Nikon), and the video was subsequently cropped and processed.
680

681 **Analysis of Wild-Type Cluster Areas and Radial Expression**

682 Tiled whole-coverslip images were cropped to a central circumscribed square (typically
683 9000μm^2) image for analysis. Wild-type compartment analysis was performed using an in-house
684 MATLAB (Mathworks) pipeline (built in MATLAB R2022b). In brief, wild-type compartments were
685 detected programmatically via thresholding the nuclear dye, f-actin stain, or mKate marker
686 channel for areas of high signal intensity. Maximum individual compartment area and average
687 cell size were defined manually. Compartments were filtered to remove those with areas close
688 to the defined individual cell area via the following equation:
689

690 $$lowerSizeLimit = \frac{maxCompartmentArea}{10^{\left(\frac{maxCompartmentArea - singleCellArea}{2}\right)}}$$

691

692 For compartment counting and wild-type area analysis, characteristics of each
693 compartment were recorded using the regionprops function in MATLAB. Additionally, the
694 distance to the nearest wild-type compartment was recorded.
695 Analysis of marker expression inward from the compartment perimeter was performed
696 via drawing a line a defined distance towards each compartment's centroid and recording the
697 intensity value of each marker at each point within this range. If a line was drawn such that it
698 would intersect the compartment edge (e.g., drawing from the outer edge of a concave shape),
699 those values were skipped. Intensity values recorded for each line were averaged to create final
700 per-compartment expression distributions. These distributions were aligned based on the point
701 of maximum CER1 expression when present, and the lengths of the distributions were equalized
702 to the length of the longest border recorded using the imresize function in MATLAB.
703
704 **Analysis of Wild-Type Cluster Features**
705 Coverage of WT clusters was evaluated via widefield images taken of the entire coverslip.
706 Areas in which EGFP expression was observed inside of WT clusters greater than 50% of the
707 cluster area were counted as being covered. Presence or absence of lumens/cavities was
708 evaluated manually, and was detected via the presence of visible rings of PODXL, ZO-1, or F-
709 actin expression within wild-type compartments.
710 ISL1 expression within the WT disc was recorded via averaging the radial expression
711 values obtained from the protocol above.
712 Polarity of TBXT and CER1 domains around WT compartments were assessed in the
713 following way: raw values of CER1 and TBXT produced by radial cluster analysis were normalized
714 to cell density by dividing by the F-actin intensity measured at each point. The average of CER1
715 and T values at all points within 1/8 of the disc radius in either direction away from a given point
716 around each disc was computed for every point recorded. Positivity for the marker was assessed
717 by comparing the minimum and maximum rolling average values for a difference of greater than
718 0.1 normalized intensity (a.u.). For islands with positivity, the radial index of the highest averaged
719 quadrant value of CER1 was compared to the radial index of the highest average quadrant value
720 of TBXT: a WT compartment with an average peak of TBXT between 135° and 225° away from
721 the CER1 peak was considered to be anti-polar; a WT compartment with a peak of TBXT within
722 45° of the CER1 peak was considered to be syn-polar. Polarity type assignments per-island were
723 verified manually by eye.
724
725 **Analysis of Hemogenic Endothelial and Hematopoietic Marker Genes**
726 Whole-coverslip immunofluorescence images of each marker were cropped into a
727 square, then divided into equivalent quadrants for ease of analysis. WT cluster masks were
728 created via manual thresholding and creation of a binary image using Fiji/ImageJ. These binary
729 images were then used to remove WT clusters from the images via the Image Calculator function
730 applied to the individual channel images. Resulting images were evaluated via using a custom
731 pipeline in CellProfiler[66]. Scatterplots shown were created using the Mean Object Intensity values
732 for each detected cell. CD43 bar graphs were produced via counts of the number of objects
733 detected by the IdentifyPrimaryObjects function applied to the CD43 channel in isolation.
734 Resulting images were evaluated via using a custom pipeline in CellProfiler[66]. Scatterplots shown

735 were created using the Mean Object Intensity values for each detected cell. CD43 bar graphs
736 were produced via counts of the number of objects detected by the IdentifyPrimaryObjects
737 function applied to the CD43 channel in isolation.
738 CD34$^+$ area was assessed via Fiji/ImageJ. The cropping and WT cluster masking steps
739 above were followed. Subsequently, the Threshold command was used to select areas positive
740 for CD34, and thresholded area was assessed via the Measure command. The area of the
741 threshold was then divided by the total cropped image area to determine percent CD34
742 coverage for each image.
743
744 **Quantitative Analysis of Z-Distribution of Blood Island-related Markers**
745 Confocal z-stack images were captured near to the center of each identified area. For
746 each stack, the bottom was defined as the lowest z-index where an in-focus marker could be
747 identified. The sum of the pixel intensity values from all pixels within each slice above the bottom
748 index were summed, then divided by pixel number. The resulting values at each corresponding
749 z-index were then averaged between all samples identified; these values were then converted
750 to a percentage of the max value to produce the graphs shown.
751
752 **Enzyme-linked Immunosorbent Assays (ELISA)**
753 Samples were assayed for AFP and APOA1 using commercially available ELISA kits (abcam).
754 Sample dilutions were optimized to attain detection in the linear range of the standard curves
755 for each individual assay.
756
757 **10x Genomics Sample Preparation for Next-generation Sequencing**
758 Samples were prepared as described by the 10x Genomics Cell Multiplexing Oligo
759 Labeling for Single Cell RNA Sequencing Protocols for cells with >80% viability. iDiscoids of each
760 day indicated were acquired in single cell suspension by incubation with Accutase for 20 minutes
761 at 37°C. The cell suspension was passed through a 40µm strainer in order to remove aggregates.
762 Each sample suspension was then centrifuged at 300g for 5 minutes at room temperature. The
763 supernatant was manually aspirated from each of the samples using a P1000 pipette. The cell
764 pellets were each resuspended in 1ml PBS+0.04% BSA, and the samples were centrifuged at
765 300g for 5 minutes at room temperature. The samples were then resuspended, targeting 1×10^6
766 cells/ml based on expected densities of cells at each day, and the cell suspension was passed
767 through a 40um strainer again. Each cell suspension was counted using a hemocytometer, and
768 a volume of cell suspension was removed as necessary to adjust the final cell count to 1×10^6 cells.
769 The volumes were then adjusted to 1ml by adding additional PBS+0.04% BSA to replace the
770 removed volumes.
771 *For Multiplexing Oligo Labeling:* Samples were then centrifuged at 300g for 5 mins at
772 room temperature. Supernatant was aspirated manually with a P1000 pipette. Samples were
773 resuspended in 50µl of CellPlex Multiplexing Solution (10x Genomics), with unique multiplexing
774 oligo solutions assigned to each sample. A maximum of 12 samples were labeled in parallel.
775 Samples were incubated for 5 minutes, starting after the oligo solution was added to the last
776 sample.

Following labeling, 1.95ml 1x PBS + 1% BSA was added to each sample, and the solution was mixed thoroughly. Each sample was centrifuged at 300g for 5 minutes at 4°C. The supernatant was manually aspirated with a P1000 pipette, taking care to leave less than 10µl of supernatant remaining in each sample when possible. Tthe samples were resuspended in 2ml 1x PBS + 1% BSA and were mixed thoroughly to wash. This wash and centrifugation step was repeated two additional times; at the last resuspension, enough wash buffer was added to reach a cell count of 1×10^6 cells/ml, assuming 50% cell loss from the count at the end of the first protocol. Labeled cell suspensions were then put on ice for transfer to the Pitt Single Cell Core for library creation.

Following a final counting and viability analysis, cells and 10x Genomics reagents were loaded into the single cell cassette, with a target of 25000 single cells for analysis, accounting for predicted cell loss and doublets resulting from multiplexing as laid out in the user guide for the Chromium Single-Cell 3' Reagent Kit (10x Genomics). After generation of GEMs, the cDNA library was prepared by Pitt Single Cell Core staff following the appropriate steps determined by the 10x Genomics user guide. Libraries were sent to the UPMC Genome Center for sequencing by a NovaSeq S4-200 for an intended read depth of 100,000 reads per cell with 150 bp paired end reads. Our downstream analysis from the sequencing data yielded between 40k-150k mean reads per cell in different samples.

scRNA-seq Sample Processing and QC

The 10x Genomics CellRanger pipeline was used to align reads to the reference genome (GRCh38.84) appended with transgene sequences, to assign reads to individual cells, and to estimate gene expression based on UMI counts[67].

Single cell data was excluded based on high mitochondrial genome transcript ratio and either high or low feature or UMI counts. Genes with UMI counts in fewer than 5 cells were removed from consideration. For scRNA-Seq data processing and cluster analysis using Seurat[68,69], we used the following general standardized pipeline for processing of the Cell Ranger output: SCTransform regressing percent mitochondrial genes, principal component analysis (PCA), and clustering. Jackstraw plots and permuted p-values were used to assist in determining the optimal number of principal components (PCs) needed to summarize the datasets without losing a significant amount of variation. The quality of a range of clustering resolution values was assessed using enrichment of cluster marker genes (genes differentially upregulated in a given cluster relative to all other clusters) with embryo cell type-specific genes. As a quality check, PC and resolution metrics were modulated to yield fewer or additional clusters to confirm that chosen parameters resulted in the most biologically relevant clustering. Visualization was achieved by the use of UMAP plots identifying cells, clusters, and selected gene expression in each cell, as well as heatmaps and violin plots showing the expression level of genes by cluster.

Subclustering the WT in the D3 dataset was achieved by subsetting the WT cluster, then computing the distance between rows of the data matrix using Euclidean measurement (dist function in R). Hierarchical clustering was applied using this distance matrix (hclust in R), and used to create a dendrogram of the distance explained by different numbers of subclusters. The number of subclusters was selected based on the highest level of distance observed and applied

819 to the data (clustercut in R). The cluster was then added back to the rest of the dataset and
820 markers were determined.
821
822 **Statistical analysis**
823 *Similarity quantification between lists of marker genes*: We quantify the similarity
824 between two lists of marker genes A and B by using the Hypergeometric test for
825 overrepresentation. This is equivalent to a one-tailed Fisher's Exact Test.
826
827 *Statistical Tests for Bar Graphs:* The statistical test used to calculate p-value for the bar
828 graph in Figure 4I was an unpaired, two-tailed t-test. The statistical test used to calculate p-value
829 for the bar graph in Supplementary Figure 13G is a one-way ANOVA with Dunnett's multiple
830 comparisons test.
831
832 ***in vitro* Colony-Forming Unit Assays**
833 Cells were dissociated into single cell solution using accutase (StemCell Technologies).
834 $CD34^+$ cells were purified from a portion of the cells using the CD34 Microbead Kit Ultrapure
835 and LS columns (Miltenyi Biotec) according to the manufacturer's instructions. Colony forming
836 unit assays were performed using methylcellulose containing media (MethoCult H4434 Enriched
837 and MethoCult SF H4636, StemCell Technologies) according to the manufacturer's instructions.
838 After 14 days in culture, cells were analyzed immunophenotypically via flow cytometry.
839
840 **Flow Cytometry Analysis**
841 Input and output cells to the colony forming unit assays were analyzed
842 immunophenotypically via flow cytometry. Cells were stained with human anti-CD235a, anti-
843 CD71, anti-CD45, anti-CD11b, and anti-CD34 antibodies (Biolegend) for 30 minutes on ice. Cells
844 were analyzed using an LSR II flow cytometer (BD Biosciences) using 7-AAD (ThermoFisher
845 Scientific) as a dead cell stain. Data analysis was performed using FlowJo software. An example
846 of gating strategy can be found in Supplementary Figure 21.
847 To prepare samples, iDiscoids were developed in 4-6 wells of a 6 well plate, and samples
848 were pooled prior to $CD34^+$ cell isolation using the MACS separation system (Stemcell Tech).
849 For flow cytometry analysis of $CD34^+$ cells on D5 and D12, iDiscoids were seeded on three
850 wells of a 12 well plate. On the day of harvest, accutase was used to detach the cells. Cells were
851 stained with anti-CD34 (APC) (Biolegend) at the concentration of 1:600 for 30 minutes on ice.
852 Cells were analyzed using an LSR II flow cytometer (BD Biosciences) using 7-AAD (ThermoFisher
853 Scientific) as a dead stain.
854
855 **Data Availability**
856 The datasets generated and/or analyzed during the current study are available from the
857 corresponding author on reasonable request.
858
859 **Code Availability**
860 All scripts used have been deposited at https://github.com/AmirAlavi/GATA6_R and are
861 available from the corresponding author upon request.

Methods References

61 Daley, G. Q. *et al.* Setting Global Standards for Stem Cell Research and Clinical Translation: The 2016 ISSCR Guidelines. *Stem Cell Reports* **6**, 787-797, doi:https://doi.org/10.1016/j.stemcr.2016.05.001 (2016).

62 Lovell-Badge, R. *et al.* ISSCR Guidelines for Stem Cell Research and Clinical Translation: The 2021 update. *Stem Cell Reports* **16**, 1398-1408, doi:https://doi.org/10.1016/j.stemcr.2021.05.012 (2021).

63 Guye, P. *et al.* Genetically engineering self-organization of human pluripotent stem cells into a liver bud-like tissue using Gata6. *Nature Communications* **7**, 10243, doi:10.1038/ncomms10243 (2016).

64 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).

65 Stirling, D. R. *et al.* CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**, 433, doi:10.1186/s12859-021-04344-9 (2021).

66 Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nature Communications* **8**, 14049, doi:10.1038/ncomms14049 (2017).

67 Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411-420, doi:10.1038/nbt.4096 (2018).

68 Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology* **33**, 495-502, doi:10.1038/nbt.3192 (2015).

Acknowledgements

This work was supported by a Startup Fund from the Department of Pathology at University of Pittsburgh. We acknowledge support from the Center for Biological Imaging and the University of Pittsburgh Flow Cytometry Core facilities. Characterization of hematopoietic cellular subsets was partly supported by an R01 from the National Heart Lung and Blood Institute (HL141805) to M.R.E. Development and characterization of the iPSC lines used in these studies were partly supported by an R01 from the National Institute of Biomedical Imaging and Bioengineering (EB028532) and NSF award # 2134999 to M.R.E. The computational analysis of published embryo datasets was partly supported by NIH grants 1R01GM122096, OT2OD026682, 1U54AG075931 and 1U24CA268108 to Z.B.J. This work is also partly supported by the Pittsburgh Liver Research Center (NIHNIDDK P30DK120531), the Cellular Approaches to Tissue Engineering and Regenerative Medicine Training Grant (T32 EB001026), and NIH 1S10OD019973-01. We would like to thank Katherine Helfrich and Michael Calderon at the Center for Biological Imaging for their assistance related to imaging. We also thank Alaina Sutherland, Chloe Hislop, and Raini Porterfield for helpful discussion during this work. Biorender was used to create several of the schematics shown.

Author Contributions

J.H. and M.R.E. conceived the study; M.R.E., Z.B.J., B.S., and S.K. supervised the project; J.H., M.R.E., S.K., A.A., K.K.F., J.V., R.L., D.S., Z.B.J. conceived the methodology for experiments; J.H., R.S., K.K.F., J.V., R.L., M.N.T., and M.R. performed iDiscoid experiments; J.H., R.S., K.K.F., T.M., M.N.T., S.W. and D.S. performed imaging and analysis; J.H., A.A., and Q.S. performed the bioinformatic analysis of scRNA-seq datasets; R.L. and K.K.F. performed flow cytometry and analysis; J.H., A.A., K.K.F., J.V., R.L., D.S., M.R.E. performed visualization of the data; M.R.E., S.K., Z.B.J., and J.H. acquired funding for the project; J.H. and M.R.E. wrote the original draft of the manuscript; J.H., R.S., K.K.F., M.R.E., B.S., S.M.C.S.L., A.A., J.V., S.K., and Z.B.J. were involved in editing and input to the final draft of the manuscript.

Competing Interests:

The authors declare no conflicts of interest. J.H., S.K., and M.R.E. have filed for IP for the technology presented here.

Figure 1

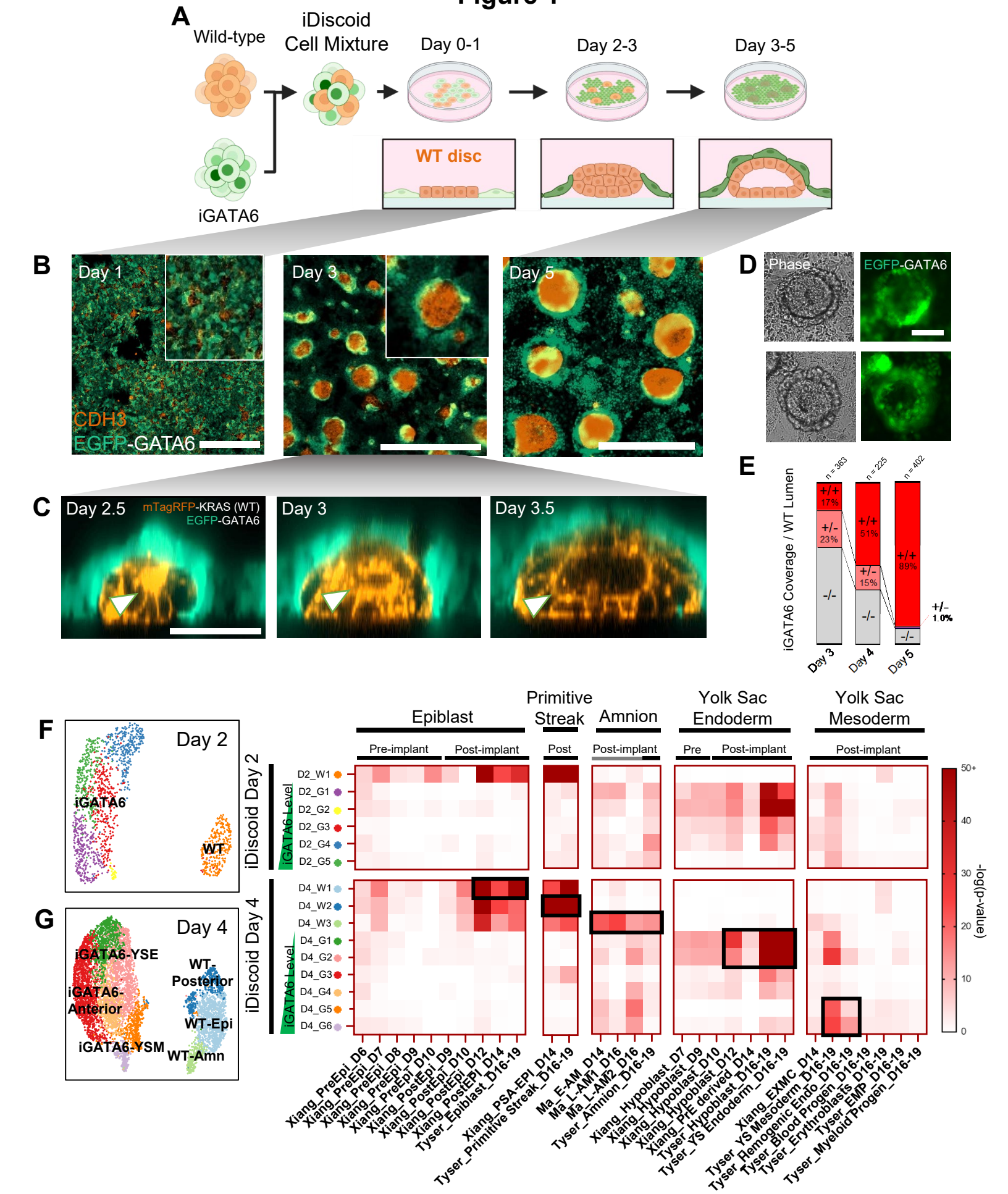

A
Wild-type
iDiscoid Cell Mixture
Day 0-1
Day 2-3
Day 3-5
iGATA6
WT disc
B
Day 1
Day 3
Day 5
CDH3
EGFP-GATA6
C
Day 2.5
mTagRFP-KRAS (WT)
EGFP-GATA6
Day 3
Day 3.5
D
Phase
EGFP-GATA6
E
iGATA6 Coverage / WT Lumen
n = 363
n = 225
n = 402
+/+ 17%
+/- 23%
-/-
+/+ 51%
+/- 15%
-/-
+/+ 89%
+/- 1.0%
-/-
Day 3
Day 4
Day 5
F
Day 2
iGATA6
WT
G
Day 4
iGATA6-YSE
iGATA6-Anterior
iGATA6-YSM
WT-Posterior
WT-Epi
WT-Amn
iDiscoid Day 2
iDiscoid Day 4
iGATA6 Level
D2_W1
D2_G1
D2_G2
D2_G3
D2_G4
D2_G5
D4_W1
D4_W2
D4_W3
D4_G1
D4_G2
D4_G3
D4_G4
D4_G5
D4_G6
Epiblast
Primitive Streak
Amnion
Yolk Sac Endoderm
Yolk Sac Mesoderm
Pre-implant
Post-implant
Post
Pre
-log(p-value)
50+
40
30
20
10
0
Xiang_PreEpi_D6
Xiang_PreEpi_D7
Xiang_PreEpi_D8
Xiang_PreEpi_D9
Xiang_PreEpi_D10
Xiang_PostEpi_D9
Xiang_PostEpi_D10
Xiang_PostEpi_D12
Xiang_PostEpi_D14
Tyser_Epiblast_D16-19
Xiang_PSA-EPI_D14
Tyser_Primitive Streak_D16-19
Ma_E-AM_D14
Ma_L-AM1_D16
Ma_L-AM2_D16
Tyser_Amnion_D16-19
Xiang_Hypoblast_D7
Xiang_Hypoblast_D9
Xiang_Hypoblast_D10
Xiang_Hypoblast_D12
Xiang_PrE derived_D14
Tyser_Hypoblast_D16-19
Tyser_YS Endoderm_D16-19
Xiang_EXMC_D14
Tyser_YS Mesoderm_D16-19
Tyser_Hemogenic Endo_D16-19
Tyser_Blood Progen_D16-19
Tyser_Erythroblasts_D16-19
Tyser_EMP_D16-19
Tyser_Myeloid Progen_D16-19

Figure 1: Engineering co-development of embryonic and extra-embryonic endoderm tissues in iDiscoids

(A) Schematic demonstrating cell mixing and subsequent organization of iGATA6 (green) and WT (strongly orange) hiPSCs after GATA6 induction by Dox.

(B) Immunofluorescence staining demonstrating self-organization of iDiscoids in the cultures. Scale bar = 1000 μm

(C) Live time lapse images of one iDiscoid showing growth of a central lumen from a rosette indicated by an arrow. Scale bar = 50 μm.

(D) Phase and fluorescence images from live day 4 iDiscoid cultures showing developed iDiscoid morphologies with EGFP-expressing iGATA6 around a WT cluster. Scale bar = 100 μm

(E) Quantification of characteristics of WT clusters possessing different iGATA6 coverage and lumen formation characteristics. No islands without coverage were observed possessing a lumen (-/+ = 0%)

(F) Single cell UMAP and hypergeometric statistical comparisons of differentially expressed gene (DEG) lists between day 2 iDiscoid and human and NHP embryo single cell datasets. Labels above heatmaps indicate pre-implantation versus post-implantation embryo sample comparisons; gray bars under labels indicated NHP comparisons.

(G) Single cell UMAP and hypergeometric statistical comparisons of differentially expressed gene (DEG) lists between day 4 iDiscoid and human and NHP embryo single cell datasets. Black boxes highlight high DEG similarity (high P values) to relevant populations of interest.

Figure 2

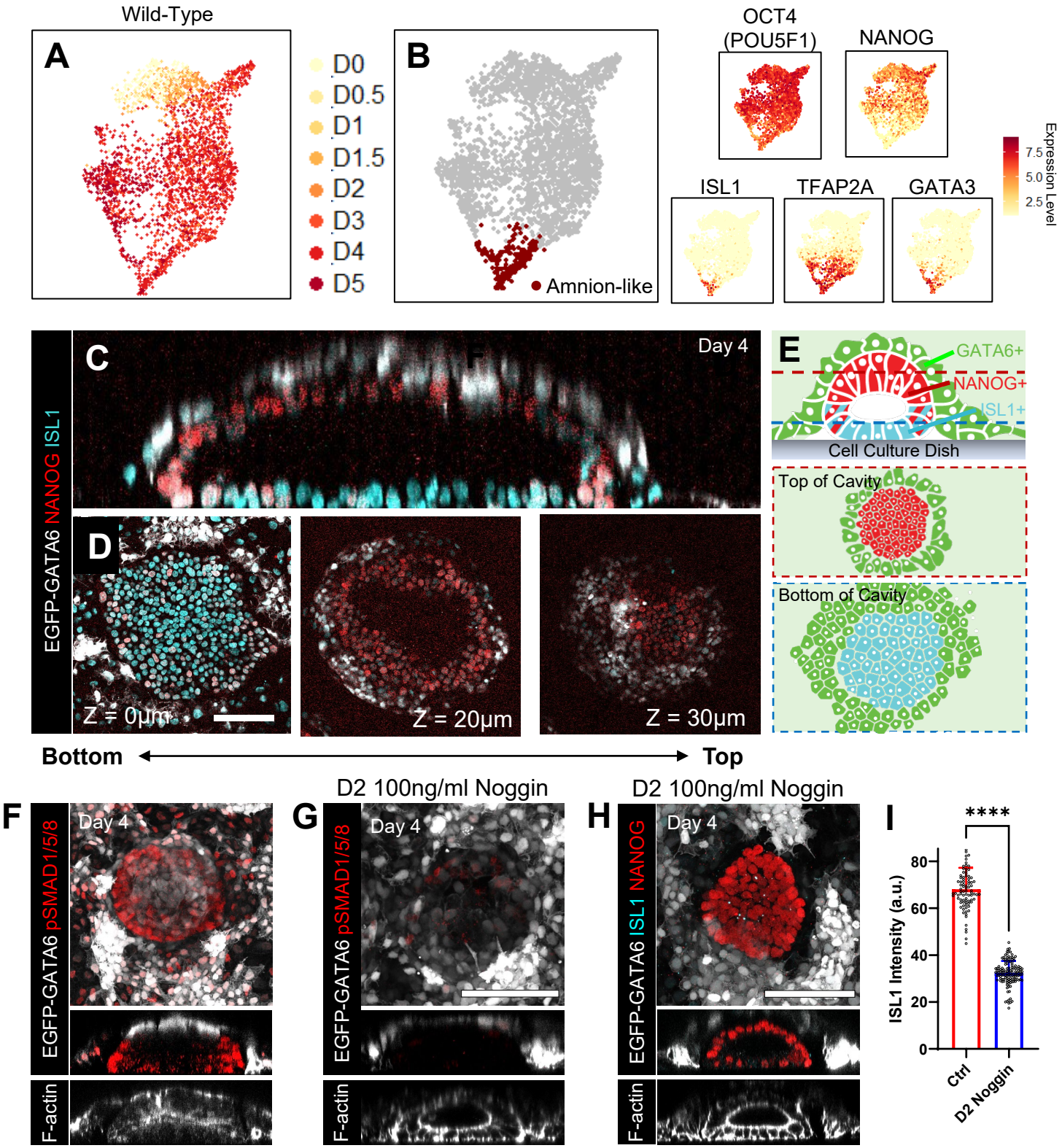

Wild-Type
A
D0
D0.5
D1
D1.5
D2
D3
D4
D5
B
Amnion-like
OCT4 (POU5F1)
NANOG
ISL1
TFAP2A
GATA3
Expression Level
7.5
5.0
2.5
C
Day 4
EGFP-GATA6 NANOG ISL1
D
Z = 0μm
Z = 20μm
Z = 30μm
E
GATA6+
NANOG+
ISL1+
Cell Culture Dish
Top of Cavity
Bottom of Cavity
Bottom
Top
F
Day 4
EGFP-GATA6 pSMAD1/5/8
F-actin
G
D2 100ng/ml Noggin
Day 4
EGFP-GATA6 pSMAD1/5/8
F-actin
H
D2 100ng/ml Noggin
Day 4
EGFP-GATA6 ISL1 NANOG
F-actin
I
****
ISL1 Intensity (a.u.)
80
60
40
20
0
Ctrl
D2 Noggin

Figure 2: Amniotic-like cavity formation and expansion

(A) Merged UMAP of all WT lineages from iDiscoid labeled by day of development.

(B) The merged WT population showing the compartment expressing markers of amnion. *ISL1*, *TFAP2A* (AP-2$^{\alpha}$), and *GATA3* are expressed in this area, while *NANOG* is negative and OCT4 is low.

(C) An orthogonal slice of an individual iDiscoid showing top-bottom compartmentalization of ISL1^{+} (cyan) and NANOG^{+} (red) cells while iGATA6 cells cover the top (white).

(D) Horizontal z-slices at the indicated distance from the dish of the WT cluster from (C).

(E) Schematic showing the position of each population within a single iDiscoid. Dotted lines indicate the area of slices shown.

(F) Expression patterns of BMP4 effectors (phosphorylated SMAD1, SMAD5, and SMAD8/9) in iDiscoid. Lower images show a lateral slice of the WT disc shown.

(G) Immunofluorescence staining for the BMP4 effectors (phosphorylated SMAD1, SMAD5, and SMAD8/9) in iDiscoid at day 4 after application of the inhibitor Noggin at Day 2 of development. Lower images show a lateral slice of the WT disc shown.

(H) Immunofluorescence staining for ISL1 and NANOG in iDiscoid at day 4 after application of Noggin at Day 2 of development. Lower images show a lateral slice of the WT disc shown.

(I) Scatterplot showing ISL1 expression (D4) intensities within WT clusters; control (Ctrl) iDiscoid versus BMP4 inhibition (Noggin) conditions. n[control] = 87, n[Noggin] = 134, **** $p<0.0001$

Scale bars = 100 µm. n represents separate biological samples from a single experiment harvested on the corresponding day.

Figure 3

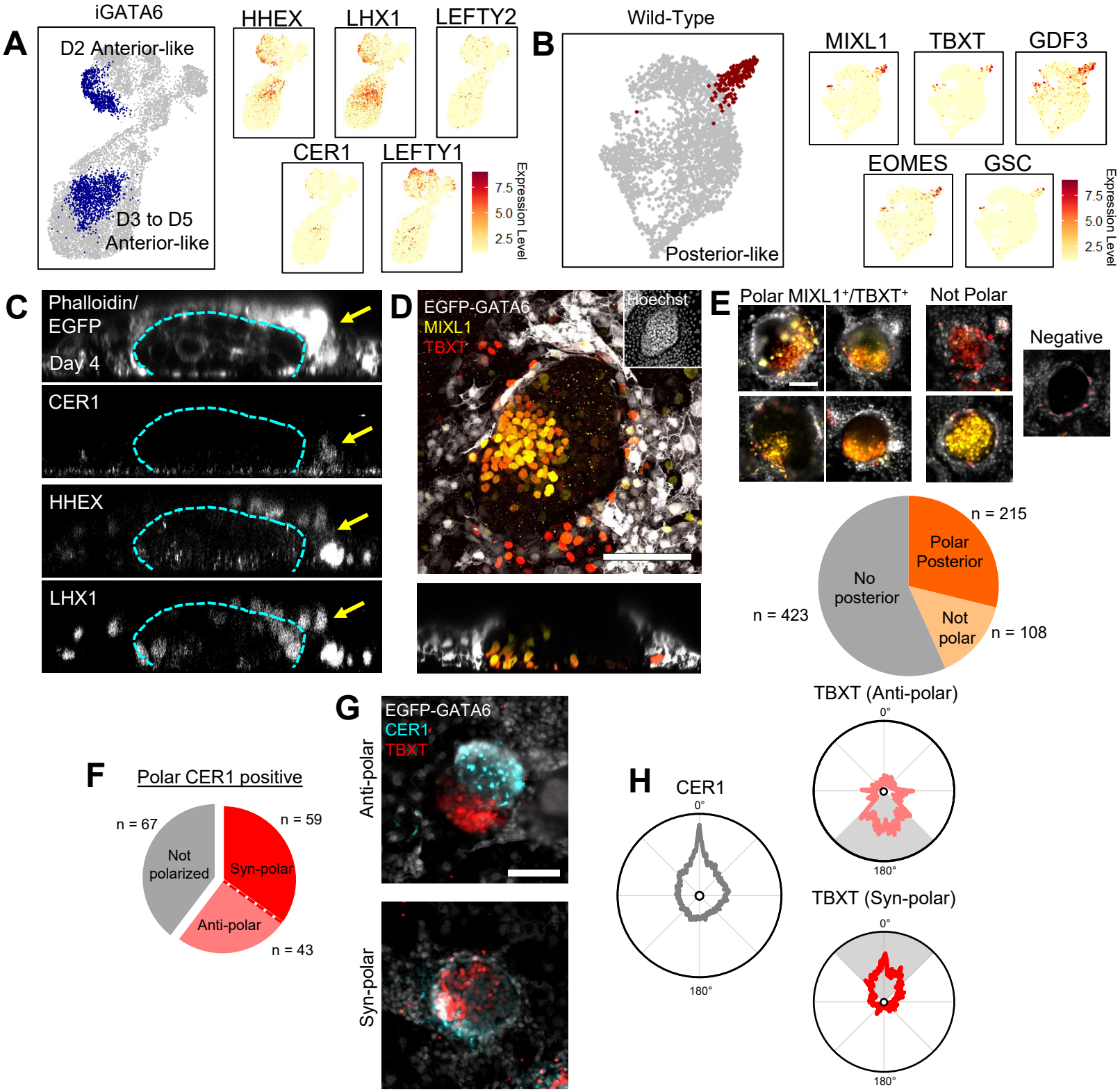

A
iGATA6
D2 Anterior-like
D3 to D5
Anterior-like
HHEX
LHX1
LEFTY2
CER1
LEFTY1
7.5
5.0
2.5
Expression Level
B
Wild-Type
Posterior-like
MIXL1
TBXT
GDF3
EOMES
GSC
7.5
5.0
2.5
Expression Level
C
Phalloidin/
EGFP
Day 4
CER1
HHEX
LHX1
D
EGFP-GATA6
MIXL1
TBXT
Hoechst
E
Polar MIXL1+/TBXT+
Not Polar
Negative
n = 215
Polar
Posterior
No
posterior
n = 423
Not
polar
n = 108
F
Polar CER1 positive
n = 67
Not
polarized
n = 59
Syn-polar
Anti-polar
n = 43
G
EGFP-GATA6
CER1
TBXT
Anti-polar
Syn-polar
H
CER1
0°
180°
TBXT (Anti-polar)
0°
180°
TBXT (Syn-polar)
0°
180°

Figure 3: Anterior hypoblast-like domain and posterior axis in iDiscoids

(A) Merged UMAP of all iGATA6 lineages showing the compartments expressing markers of anterior hypoblast. Two separate domains of these markers were observed, one within day 2 iDiscoid cells and one within day 3-5 iDiscoid cells. (UMAP labeled by day can be seen in Supplementary Figure 5)

(B) Merged UMAP of all WT lineages showing the compartment expressing markers of the posterior axis and primitive streak.

(C) Immunofluorescence staining showing a z-slice of an iDiscoid with a HHEX/LHX1/CER1 co-positive domain adjacent to a WT clusters. Dotted line indicates boundaries of the WT cluster in each image, arrow indicates co-positive domain.

(D) Immunofluorescence staining showing a TBXT/MIXL1 co-positive domain within the WT clusters of the iDiscoid.

(E) Examples of WT clusters with polarized TBXT/MIXL1 domains (top) or without co-expression of markers (bottom). Venn diagram shows proportions of cluster types observed in iDiscoid cultures, n = 746 total.

(F) Venn diagram showing the proportion of clusters with a particular TBXT/CER1 polarity type when CER1 confined to one pole is present. n = 169 total.

(G) Representative WT clusters showing TBXT and CER1 expression domains at opposing poles of the cluster (anti-polar, top) or at the same pole of the cluster (syn-polar, bottom).

(H) Diagrams indicating the average radial expression patterns of TBXT in each polarity pattern from WT clusters with a polarized CER1-expressing domain indicated in (F). All diagrams are scaled to the same expression intensity value (1 a.u.). Degrees indicate radial distance around the circularized perimeter of a WT cluster from the CER1 peak shown. Shaded areas indicate region of highest average TBXT polarity corresponding to the polarity types.

Scale bars = 100 μm. n represents separate biological samples from a 3-4 separate experiments harvested on day 4 after iDiscoid induction.

Figure 4

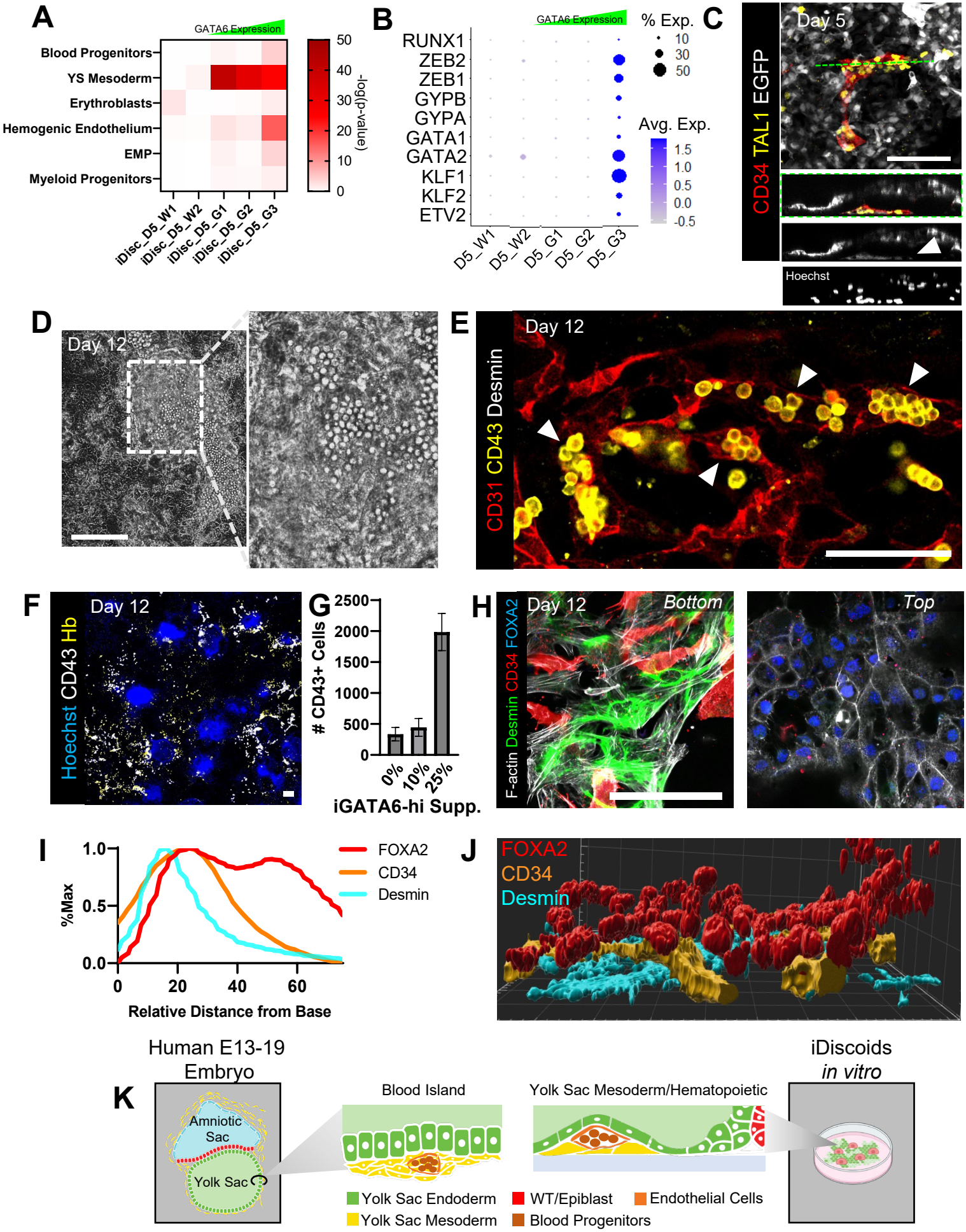

A
GATA6 Expression
Blood Progenitors
YS Mesoderm
Erythroblasts
Hemogenic Endothelium
EMP
Myeloid Progenitors
iDisc_D5_W1
iDisc_D5_W2
iDisc_D5_G1
iDisc_D5_G2
iDisc_D5_G3
-log(p-value)
50
40
30
20
10
0
B
GATA6 Expression
% Exp.
10
30
50
Avg. Exp.
1.5
1.0
0.5
0.0
-0.5
RUNX1
ZEB2
ZEB1
GYPB
GYPA
GATA1
GATA2
KLF1
KLF2
ETV2
D5_W1
D5_W2
D5_G1
D5_G2
D5_G3
C
Day 5
CD34 TAL1 EGFP
Hoechst
D
Day 12
E
Day 12
CD31 CD43 Desmin
F
Day 12
Hoechst CD43 Hb
G
CD43+ Cells
2500
2000
1500
1000
500
0
0%
10%
25%
iGATA6-hi Supp.
H
Day 12
Bottom
Top
F-actin Desmin CD34 FOXA2
I
%Max
1.0
0.5
0.0
0
20
40
60
Relative Distance from Base
FOXA2
CD34
Desmin
J
FOXA2
CD34
Desmin
Human E13-19 Embryo
K
Amniotic Sac
Yolk Sac
Blood Island
Yolk Sac Mesoderm/Hematopoietic
iDiscoids in vitro
Yolk Sac Endoderm
WT/Epiblast
Endothelial Cells
Yolk Sac Mesoderm
Blood Progenitors

Figure 4: iDiscoid hematopoietic lineages and blood island-like structures

(A) Heatmap showing day 5 iDiscoid scRNA-seq populations compared to hematopoietic populations from the human E16-19 embryo.

(B) Dot plot showing the expression pattern of hematopoietic markers in day 5 iDiscoid scRNA-seq populations.

(C) Immunofluorescence image showing the distribution of cells expressing CD34 and TAL1 (scl) in iDiscoid culture. Cells expressing TAL1 localize between the yolk sac endoderm compartment and the tissue culture dish and form arrangements of spindle-shaped cells. Orthogonal slice shows the position of spindle cells against the dish. Dashed line indicates the position from which the slice was taken.

(D) Live phase image taken at day 12 showing cells spherical cells generated within iDiscoid tissue. Dotted box indicates area of inset.

(E) Immunofluorescence image of day 12 iDiscoid showing the generation of $CD43^+$ spherical cells within $CD31^+$ endothelial cells.

(F) Immunofluorescence image of a day 12 iDiscoid with 25% GATA6-high cells substituted into the normal iDiscoid ratio.

(G) Histogram showing the number of $CD43^+$ cells detected in experiments in which the given percentages of GATA6-hi cells were substituted. n[0%] = 4 experimental replicates, n[10%] = 2 experimental replicates, n[25%] = 2 experimental replicates.

(H) Two slices from a single blood island-like structure demonstrating $Desmin^+$ mesoderm cells and $CD34^+$ endothelial cells localized underneath a $FOXA2^+$ endoderm.

(I) Histogram showing the z-distribution of the indicated markers between the bottom of the dish and the top of the culture. Bimodal distribution of FOXA2 indicates areas of expression outside of blood island-like foci. Distributions are averaged from 9 foci from 3 technical replicates.

(J) 3D reconstruction of one blood island area showing positioning of endoderm (FOXA2), endothelial (CD34), and mesoderm (Desmin) markers

(K) Schematic depicting in vivo embryonic yolk sac blood islands and in vitro iDiscoid yolk sac blood island-like structures.

Scale bars = 100 μm.

Extended Data Figures

Modelling Human Post-Implantation Development via Extra-Embryonic Niche Engineering

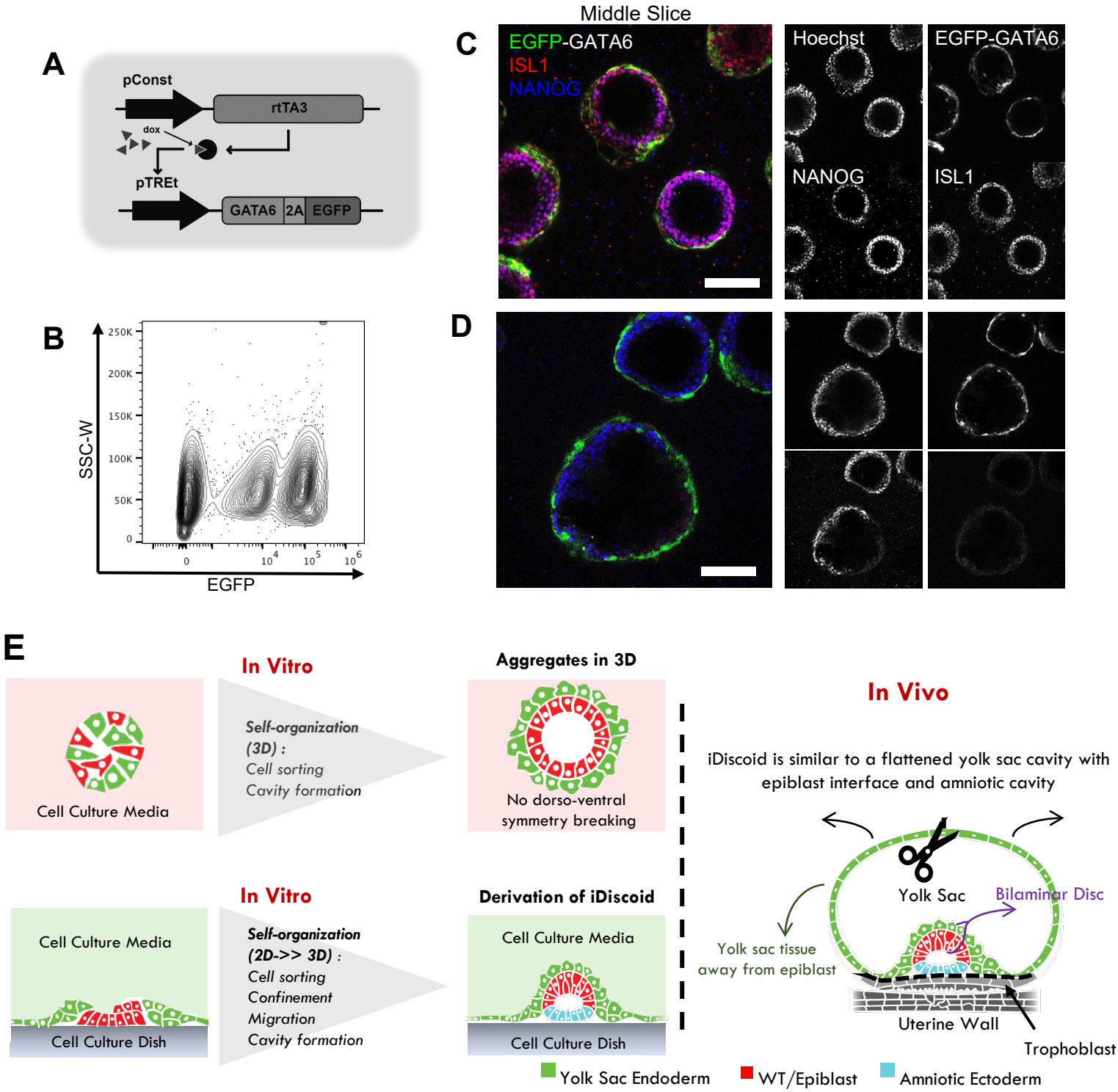
Joshua Hislop[1-3], Amir Alavi[4,5], Qi Song[4,5], Rayna Schoenberger[1,3], Kamyar Keshavarz F.[1-3], Ryan LeGraw[2,3], Jeremy Velazquez[2,3], Tahere Mokhtari[1-3], Mohammad Nasser Taheri[1-3], Matthew Rytel[2,3], Susana M Chuva de Sousa Lopes[8], Simon Watkins[6,7], Donna Stolz[6,7], Samira Kiani[1-3,9], Berna Sozen[10], Ziv Bar-Joseph[4,5], Mo R. Ebrahimkhani[1-3,9]*

Affiliations:

1. Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA
2. Department of Pathology, Division of Experimental Pathology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15213, USA
3. Pittsburgh Liver Research Center, University of Pittsburgh, Pittsburgh, PA 15261, USA
4. Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
5. Machine Learning Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA.
6. Center for Biologic Imaging, University of Pittsburgh, Pittsburgh, PA, USA.
7. Department of Cell Biology and Molecular Physiology, University of Pittsburgh, Pittsburgh, PA, USA
8. Department of Anatomy and Embryology, Leiden University Medical Center, Einthovenweg, 2333 ZC Leiden, the Netherlands
9. McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA
10. Department of Genetics, Yale School of Medicine, Yale University, New Haven, CT, 06510, USA.

*Correspondence: mo.ebr@pitt.edu

Extended Data Figure 1



Extended Data Figure 1: 3D iDiscoid cultures.

(A) The gene circuit used to create inducible GATA6-expressing iPSCs. pConst is a constitutively active promoter.

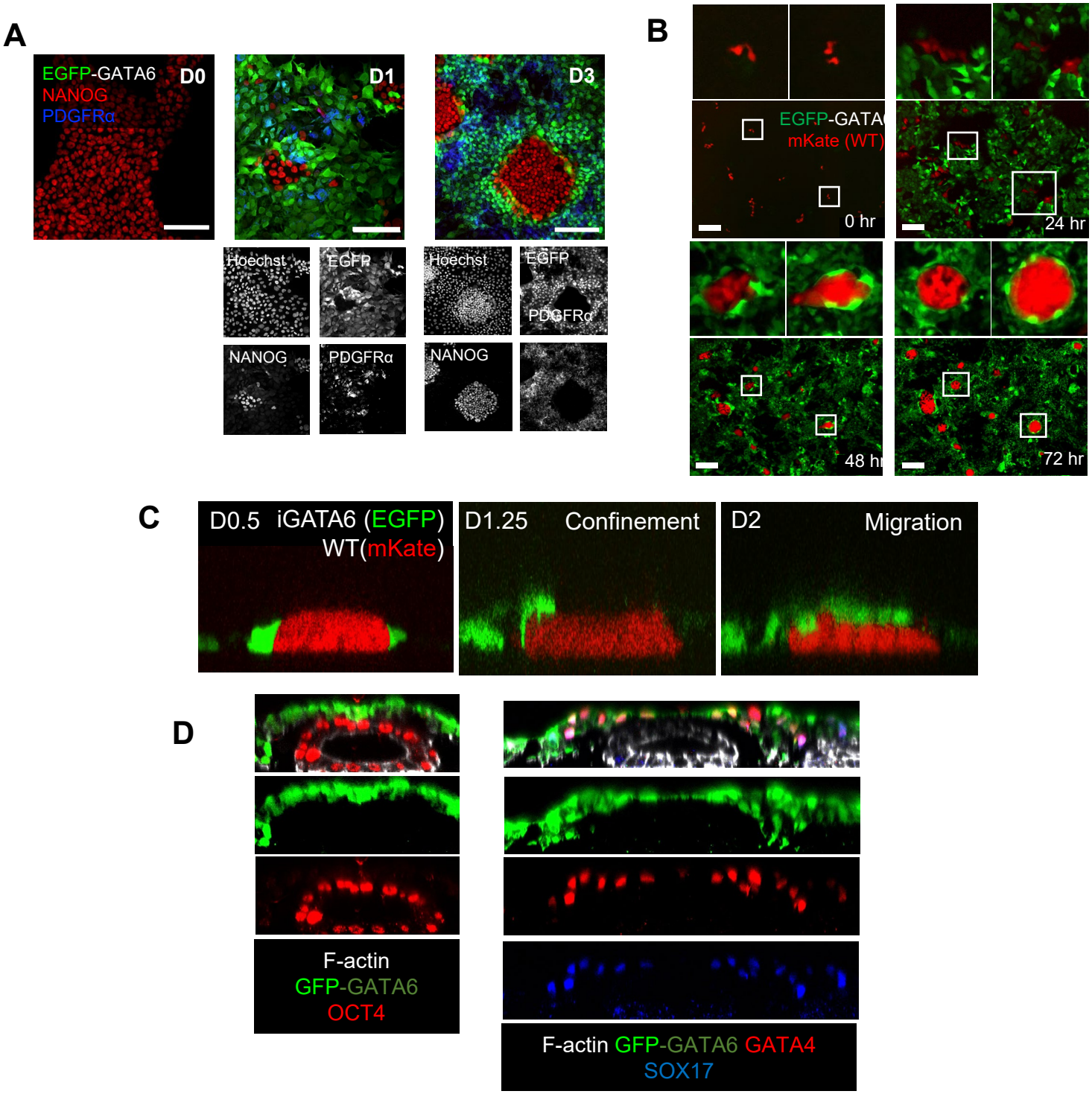
(B) Heterogeneity of EGFP (GATA6) activation in iGATA6 cells, detected via flow cytometry analysis. Higher gene circuit copy numbers leads to higher expression level of EGFP and GATA6.

(C) 3D culture of iGATA4 and WT showing co-expression of the amnion marker ISL1 and the pluripotency marker NANOG spread fully throughout the WT layer without D-V polarization. Middle slice shows the development of a central lumen.

(D) 3D culture of iGATA4 and WT showing expression of the pluripotency marker NANOG throughout the WT layer but lack of ISL1 expression in notable subset of these 3D tissues. These spheres do not exhibit polarization as well.

(E) Schematic depicting iDiscoid development in 3D versus from 2D>>3D in comparison to embryo morphology. iDiscoid is similar to a flattened yolk sac cavity with epiblast interface and amniotic cavity.

Extended Data Figure 2



Extended Data Figure 1: Sorting and symmetry breaking events following GATA6 induction.

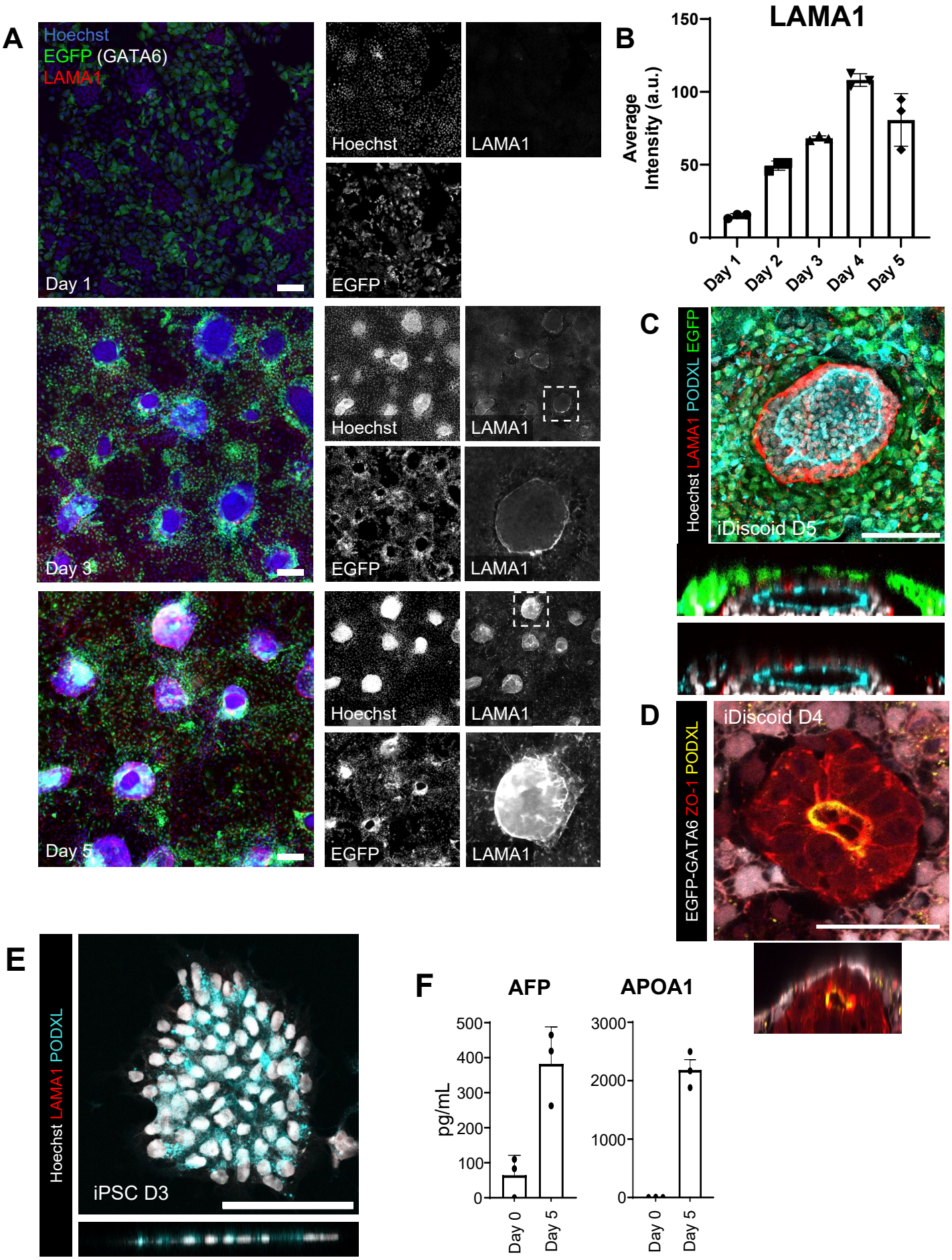
(A) Immunofluorescence images of fixed cultures demonstrating cell organization of iGATA6 (green) and WT (NANOG) hiPSCs between day 0 and day 3 after GATA6 induction. PDGFRα rises within the iGATA6 cells as they acquire a more yolk sac endoderm-like morphology.

(B) Time lapse images of a single position within the iGATA6/WT co-culture from day 0 to day 3 after GATA6 induction. Top cropped images correspond to positions within the white boxes in the images below. Scale bar = 200 μm.

(C) Time lapse images showing the initial confinement of red WT cells with green iGATA6 cells, followed by migration of iGATA6 cells over the WT disc prior to D2 after iDiscoid induction.

(D) Z-slices of two representative iDiscoids showing localization of OCT4, GATA4, and SOX17 within the bilaminar disc-like area of iDiscoid culture, as well as development of a central lumen by day 5.

Extended Data Figure 3



Extended Data Figure 3: Lumen development within WT of iDiscoids

(A) Immunofluorescence images showing dynamic of LAMA1 deposition on days 1, 3, and 5 of iDiscoid development after induction. Dotted boxes show the areas of inset in the fourth panels of the day 3 and day 5 images.

(B) Timecourse graphs showing the increase in LAMA1 signal in immunofluorescence images over the first five days of iDiscoid development. Dots represent randomly sampled areas from a single experiment harvested at each day

(C) Immunofluorescence image showing the deposition of laminin around a WT cluster with a central lumen as well as polarization of PODXL .

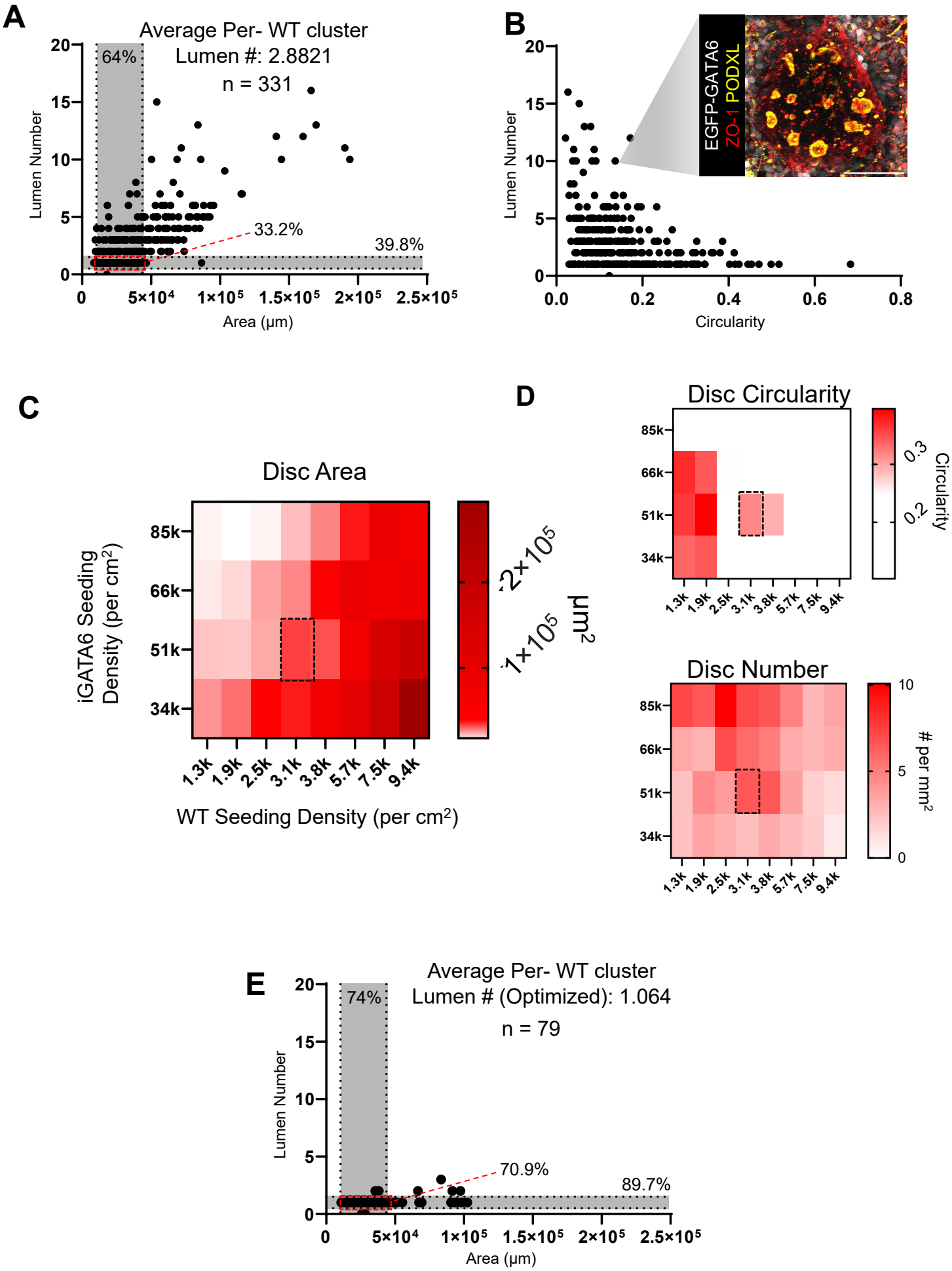
(D) Immunofluorescence showing horizontal and lateral slices of a representative WT cluster with polarization of PODXL and ZO-1 towards a central lumen.

(E) A representative cluster of WT iPSCs at day 3. No laminin deposition is observable in the vicinity of the cluster.

(F) ELISA comparing secreted AFP and APOA1 detected on D0 and D5 after GATA6 induction with Dox. Individual dots represent biological replicates.

Scale bars = 100μm

Extended Data Figure 4



Extended Data Figure 4: Lumen development within WT of iDiscoids

(A) Distribution of WT cluster areas versus the number of lumens observed in iDiscoids with iGATA6 coverage. Shaded area indicates the areas of the bilaminar disc between E9 and E17 as recorded by Hertig and Rock, 1949 and Heuser et al. 1945; 64% of discs observed have areas that fall into this range of values, and 39.8% have a single lumen. 33.2% fall into both of these categories. n = 331 total.

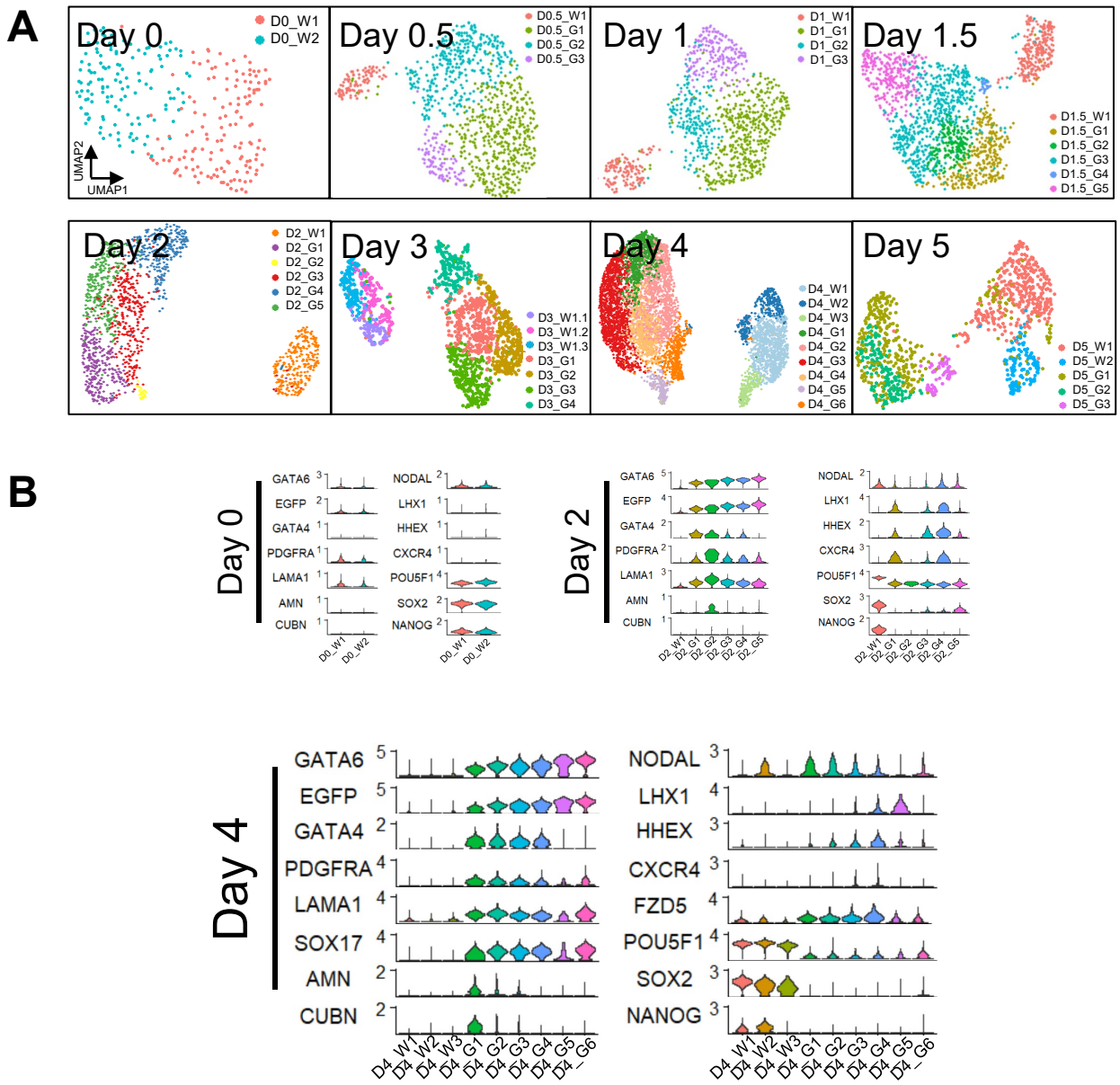
(B) Distribution of WT cluster circularity versus the number of lumens observed in iDiscoids with iGATA6 coverage, with a representative image of an iDiscoid with the indicated characteristics. n = 331, Scale bar = 100μm

(C) Heatmaps displaying the average area of iDiscoids resulting from different initial seeding densities of iGATA6 and WT. Dotted box shows optimized seeding density used for iDiscoid experiments.

(D) Heatmaps displaying the average circularity and average resulting disc number pre mm^2 of iDiscoids resulting from different initial seeding densities of iGATA6 and WT. Dotted box shows optimized seeding density used for iDiscoid experiments.

(E) Distribution of WT cluster areas versus the number of lumens observed in iDiscoids with iGATA6 coverage when seeded at the optimized ratio indicated in C. Shaded area indicates the areas of the bilaminar disc between E9 and E16-19 as recorded by Hertig and Rock, 1949 and Heuser et al. 1945; 74% of discs that fall into this area range. 89.7% of discs have a single lumen. 70.9% of discs fall into both of these categories. n = 79 total.

Extended Data Figure 5



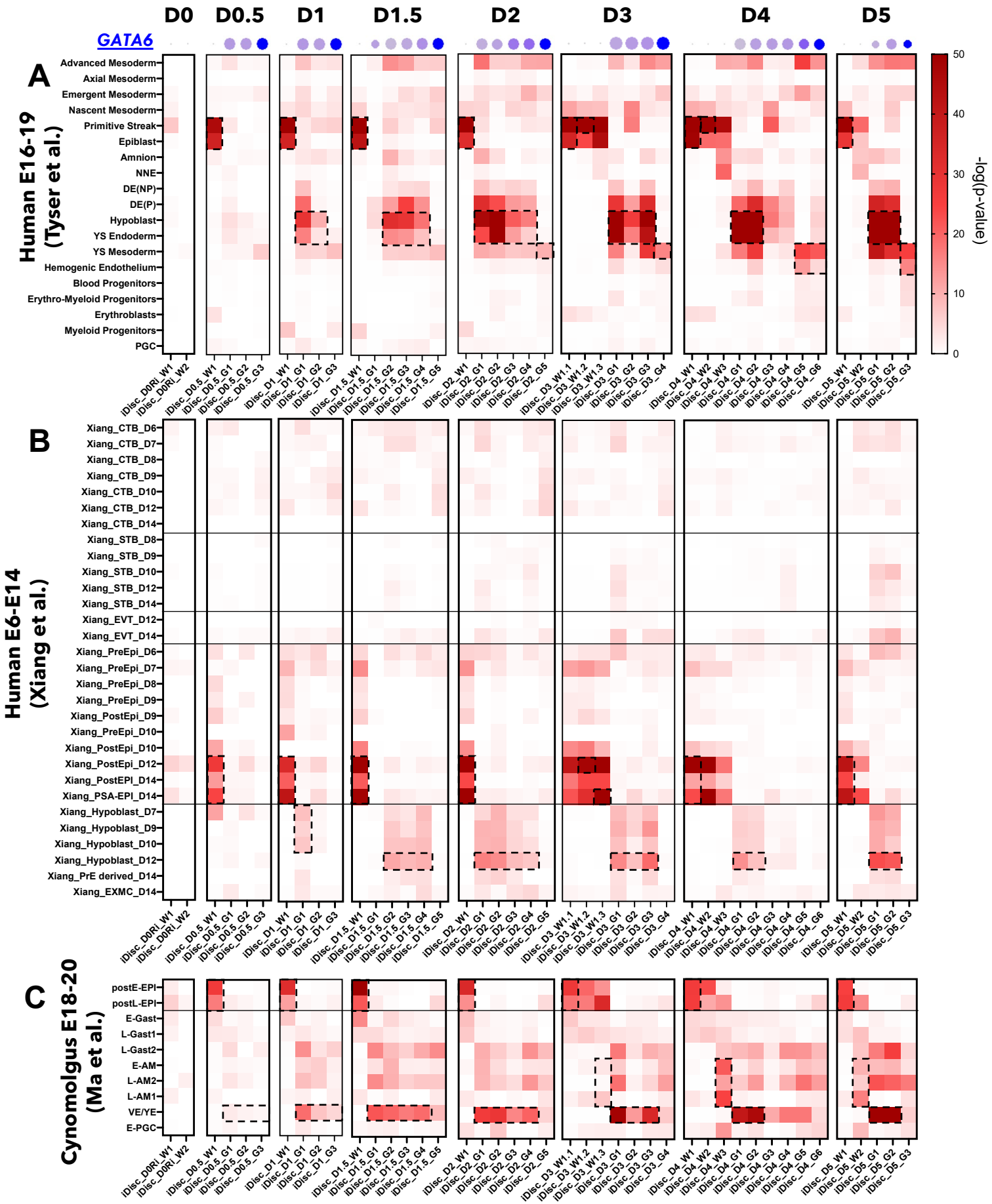
Extended Data Figure 5: Merged clustering of D0-D5 iDiscoid single cell populations

(A) Individual UMAP projections and clustering for each time point recorded through day 5 in iDiscoid development.

(B) Violin plots showing a curated set of genes in iDiscoid day 0, day 2, and day 4 clusters.

iDiscoid clusters are ordered by lowest to highest GATA6 expression level. "W" clusters are clusters with putative wild-type lineage; "G" clusters are clusters with putative iGATA6 lineage.

Extended Data Figure 6



Extended Data Figure 6: Hypergeometric statistical comparison of iDiscoid time points to human and NHP embryo data.

(A) Hypergeometric statistical comparison of each iDiscoid day to the annotated human embryo populations from Tyser et al. 2022. Blue dots above each column indicates the relative GATA6 expression level of each population indicated per day.

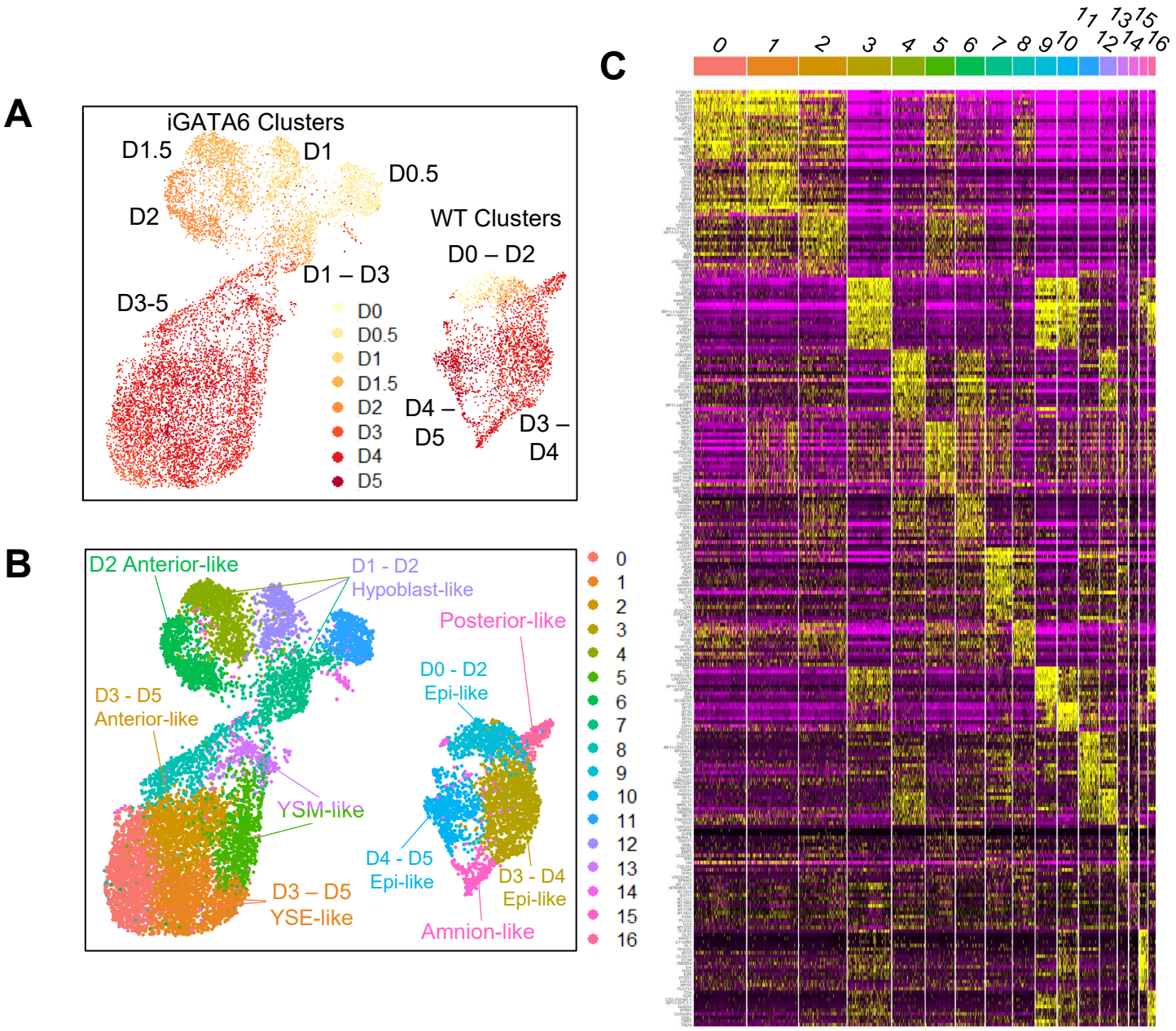
(B) Hypergeometric statistical comparison of each iDiscoid day to the annotated human embryo populations from Xiang et al. 2020. Scale used is the same as shown in panel A.

(C) Hypergeometric statistical comparison of each iDiscoid day to the annotated cynomolgus embryo populations from Ma et al. 2019. Scale used is the same as shown in panel A.

iDiscoid clusters correspond to those in the individual day-by-day clustering in Extended Data Figure 5 and are ordered from left to right on the x-axis by lowest to highest GATA6 expression level. "W" clusters are clusters with putative wild-type lineage; "G" clusters are clusters with putative iGATA6 lineage. Dotted outlines indicate fate comparisons of the most interest for each cluster.

Abbreviations from other datasets: NNE = Non-neural ectoderm, DE (P) = Definitive endoderm (proliferative), DE (NP) = Definitive endoderm (not proliferative), YS = Yolk sac, PGC = Primordial germ cell, CTB = Cytotrophoblast, STB = Syncytiotrophoblast, EVT = Extravillous trophoblasts, PSA-EPI = Primitive streak anlage in the epiblast, EXMC = Extraembryonic mesoderm cells, E- = Early, L- = Late, Gast = Gastrulating cells, AM = Amnion, VE/YE = Visceral endoderm/Yolk sac endoderm

Extended Data Figure 7



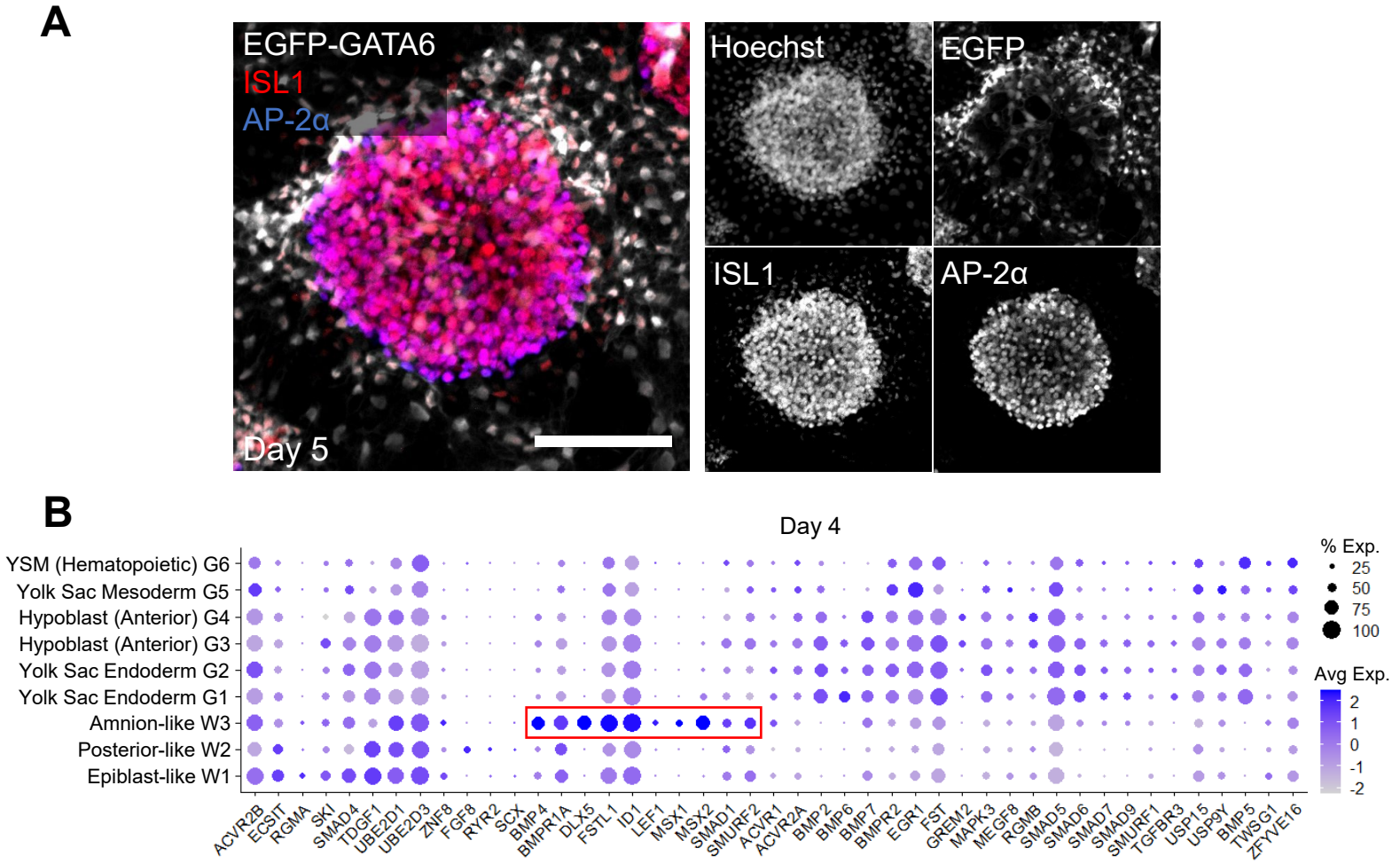
Extended Data Figure 7: Merged clustering of Day 0 – Day 5 iDiscoid single cell populations

(A) UMAP showing the merged iDiscoid clusters annotated by day.

(B) UMAP showing the merged iDiscoid clusters with unsupervised clustering applied. (*repeated figures*: Cluster 14 is the amnion-like cluster shown in Figure 2B; Cluster 6 is the D2 anterior-like cluster shown in Figure 3A; Cluster 2 is the D3-5 anterior-like cluster shown in Figure 3A; and Cluster 15 is the posterior-like cluster shown in Figure 3B.)

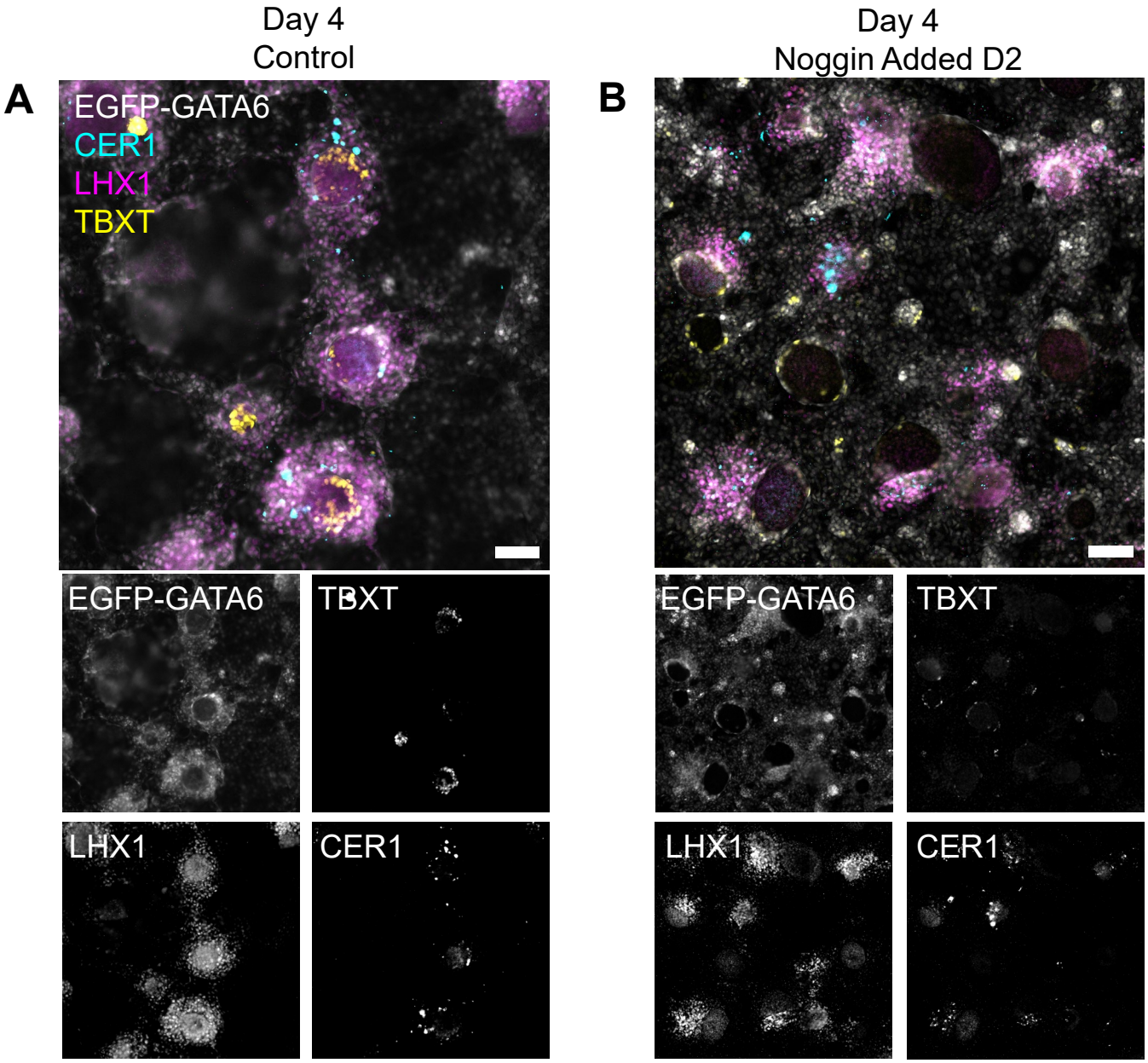
(C) Heatmap showing the top 20 genes corresponding to each cluster of the merged day 0 through day 5 dataset.

Extended Data Figure 8



Extended Data Figure 8: Major signaling interactions between day 5 iDiscoid cell clusters.
(A) Immunofluorescence staining for the amnion markers ISL1 and AP-2α at day 4. Top-down widefield image of a flattened coverslip. Scale Bar = 100 μm
(B) Dot plot of marker genes from the BMP pathway from day 4 iDiscoid scRNA-seq. BMP4 expression and a number of associated genes (boxed in red) are highest in the amnion-like population.

Extended Data Figure 9



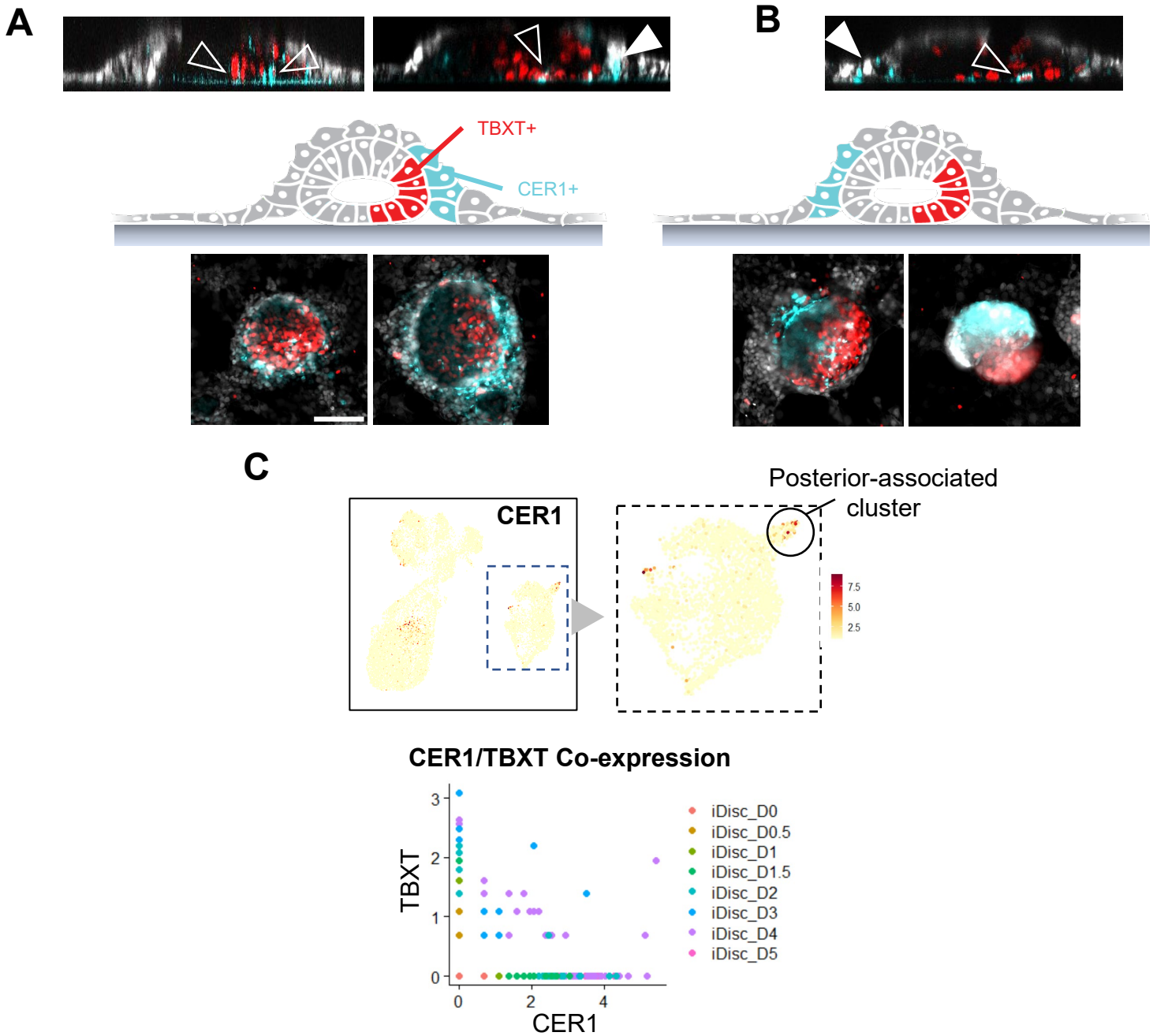
Extended Data Figure 9: Effect of BMP4 inhibition on anterior-like and posterior domains in iDiscoids.

(A) Control iDiscoids showing development of TBXT$^+$ posterior domains and LHX1$^+$ areas expressing CER1.

(B) iDiscoids with 100ng/mL Noggin added at day 2 showing presence of LHX1$^+$ areas expressing CER1, but no expression of TBXT$^+$ domains. TBXT expression is seen in iGATA6-lineage cells at the periphery of the WT disc.

Scale bars = 100 μm

Extended Data Figure 10



Extended Data Figure 10: Examples of TBXT and CER1 observed in iDiscoids.

(A) A representative WT cluster showing syn-polarity of TBXT and CER1 within the WT cluster and iGATA6 layers, respectively. Filled arrow indicates CER1-expressing cells within the iGATA6 layer; empty arrow indicates CER1 and TBXT co-expressing cells within the WT layer. Z-slices are representative slices from the center of two different WT discs.

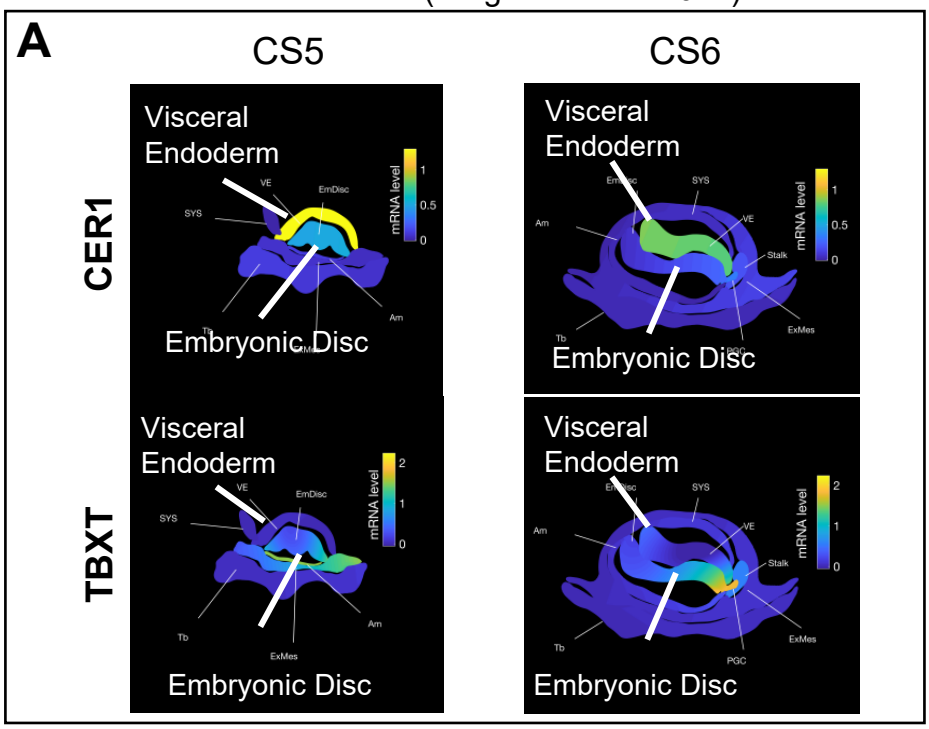
(B) Representative WT clusters showing anti-polarity of TBXT and CER1 within the WT cluster and iGATA6 layers, respectively. Filled arrow indicates CER1-expressing cells within the iGATA6 layer; empty arrow indicates CER1 and TBXT co-expressing cells within the WT layer. Z-slice is a representative slice from the center of a WT disc.

(C) Top shows a merged UMAPs of all D0-D5 scRNA-seq data showing the posterior compartment expressing the inhibitor CER1. Dotted boxes show WT lineages. Posterior region was illustrated in Figure 3B. Bottom shows a scatterplot from merged D0-D5 iDiscoid scRNA-seq showing a population of cells derived from D3-D4 iDiscoid that co-express CER1 and TBXT. Axes represent scaled expression levels of each marker; dots represent individual cells colored by day.

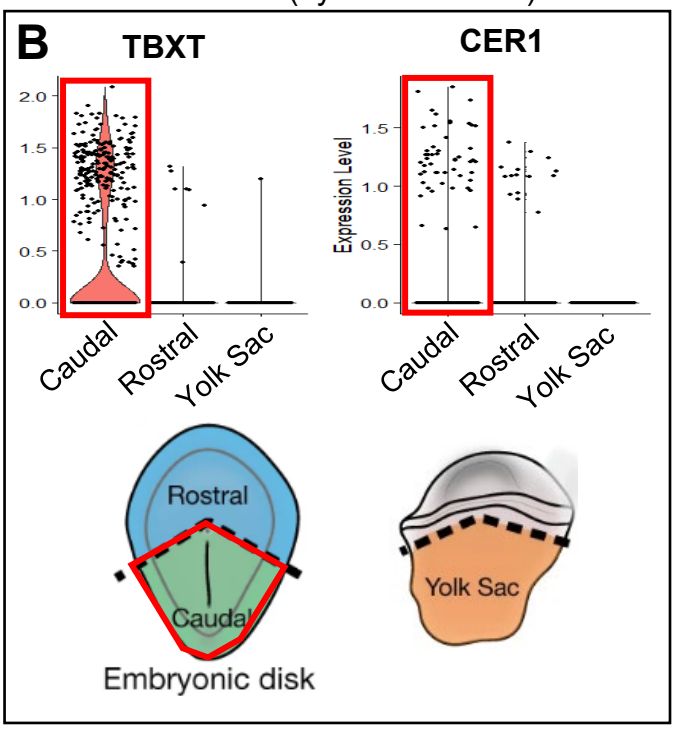
Scale Bars = 100 μm

Extended Data Figure 11

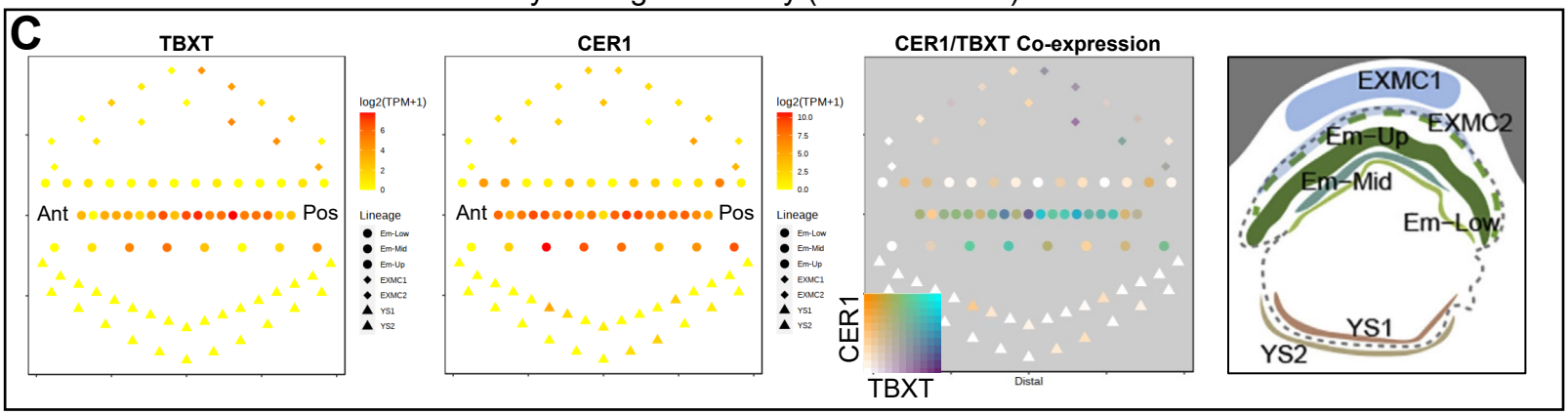
Marmoset (Bergmann et al 2022)



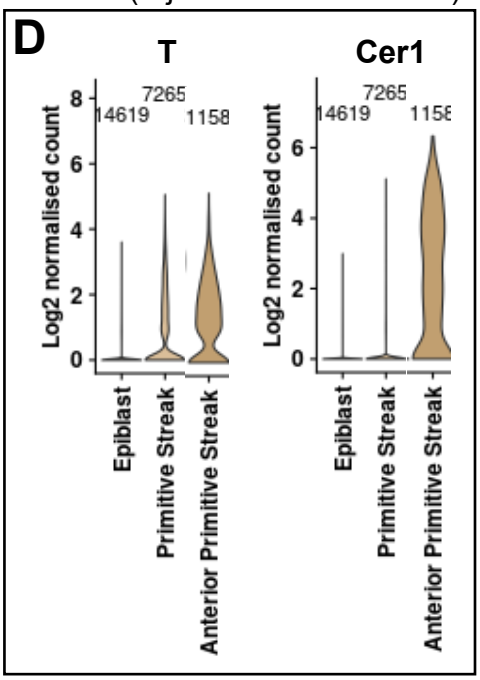
Human (Tyser et al 2021)



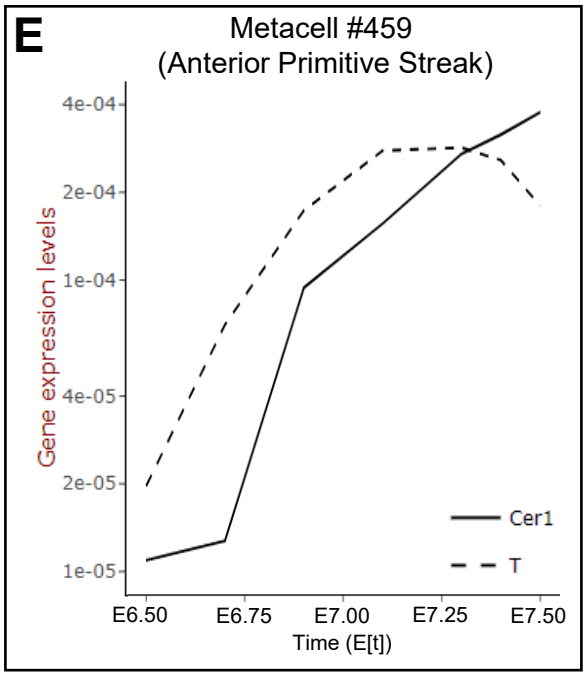
Cynomolgus Monkey (Cui et al 2022)



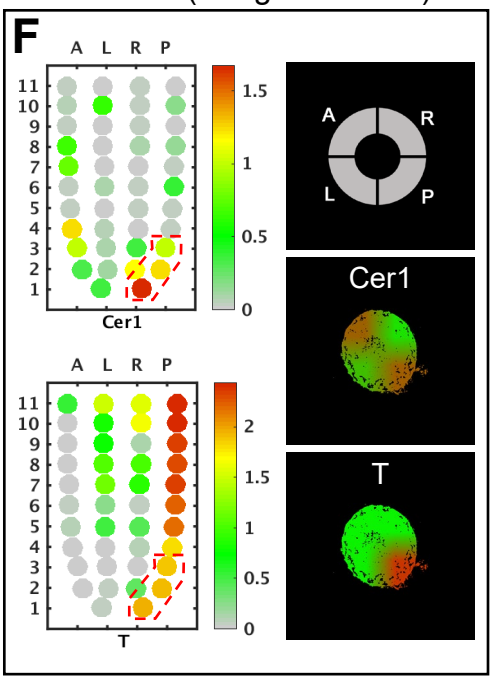
Mouse (Pijuan-Sala et al 2018)



Mouse (Mittnenzweig et al 2021)



Mouse (Peng et al 2016)



Extended Data Figure 11: Positioning of TBXT and CER1 expression in transcriptomic human, NHP and mouse datasets.

(A) Spatial expression of TBXT and CER1 from the CS5 and CS6 marmoset embryo from Bergmann et al. 2022. In marmoset CS5, CER1 has high and unpolarized expression in the extraembryonic endoderm overlaying the embryonic disc; by CS6, overall expression has decreased, and a very slight polarization away from the putative posterior is observed. From CS5 to CS6, TBXT moves from unregionalized/bipolar expression within the embryonic disc to strong expression at one pole of the embryonic disc.

(B) Regional expression of TBXT and CER1 in the E16-19 (CS7) human embryo from Tyser et al. 2022. Highest expression of TBXT and CER1 is observable in the caudal portion of the embryo, which corresponds to the area of the primitive streak and the hypoblast overlaying this area, as shown in the diagram.

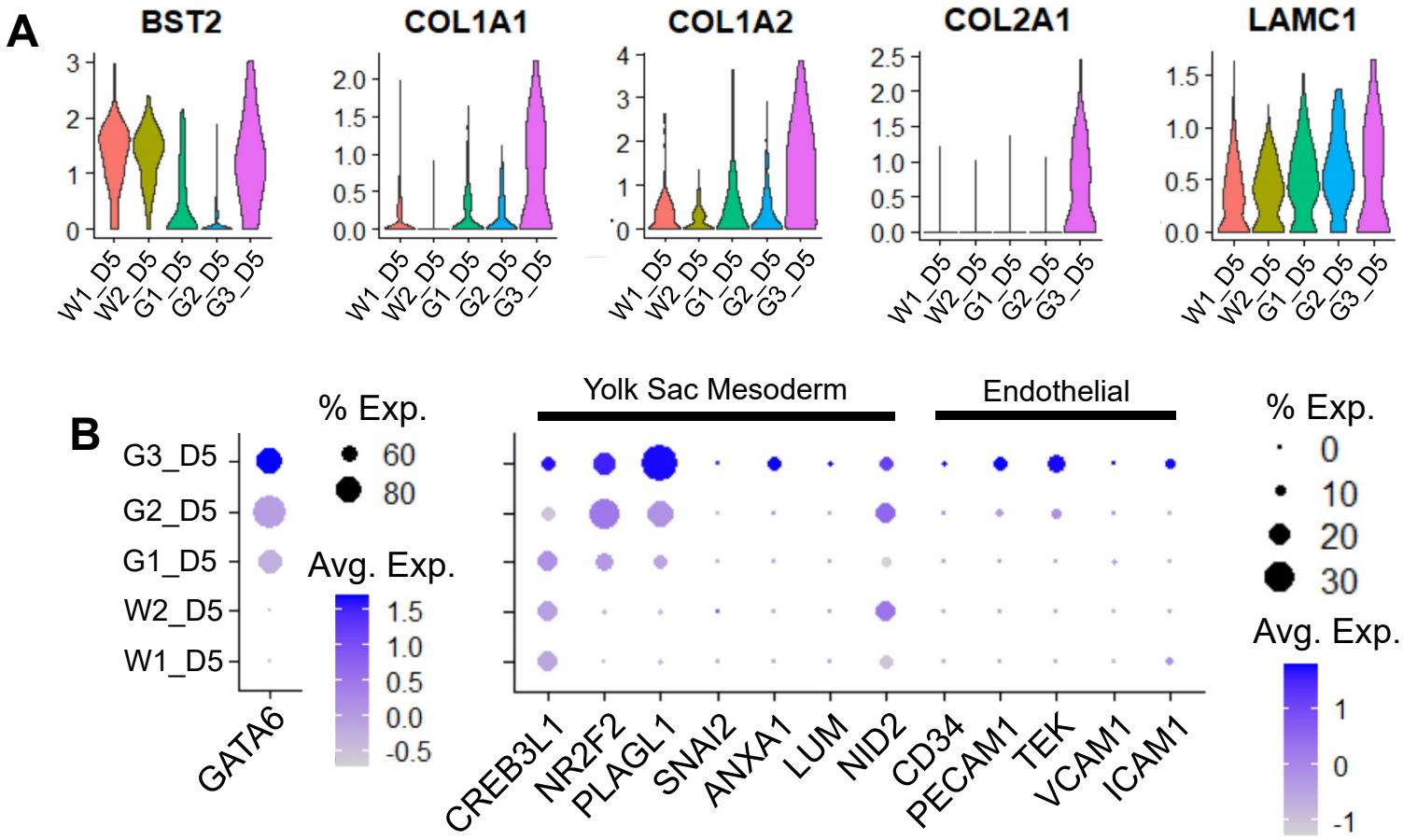
(C) Pseudospatial representation of TBXT and CER1 expression in the cynomolgus E18-E20 embryo from Cui et al. 2022. Diagram to the right shows the microdissected embryo sections with labels corresponding to each identified tissue layer. CER1 expression strongly overlays TBXT expression in the lower embryonic (EM-Low) and middle embryonic (EM-Mid) regional layers.

(D) Expression levels of T and Cer1 from the E6.5-E8.5 mouse embryo. There is a pronounced co-expression within the anterior domain of the primitive streak.

(E) Expression levels of T and Cer1 from an anterior primitive streak metacell from 153 mouse embryos spanning E6.5 – E7.5. Within this representative cell/cell type, T and Cer1 rise together over time at the anterior primitive streak.

(F) Spatial expression levels of T and Cer1 within the ~E6.5-E7.5 mouse embryo. These markers are co-expressed in the area near the anterior tip of the posterior domain. Boxed area in the diagram corresponds highlights this range. Right diagrams show expression patterns of each marker at the index 3 plane in the Y axis of left diagrams.

Extended Data Figure 12

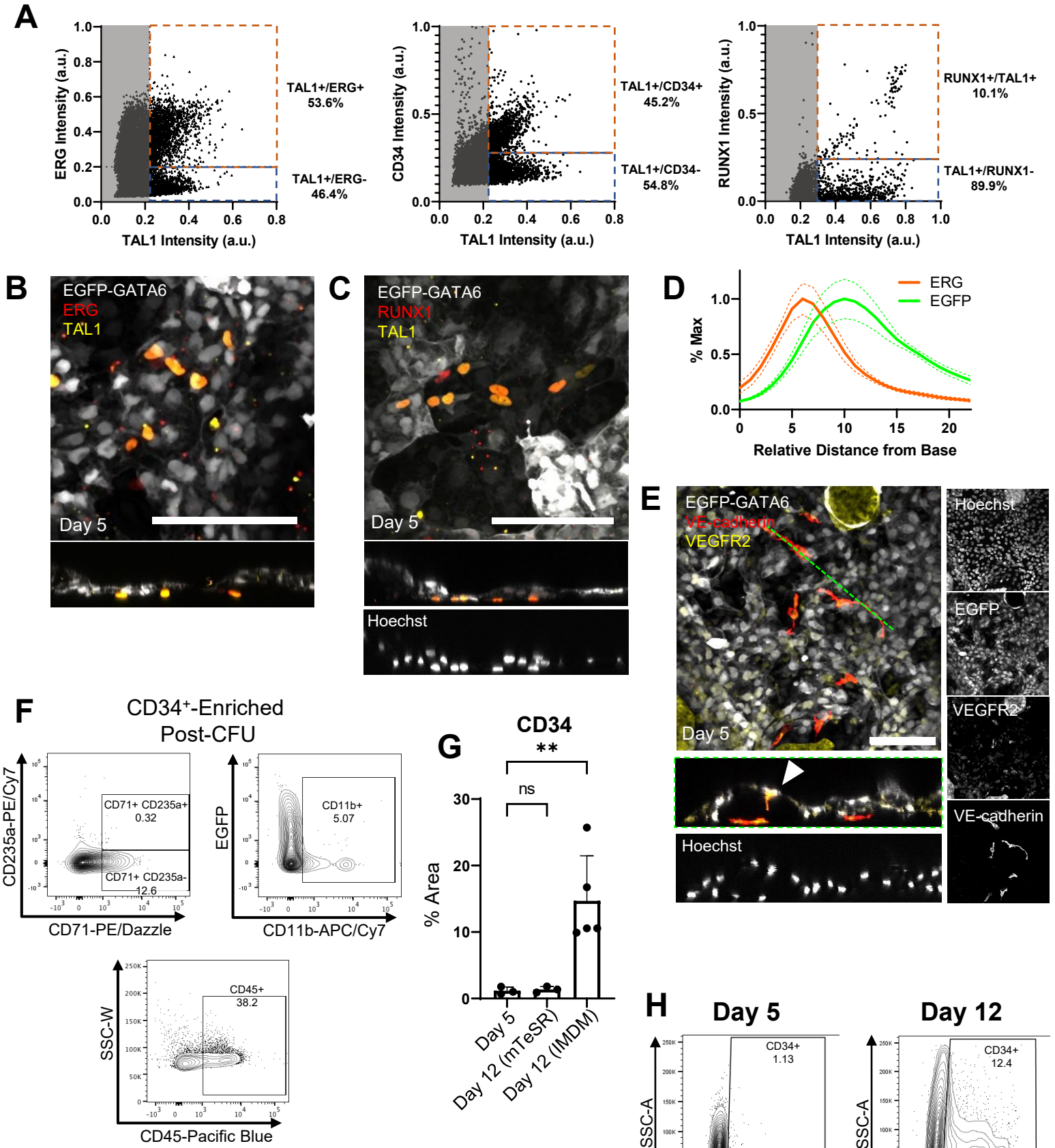


Extended Data Figure 12: Expression of yolk sac mesoderm and extracellular matrix genes in day 4 iDiscoid scRNA-seq.

(A) Violin plots showing the distribution of ECM genes as well as the yolk sac mesoderm marker gene BST2 in day 5 iDiscoid scRNA-seq clusters.

(B) Dot plot showing a subset of day 5 yolk sac mesoderm and endothelial marker genes. The highest GATA6 population expresses the highest levels of YSM marker genes. The GATA6 expression is shown to the left with different scaling of circle sizes than the marker list to the right.

Extended Data Figure 13



Extended Data Figure 13: Identification of endothelial/hematopoietic populations.

(A) Scatterplots showing the distribution of markers obtained via image analysis of 3 independent experiments at day 5. Percentages correspond to the fraction of cells recorded with corresponding marker expression levels above the thresholds represented by the dotted lines.

(B) Immunofluorescence image showing cells expressing hematopoietic markers in iDiscoid culture. Cells expressing the hematopoietic marker TAL1 (Scl) localize between the yolk sac endoderm compartment and the tissue culture dish. Scale bar = 100 μm

(C) Immunofluorescence image showing cells expressing RUNX1 and TAL1 positioned against the dish. Scale bar = 100 μm

(D) Image analysis of z-slices from 5 areas in 3 biological replicates. The peak of ERG expression is underneath the peak of EGFP expression representing the iGATA6 endoderm layer. Dotted curves represent SEM calculated at each point.

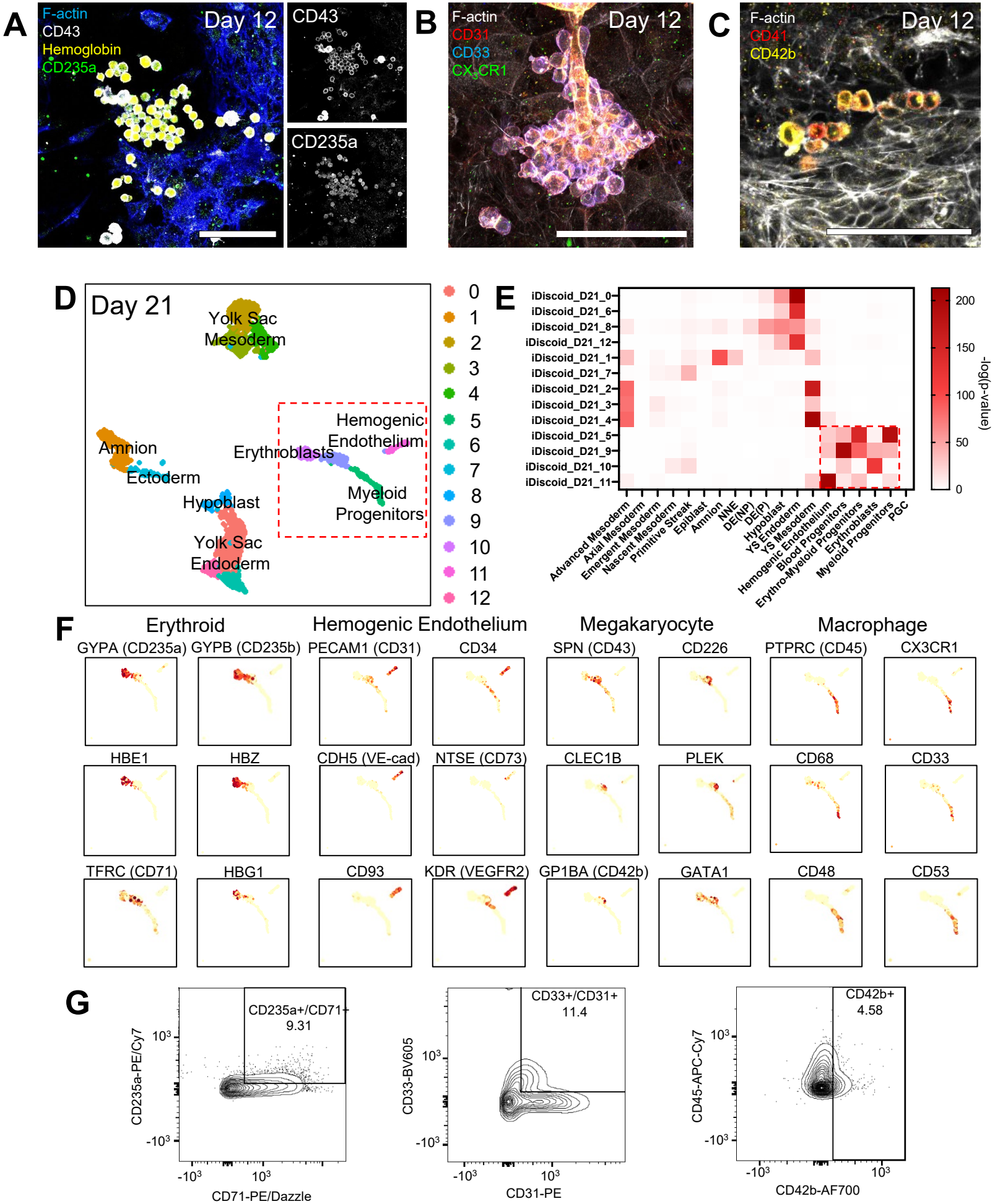
(E) Immunofluorescence image showing cells expressing VE-cadherin and VEGFR2 positioned against the dish. Arrow indicates likely differentiating endothelial cell from the overlaying iGATA6 layer. Scale bar = 100 μm

(F) Flow cytometry plots showing the presence of hematopoietic markers on day 14 after a CFU assay initiated CD34-expressing cells (enriched from iDiscoids).

(G) Results from image analysis demonstrating the change in area of $CD34^+$ cells between day 5 and day 12 of iDiscoid culture, assessed via analysis of immunofluorescence images. Individual dots represent biological replicates. **: P value=0.0055

(H) Representative flow cytometry plot on day 5 and day 12 of iDiscoid showing expansion of the $CD34^+$ population by day 12 (n=3).

Extended Data Figure 14



Extended Data Figure 14: Hematopoietic cells in iDiscoid.

(A) Erythroid markers (CD235a and Hemoglobin) expressed in a subset of $CD43^+$ round cells in day 12 iDiscoids. Scale bar = 100μm

(B) Cells expressing myeloid (CD31, CD33) and macrophage (CX_3CR1) markers in a representative cluster in day 12 iDiscoid. Scale bar = 100μm

(C) Cells expressing megakaryocyte markers in day 12 iDiscoid. Scale bar = 100μm

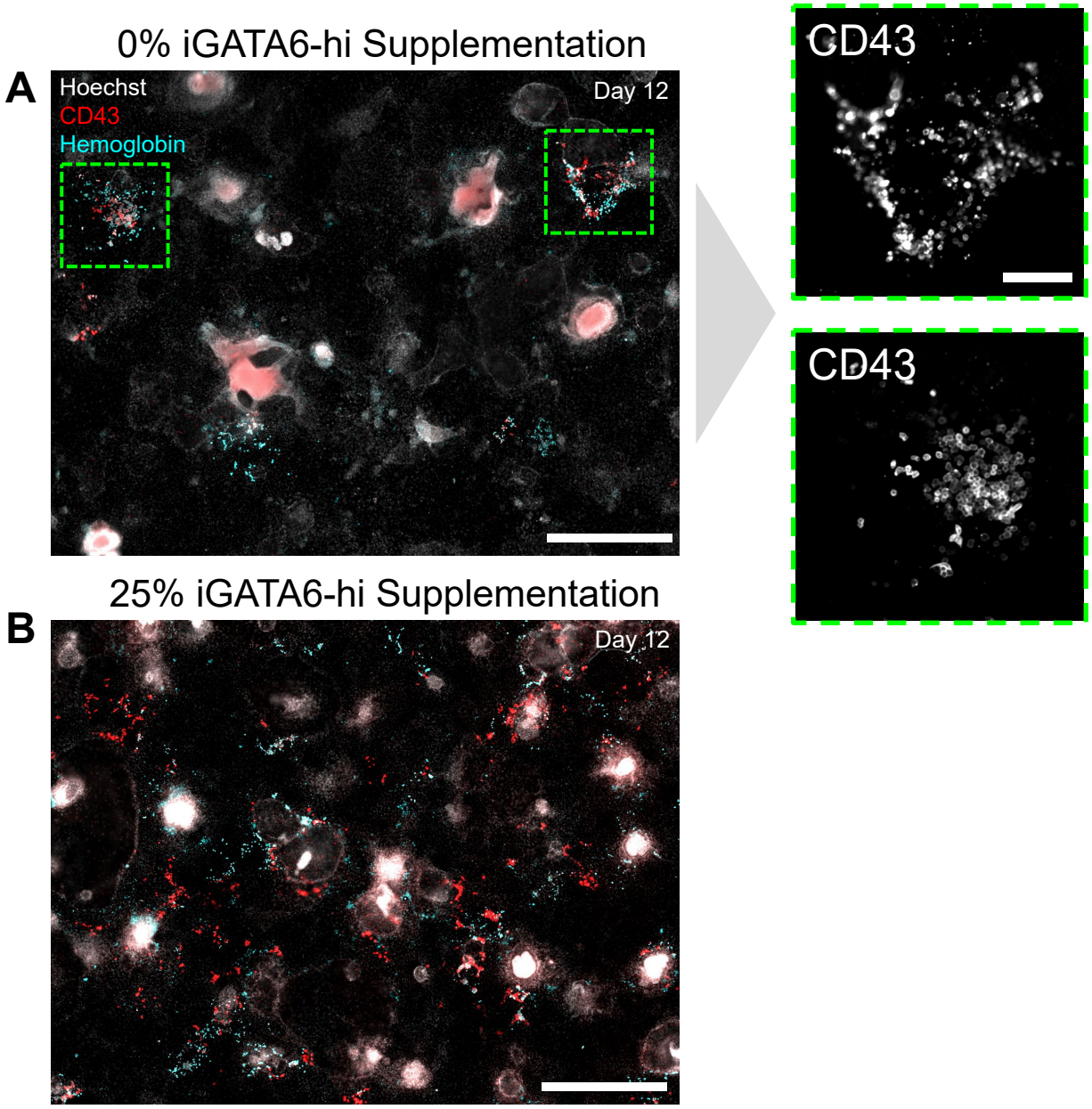
(D) UMAP showing unbiased clustering of scRNA-seq performed on day 21. Boxed area shows populations expressing hematopoietic marker genes.

(E) Hypergeometric statistical comparison of day 21 iDiscoid populations to the E16-19 human embryo populations from Tyser et al. 2022. The red box shows populations of interest with similarity to in vivo hematopoietic lineages. iDiscoid clusters are grouped by similar fates on Y axis.

(F) Expression of hematopoietic marker genes within the red boxed population in panel D. The four clusters show distinct marker gene expression profiles associated with yolk sac-derived hematopoietic populations.

(G) Flow cytometry analysis on day 21 iDiscoid. Cells were gated for positive CD43 expression prior to gating for the markers shown (n=2, representative sample shown).

Extended Data Figure 15

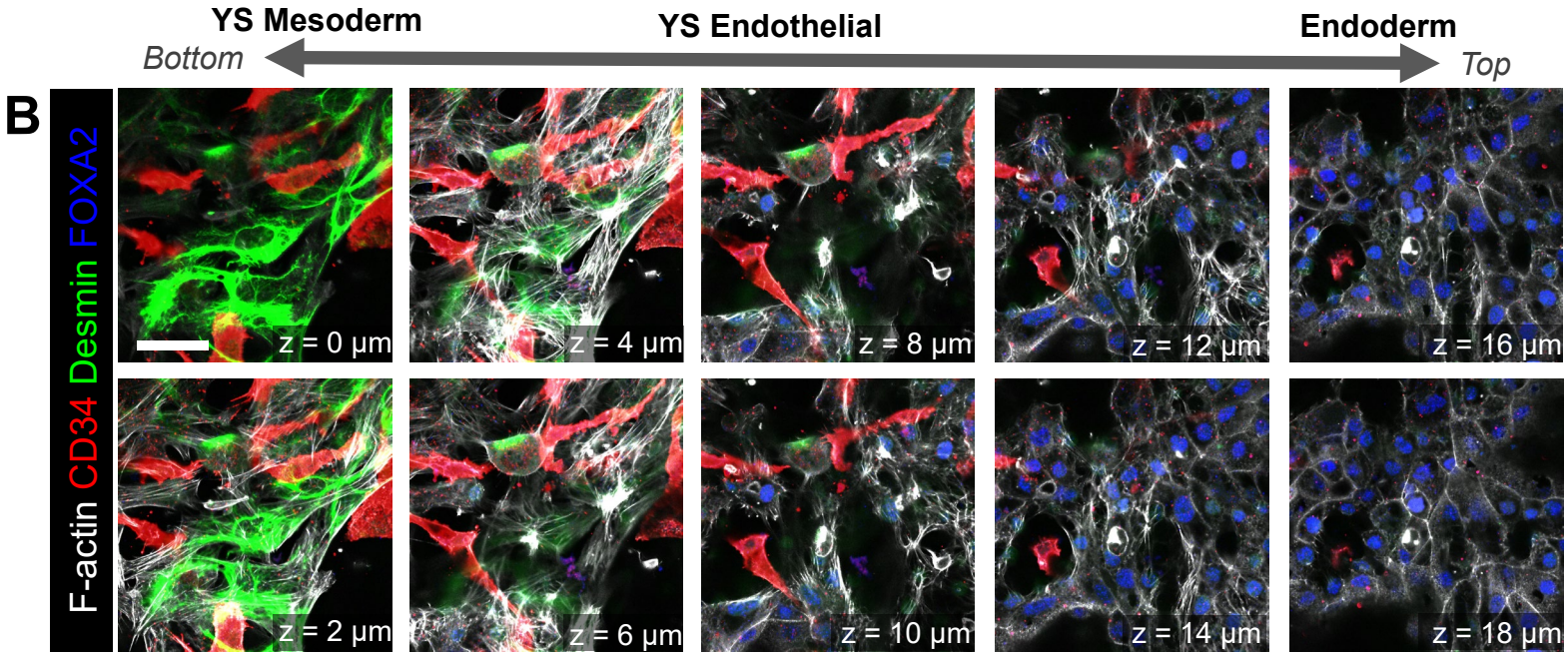


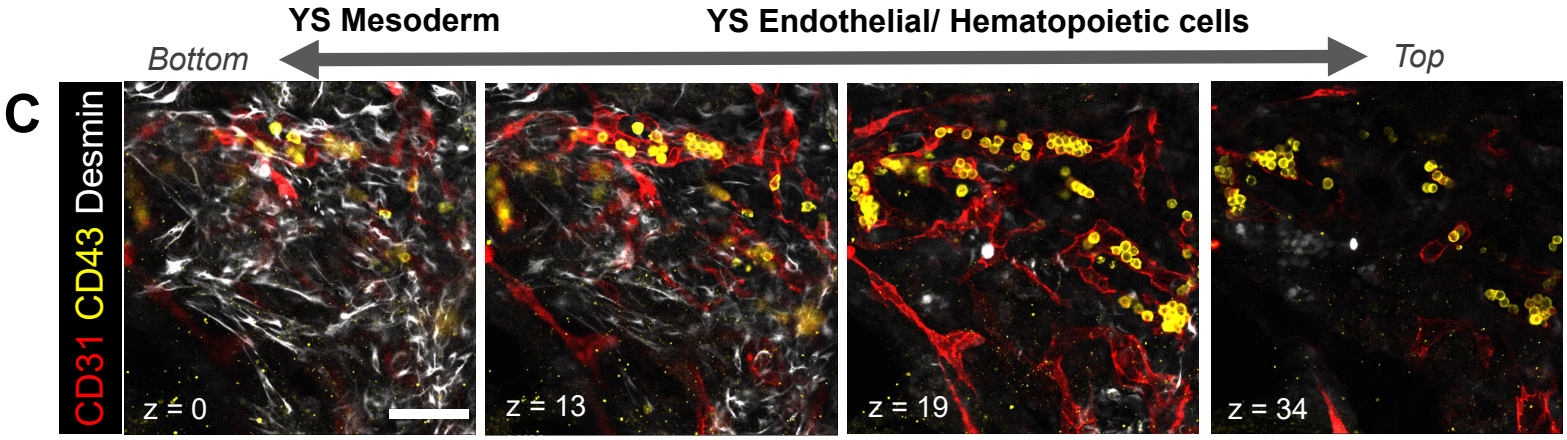
Extended Data Figure 15: Comparison of iDiscoids with and without supplementation with high iGATA6 expressing (iGATA6-hi) cells.

(A) Expression of CD43⁺ and Hemoglobin⁺ cells in day 12 iDiscoid without additional supplementation of GATA6-hi cells. Red color in high Hoechst areas is nonspecific staining. Scale bar = 500 μm. Insets show areas of CD43⁺ cells. Dotted outlines on right show zoomed in areas shown within the image. Scale bar = 100 μm

(B) Expression of CD43⁺ and Hemoglobin⁺ cells in day 12 iDiscoid with supplementation of 25% GATA6-hi cells at initial seeding. A significant expansion of CD43-expressing cells is observable. Scale bar = 500 μm. A representative image of 2-4 biological replicates.

Extended Data Figure 16



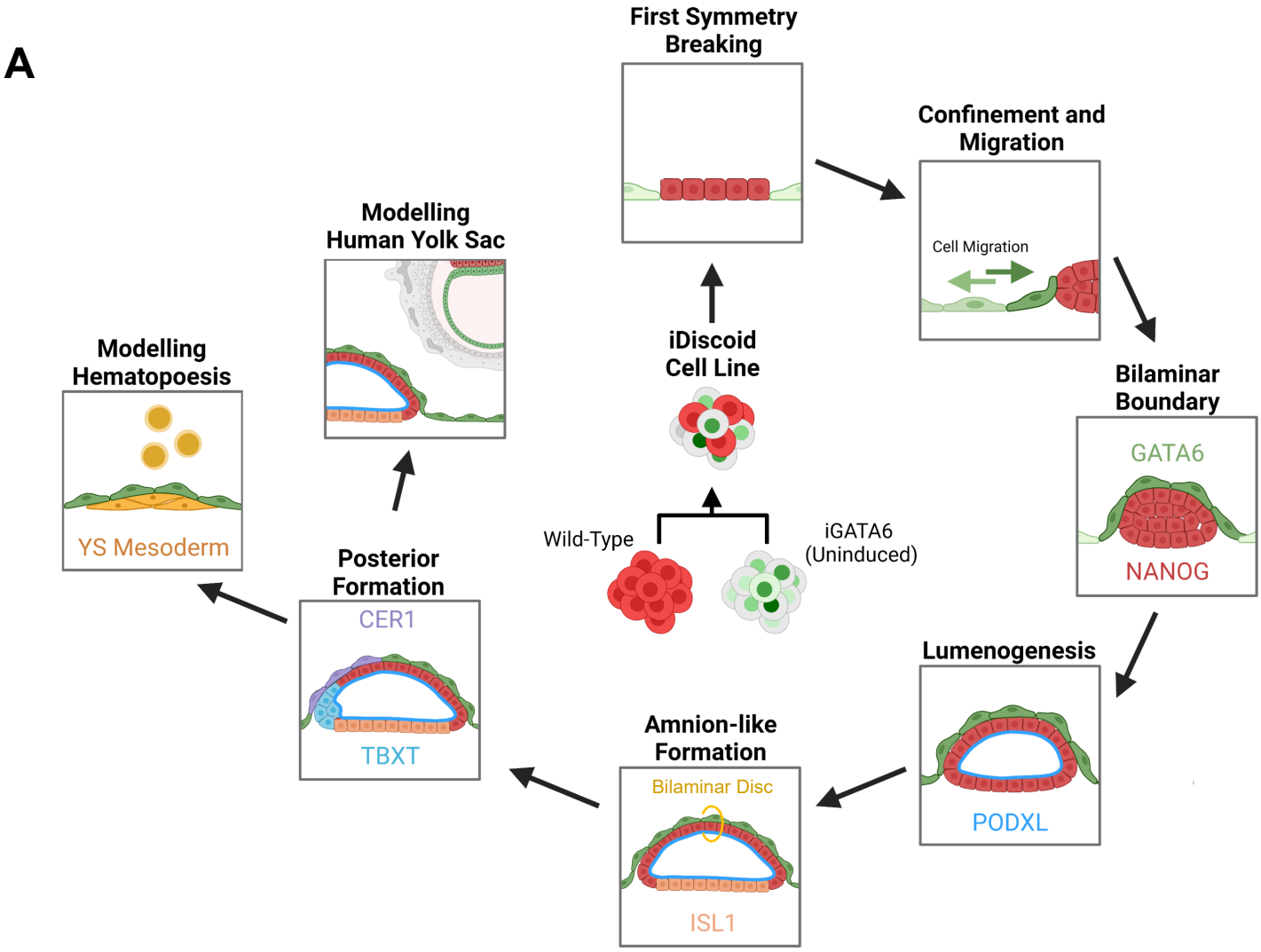


Extended Data Figure 16: Structure of blood island-like associated tissues within iDiscoid.

(A) Z-slices of a day 12 iDiscoid at the indicated height above the bottom focal plane. Expression of the mesodermal marker Desmin is regionalized close to the bottom of the tissue; endothelial marker CD34 is expressed near the bottom and middle of the tissue, with highest expression just above Desmin (z = 6-8); endoderm marker FOXA2 is exclusively expressed above the other tissues. Scale bar = 50 µm

(B) Z-slices around the blood island-like area shown in Figure 4E. Desmin is observed regionalized below the endothelial ($CD31^+$) and hematopoietic ($CD43^+$) tissues. No stain for endoderm is shown. Scale bar = 100 µm

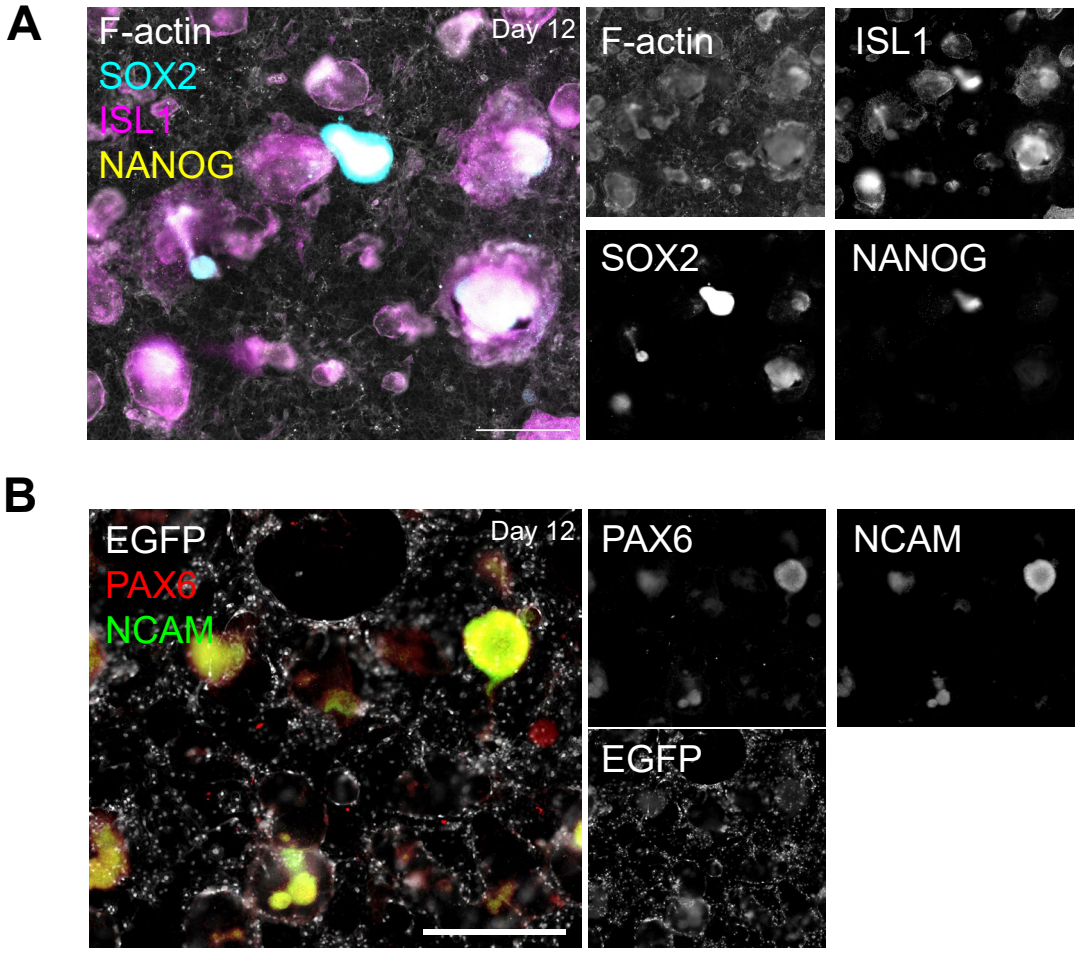
Extended Data Figure 17



Extended Data Figure 17: iDiscoid formation from hiPSCs to model human early post-implantation development *in vitro*.

(A) From an initially mixed state, iDiscoid cells segregate into WT clusters surrounded by iGATA6 cells. These iGATA6 cells migrate laterally to create a bilaminar boundary on top of the WT clusters. These clusters then undergo lumenogenesis, specification of amnion-like cells, and formation of anterior hypoblast-like and posterior domains. In iGATA6-only areas, yolk sac mesoderm specifies and hematopoietic cell specification is observed.

Extended Data Figure 18



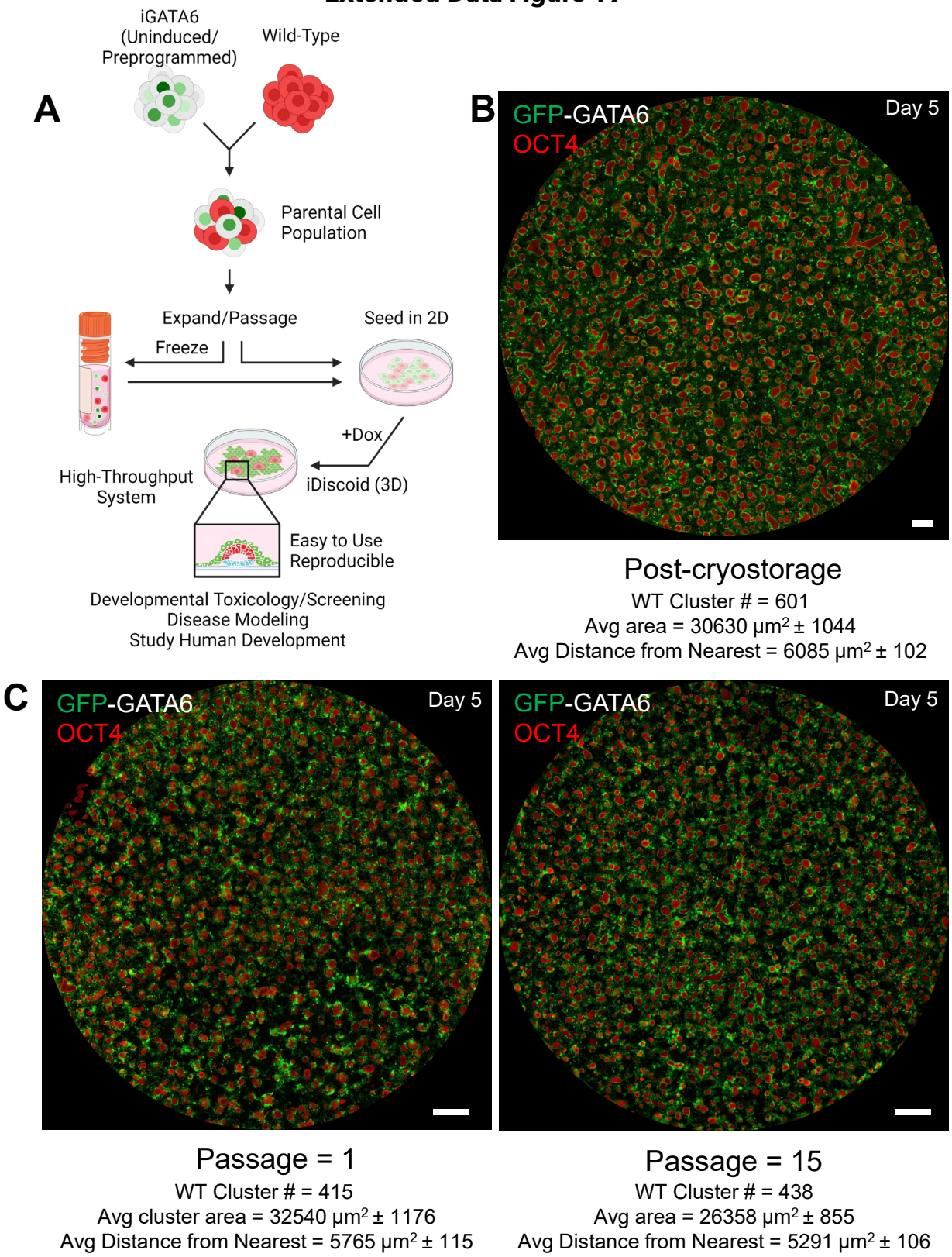
Extended Data Figure 18: Investigation of WT fates in day 12 iDiscoid.

(A) Most structures corresponding to former WT clusters at day 12 in iDiscoid have taken on expression of ISL1 and lost expression of the pluripotency markers SOX2 and NANOG. A limited number of cells express SOX2 at the core of former WT clusters, potentially indicating the specification of an ectoderm-like fate in a small number of WT-lineage cells.

(B) A subset of former WT clusters have taken on markers of ectoderm differentiation.

Scale bars = 500 μm

Extended Data Figure 19



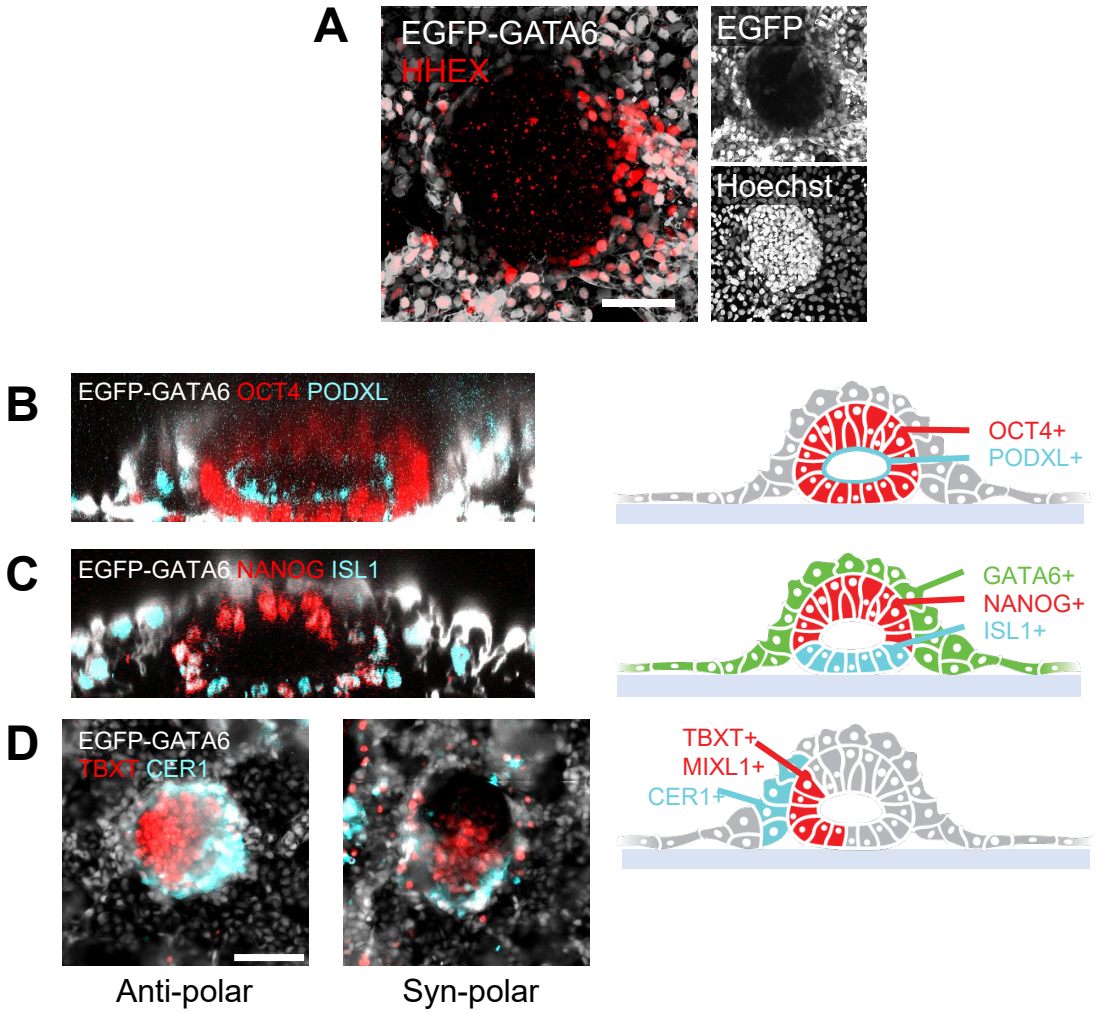
Extended Data Figure 19: Mixed iDiscoid cell lines exhibit consistent cluster formation after repeated passaging and cryostorage prior to induction.

(A) Schematic showing the creation of the iDiscoid parental cell line. iGATA6 cells with heterogeneous copy numbers of the inducible GATA6 circuit are mixed with wild-type. This cell combination is then maintained together or frozen prior to induction for iDiscoid experiments.

(B) iDiscoid morphology and characteristics following cryostorage and defrosting. Scale Bar = 500 μm

(C) Characteristics and morphology of iDiscoid cultures induced immediately after mixing (passage 1) or maintained together and passaged for two months (passage 15). Scale Bars = 500 μm

Extended Data Figure 20



Extended Data Figure 20: iDiscoid engineering in a separate iPSC cell line results in the generation all major features of interest

Day 4 immunofluorescence images of iDiscoid engineered using the iPSC cell line PGP9.

(A) A portion of PGP9 iGATA6 cells expressing high levels of GATA6 (EGFP) also express the anterior endoderm marker HHEX near the edge of a WT cluster.

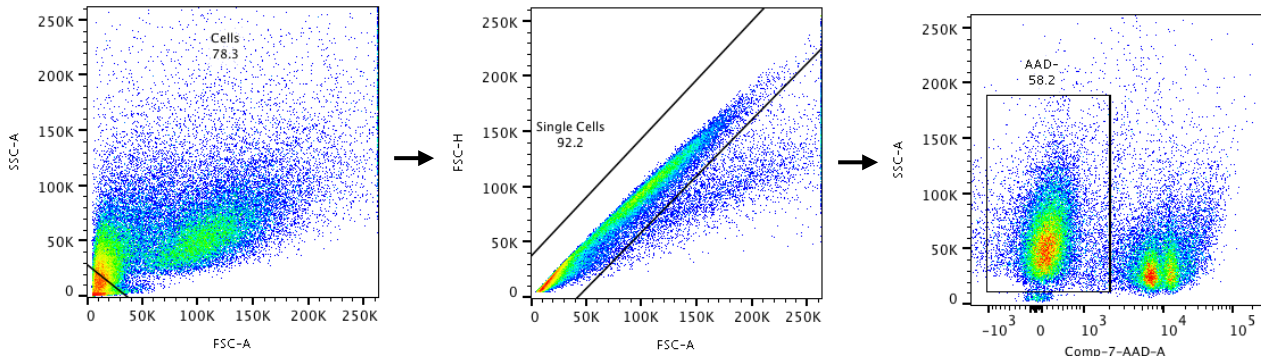
(B) A ring of PODXL expression lines the inside of a cavities formed in PGP9 WT clusters.

(C) ISL1$^+$ cells specify away from NANOG$^+$ cells, along the base of a cavity formed in PGP9 WT.

(D) TBXT$^+$ cells develop polar domains at the edge of WT clusters. Representative islands show that both syn- and anti-polar arrangement of cells are observable.

Scale Bars = 100 μm

Extended Data Figure 21



Extended Data Figure 21: Example of FACS gating strategy.
(Cited in Methods) Cell debris was excluded via an SSC-A vs FSC-A gate; aggregates were excluded by comparing FSC-A and FSC-H; dead cells were gated out using the 7-AAD stain to identify positive cells.

Supplementary Information

Modelling Human Post-Implantation Development via Extra-Embryonic Niche Engineering

Joshua Hislop[1-3], Amir Alavi[4,5], Qi Song[4,5], Rayna Schoenberger[1,3], Kamyar Keshavarz F.[1-3], Ryan LeGraw[2,3], Jeremy Velazquez[2,3], Tahere Mokhtari[1-3], Mohammad Nasser Taheri[1-3], Matthew Rytel[2,3], Susana M Chuva de Sousa Lopes[8], Simon Watkins[6,7], Donna Stolz[6,7], Samira Kiani[1-3,9], Berna Sozen[10], Ziv Bar-Joseph[4,5], Mo R. Ebrahimkhani[1-3,9]*

Affiliations:

1. Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA 15261, USA
2. Department of Pathology, Division of Experimental Pathology, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15213, USA
3. Pittsburgh Liver Research Center, University of Pittsburgh, Pittsburgh, PA 15261, USA
4. Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
5. Machine Learning Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA.
6. Center for Biologic Imaging, University of Pittsburgh, Pittsburgh, PA, USA.
7. Department of Cell Biology and Molecular Physiology, University of Pittsburgh, Pittsburgh, PA, USA
8. Department of Anatomy and Embryology, Leiden University Medical Center, Einthovenweg, 2333 ZC Leiden, the Netherlands
9. McGowan Institute for Regenerative Medicine, University of Pittsburgh, Pittsburgh, PA 15219, USA
10. Department of Genetics, Yale School of Medicine, Yale University, New Haven, CT, 06510, USA.

*Correspondence: mo.ebr@pitt.edu

Supplementary Video 1: Self-organization of hiPSC cells that result in WT discs surrounded by iGATA6 cells from day 0 to day 3

Supplementary Video 2: Migration of iGATA6 cells over the top of the WT cells to form a contiguous membrane
iGATA6 cells (green) cover the top of the WT disc (red) between day 0.5 and day 2.

Supplementary Video 3: Expansion of a lumen from a rosette in a single iDiscoid.

Supplementary Video 4: Fly-through of an iDiscoid with a single lumen
Z-slices of a single fixed iDiscoid moved through from end-to-end.
DAPI: blue, iGATA6: Green, PODXL: Red, F-actin: Gray