

Multiplexed single cell transcriptomics optimizes mesodermal patterning and hemogenic endothelial output from murine embryonic stem cells

Barbara Varnum-Finney*¹, Adam M. Heck*¹, Sanjay R. Srivatsan², Stacey Dozono¹, Rachel Wellington¹, Cynthia Nourigat-McKay¹, Tessa Dignum¹, Cole Trapnell³, Brandon Hadland^{1, 4}

1. Translational Science and Therapeutics Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

2. Basic Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, USA

3. Department of Genome Sciences, University of Washington, Seattle, WA, USA

4. Division of Pediatric Hematology/Oncology, University of Washington School of Medicine, Seattle, WA, USA

*Co-first authors

Correspondence: Brandon Hadland, M.D, Ph.D.

Fred Hutchinson Cancer Center, 1100 Fairview Ave N., Seattle, WA 98109, USA

(bhadland@fredhutch.org), Tel: 206.667.1095, Fax: 206.667.6084

Summary

Early patterning of mesodermal precursor populations is a key step of hematopoietic development in the embryo. To better understand this process, we employed sci-Plex, a high-throughput method of measuring multiplexed perturbations at the single-cell level, to evaluate the transcriptional response of mouse embryonic stem cells subjected to a gradient of two key morphogens in early mesoderm/hematopoietic development, Activin and BMP4. sci-Plex revealed varying combinations of Activin and BMP4 temporally influenced mesoderm patterning in vitro and subsequent production of cell types reflecting their in vivo counterparts. We leveraged sci-Plex data to further optimize the generation of intraembryonic-like hemogenic endothelial cells that serve as the precursors of hematopoietic lineages, including hematopoietic stem cells. This study highlights the utility of sci-Plex to dissect how dose and temporal integration of interacting signal pathways determines cell fates and serves as a resource to analyze cell fate choices in early mesoderm patterning at single cell resolution.

Introduction

Hematopoietic stem cells (HSCs) are clinically valuable due to their distinctive capacity for life-long reconstitution of all hematopoietic lineages in the marrow of conditioned transplant recipients. Yet, it has proven particularly difficult to derive HSCs from either murine embryonic stem cells (mESC) or human induced pluripotent stem cells (iPSC) despite decades of efforts (Freire and Butler, 2020). This is primarily due to the complexity of embryonic hematopoiesis such that attempts at recapitulating HSC genesis in ESC culture has resulted in a multiplicity of false leads (Ditadi et al., 2017). Embryonic hematopoiesis occurs in temporal waves of different hematopoietic progenitor types prior to the emergence of the first HSCs (Hadland and Yoshimoto, 2018). The initial wave, referred to as the primitive wave, occurs almost immediately after gastrulation in extraembryonic tissues, generating the first erythroid cells and megakaryocytes (Palis et al., 2010; Tober et al., 2007). Subsequent waves of hematopoiesis are derived from specialized endothelial cells known as hemogenic endothelium (HE), giving rise to blood cell fates through a process referred to as the endothelial to hematopoietic transition (EHT). EHT occurs in various embryonic sites including the yolk sac, placenta, endocardium, and the aorta-gonad-mesonephros region (AGM), first generating HSC-independent oligopotential and multipotent progenitor types with various erythroid, myeloid and lymphoid potential, followed by the emergence of the first long-term engrafting HSCs (Dignum et al., 2021; Frame et al., 2016; Müller et al., 1994; Yoshimoto et al., 2011; Zhu et al., 2020). Many mESC and iPSC hematopoietic differentiation protocols have recapitulated the HSC-independent progenitor waves in vitro. However, protocols developed to date have failed to robustly

and reproducibly generate HSCs. This argues the need to better understand the sequential embryonic cell types generated in current differentiation protocols in vitro and how they relate to the populations in the embryo that serve as precursors to the successive waves of hematopoiesis in vivo, including HSCs.

Mouse gastrulation starts with primitive streak (PS) formation at the extra-/intra-embryonic boundary around embryonic day 6-6.5. The position of PS formation defines the posterior pole of the embryo, and expression of the T-box transcription factor Brachyury is an early indicator of PS formation. As the PS elongates from this posterior position anteriorly, cells continuously exit and migrate from the elongating PS (Rivera-Pérez and Hadjantonakis, 2014). Those migrating anteriorly have an embryonic fate (intra-embryonic cells (IEC)) while those migrating proximally have an extra-embryonic fate (extra-embryonic cells (EEC))(Saykali et al., 2019). Following departure from the PS, IEC migrate toward the embryo proper, forming the bilateral wings of embryonic mesoderm that eventually gives rise to paraxial mesoderm, pharyngeal mesoderm, and other embryonic tissues as well as lateral plate mesoderm (LPM), that produces cardiac cells and contributes to multilineage hematopoiesis including the first HSCs (Prummel et al., 2020). EECs migrate toward the extra-embryonic area and give rise to extra-embryonic tissues including the hematopoietic cells of the first primitive waves as well as the extra-embryonic mesoderm that gives rise to a subsequent yolk sac wave of HE-derived erythromyeloid progenitors (EMP). Morphogens, such as BMP4, Wnt and Activin/Nodal, established in concentration gradients within the embryo, direct these migrations and further promote differentiation to specific cell types that differentially contribute to intraembryonic and extraembryonic tissues. It is well established that

mesoderm differentiation in mESC and huiPSC in vitro culture also requires these morphogens (BMP4, Wnt, and Activin/Nodal) in a time and concentration-dependent manner (Jackson et al., 2010; Kattman et al., 2011; Nostro et al., 2008; Palpant et al., 2015). Protocols for in vitro differentiation of mESC and huiPSC have not been able to fully assess the outcome of morphogen-dependent gradients on mesodermal/hematopoietic differentiation due to a lack of available high-throughput tools to characterize cell fate responses to multiplexed combinations of morphogen concentrations at the single cell level. To overcome this barrier, we applied sci-Plex, a novel method to measure transcriptional responses to highly multiplexed perturbations at single-cell resolution (Srivatsan et al., 2020), to examine the combinatorial dose and time-dependent effects of exogenous Activin and BMP4 gradients on mesodermal patterning and hematopoietic differentiation from mESC. This approach identified mESC-derived cell fates responsive to combinatorial Activin and BMP4 gradients that recapitulate those emerging in both the intra-embryonic and extra-embryonic regions during early murine embryogenesis. Furthermore, the technique facilitated optimization of a differentiation protocol to enhance LPM and HE output as an essential initial step towards HSC production from pluripotent stem cells in vitro.

Results

sci-Plex reveals early embryonic cell fates resulting from exposure of differentiating mESC to combinatorial doses of Activin and BMP4

In order to control for exogenous factors, we built upon a serum-free protocol for mESC differentiation that was previously shown to generate mesoderm capable of producing multilineage hematopoietic cells (Clarke et al., 2013; Irion et al., 2010). To mimic Activin/Nodal and BMP4 gradients that pattern the embryo during PS formation, we applied 16 different combinations of Activin and BMP4 concentrations (Kattman *et al.*, 2011; Palpant *et al.*, 2015) and used sci-Plex to trace mesodermal differentiation and subsequent cell fates at single cell resolution at various time points in the differentiation protocol (Figure 1A). Briefly, mESCs were cultured as embryoid bodies (EBs) for two days without exogenous growth factors, since previous studies have shown that mESC are unresponsive to Activin or BMP4 for 24 to 48 hours following removal of LIF or growth media (Jackson *et al.*, 2010). At 48 hours, EBs were dissociated to single cells and re-suspended to form EBs in a 4x4 gradient of Activin and BMP4, along with a consistent concentration of VEGF required for hematovascular fates (Giles et al., 2005). On days 4, 5 and 6 of culture (days 2, 3 and 4 following addition of growth factors), EBs were dissociated, and a portion of each sample was processed for both sci-Plex and analyzed for immunophenotypic markers by flow cytometry. As expected, the immunophenotypic data indicates a range of mesodermal and endothelial outputs responsive to different Activin and BMP4 concentrations over the period of differentiation analyzed, including $\text{Pdgfra}^+\text{Flk1}^+$ LPM-like cells previously shown to generate both cardiac and hematovascular lineages (Kataoka et al., 2011; Nostro *et al.*,

2008) (Figure S1A) and Flk⁺VE-Cadherin⁺ cells indicating endothelial differentiation (Figure S1B).

sci-Plex data was analyzed using Monocle3 (Cao et al., 2019) by first performing dimensionality reduction by uniform manifold approximation and projection (UMAP) and unsupervised cluster analysis of the single-cell transcriptomes (Figure 1B). Gene expression levels for previously published diagnostic marker genes from early mouse embryonic development were used to assign cell types to each cluster (Figure 1C; Table 1; Figure S2A) (Pijuan-Sala et al., 2019). As expected, analysis across time points revealed a clear progression from PS to mesoderm to differentiated mesoderm-derived cell types (Figure 1D). Moreover, different transcriptionally defined cell types were observed across the spectrum of Activin and BMP4 concentrations (Figure S2B) and depiction of the relative cell number by day and concentration demonstrated concentration-dependent effects on the generation of each cell type (Figure 1D). Mesoderm was maximally generated on days 4 and 5 at the highest Activin/BMP4 concentrations. Endothelial cell types (HE/EC) were detected on days 5 and 6 at all Activin concentrations but were maximally generated at the highest Activin and BMP4 concentrations. Differentiating blood progenitors and erythroid cells were maximally generated on day 6, particularly at higher BMP4 concentrations, whereas the highest Activin concentration inhibited generation of erythroid cell types. These results are consistent with previous studies that have shown hematopoiesis occurs in waves within the differentiating EB, similar to waves of hematopoiesis seen in the embryo (Irion *et al.*, 2010). This first wave resembles the yolk sac primitive hematopoiesis, generating erythroid cells predominantly expressing the characteristic Beta-H1 hemoglobin (*Hbb-*

bh1) (Figure S2C). Our observation that a primitive erythroid-like wave is maximally generated at higher BMP4 and lower Activin is also consistent with the embryonic axis where higher BMP4 is found posteriorly associated with yolk sac mesoderm and higher nodal/activin more anteriorly with intra-embryonic mesoderm formation. We found cells expressing cardiac markers predominately at lower concentration of BMP4 (Figure 1D and S2B), consistent with previous publications (Palpant *et al.*, 2015). Higher Activin levels were required for endoderm generation whereas higher BMP4 was inhibitory, consistent with distal endoderm formation in the embryo (Yamamoto *et al.*, 2009) and endoderm differentiation in huiPSC studies (Loh *et al.*, 2014). Measurable numbers of primitive germ cells (PGC) were only seen at higher BMP4 concentrations and were inhibited by the highest Activin concentration, consistent with previous studies showing BMP4 is required for PGC formation in a murine organ culture system (Dudley *et al.*, 2007). Together, these studies demonstrate that sci-Plex robustly captures transcriptionally defined cell fates during mESC differentiation *in vitro*, identifying both temporal relationships and concentration-dependent effects on cell fates that recapitulate the expected developmental trajectories of their embryonic counterparts *in vivo*.

Intra-embryonic and extra-embryonic mesodermal fates are patterned by Activin and BMP4

In the embryo, cells within the primitive streak express T-box transcription factor *Brachyury*, but downregulate its expression as they migrate within the mesoderm wings and differentiate into mesoderm (Wilkinson *et al.*, 1990). During mESC differentiation,

cells expressing *Brachyury* were identified predominantly on day 4, two days following the addition of BMP4, Activin and VEGF (Figure 2A and B), consistent with previous studies (Irion *et al.*, 2010). As expected, *Brachyury* expression was downregulated as EB cells increasingly expressed mesodermal markers *Pdgfra* and *Kdr* between day 4 and 6, which was accelerated at higher Activin concentrations. EBs exposed to the highest Activin concentration contained the most cells expressing both *Flk1* and *Pdgfra* on both days 4 and 5 (Figures 2A and B), demonstrating that the pattern of gene expression by day and concentration revealed by sci-Plex is remarkably similar to that observed based on protein expression by flow cytometric analysis (compare Figures S1A and 2B).

Embryonic day E7.5-8 EECs exiting the PS express genes associated with hematopoiesis such as *Tal1* and *Runx1*, consistent with their rapid differentiation to produce the first primitive hematopoietic wave, whereas exiting IECs express genes generally associated with mesoderm (Saykali *et al.*, 2019). To map and determine numbers of cells bearing genes characteristic of PS, IEC and EEC, we determined relative aggregate expression gene-set scores (GSS) within the subset of EB clusters representing PS and early mesodermal patterning (Figure 2C). Gene-sets were generated from previous publications (Table 2). Cells with high PS GSS were primarily observed on day 4, while cells with the highest IEC or EEC GSS were found in a generally non-overlapping pattern, as expected (Figure 2D). Most cells with a high IEC GSS at all concentrations were observed on days 4 and 5, whereas those with highest EEC GSS were observed predominantly on days 5 and 6 (Figure 2E). Higher Activin generated more cells with high IEC GSS on days 4 and 5, whereas higher BMP4

inhibited generation of IEC and promoted generation of EECs on days 5 and 6 (Figure 2E). These results are consistent with the observation that EECs exit from the posterior PS where higher BMP4 levels are present, whereas IECs exit more anteriorly at higher Activin/Nodal concentrations (Saykali *et al.*, 2019).

sci-Plex reveals optimized conditions for generation of LPM and HE

In the embryo, LPM is the source of HE that generates intra-embryonic multilineage hematopoietic progenitors and HSCs (Prummel *et al.*, 2020). Cells with high LPM GSS were observed by day 4, peaked at day 5, and were diminished by day 6 with progression to hematovascular or other mesodermal fates, including cardiac (Figures 3A). Higher Activin concentrations promoted LPM compared to lower Activin concentrations, even at the lowest BMP4 concentration, with optimal LPM generation occurring at day 5 in conditions with the highest concentration of Activin (20ng/ml) and intermediate BMP4 (5ng/ml) (Figure 3A).

In the embryo proper, LPM (co-expressing *Flk1* and *Pdgfra*) destined for hematovascular fates simultaneously downregulate *Pdgfra* and increase expression of *Etv2*, a gene essential for generating vascular and hemogenic endothelium (Chestnut *et al.*, 2020; Kataoka *et al.*, 2011). Subsequently, a subset of *Flk1*⁺*Etv2*⁺ mesoderm cells upregulate VE-Cadherin (*Cdh5*) as they mature to differentiated endothelial cell types (HE/EC). Consistent with this developmental progression, our sci-Plex data revealed substantial numbers of *Flk1*⁺*Etv2*⁺*Cdh5*⁻ hematovascular mesoderm and *Flk1*⁺*Etv2*⁺*Cdh5*⁺ immature HE/EC predominantly on day 5 (Figures 3B and C). Similar to LPM, hematovascular mesoderm and immature HE/EC were optimally generated at

20ng/ml Activin and 5ng/ml BMP4. Downregulation of *Etv2* is required for further EC/HE maturation prior to the emergence of hematopoietic and various EC (arterial, venous, lymphatic) fates. Consistent with this, *Kdr⁺Etv2⁻Cdh5⁺* cells are increased between days 5 and 6 in our EB sci-Plex data, and by day 6, a subset also express *Dll4*, a Notch ligand expressed in arterial EC and HE (Figure 3B) (Duarte et al., 2004; Hadland et al., 2017). To validate the emergence of arterial HE/EC in the EB sci-Plex data, we calculated arterial endothelium GSS using two sets of arterial endothelial markers (Figure 3D) (Kalucka et al., 2020; Xu et al., 2018). In contrast to immature HE/EC, the highest Activin concentration generated fewer arterial HE/EC than lower activin concentrations, suggesting that persistent Activin exposure may be inhibitory to generation of arterial EC/HE from immature HE/EC (Figure 3C). Based on this result, we revised our protocol to shift EBs to a lower Activin concentration on day 5 to enhance arterial EC/HE production.

EBs contain a both yolk sac-like and intraembryonic-like HE

Yolk sac HE have been show to express LYVE1 (lymphatic vessel endothelial hyaluran receptor-1), while intra-embryonic HE are negative for this marker (Lee et al., 2016). Confirming previous studies, the majority of VE-Cadherin⁺DLL4⁺ HE/EC in E8 murine yolk sac co-express LYVE1, whereas VE-Cadherin⁺DLL4⁺ HE/EC from the E9 embryo proper largely lack LYVE1 expression by flow cytometry (Figure 4A). This result is consistent with control E8.5-E9 embryo-derived cells (separated as yolk sac and embryo proper) analyzed in parallel with EBs in our sci-Plex data, in which *Lyve1*

expression is predominantly detected in yolk sac HE/EC when compared to the embryo proper (Figure 4B, S3A).

Consistent with these findings *in vivo*, we identified both yolk sac-like $Cdh5^+Dll4^+Lyve1^+$ EC/HE and $Cdh5^+Dll4^+Lyve1^-$ HE/EC in our EB dataset, a subset of which also expresses $Cxcr4$, consistent with intra-embryonic origin (Dignum *et al.*, 2021; Simic *et al.*, 2020) (Figure 4C and S3B). To verify generation of EC/HE subtypes from mESC in this protocol, we performed FACS isolation of EB cells to sort the VE-Cadherin⁺LYVE1⁺ and VE-Cadherin⁺LYVE1⁻ EC/HE (Figure S3C). The hematopoietic potential of sorted cell populations was assessed via co-culture on AGM-ECs, which we have previously shown can robustly support EHT and generation of multilineage hematopoietic progeny (Dignum *et al.*, 2021; Hadland *et al.*, 2015). Both subsets generated CD45⁺ hematopoietic cells during co-culture (Figure S3D) and myeloid/erythroid colonies in methylcellulose assays (Figure S3E). B- and T-lymphoid potential, assessed by secondary OP9 and OP9-DLL4 stromal cells co-culture, respectively, demonstrated that LYVE1⁻ HE generated abundant CD4⁺CD8⁺ T-cells, while yolk-sac like LYVE1⁺ EC/HE did not (Figure S3F). However, the LYVE1⁺ population did generate B1a-like cells (AA4.1⁺CD19⁺B220^{-/lo}; Figure S3G), an embryonic restricted B-cell subset predominantly emerging in the yolk sac (Yoshimoto *et al.*, 2011). Taken together, these findings are consistent with the hematopoietic potential of their *in vivo* counterparts.

Although we did observe the natural progression of hematovascular markers, including Runx1, within the HE/EC subset (Figure S4A) and detect a minor population of $Cdh5^+Dll4^+Cxcr4^+$ EC/HE in our sci-Plex data (Figure S3B) suggestive of HSC-

competent HE based on our prior studies *in vivo* (Dignum *et al.*, 2021; Hadland *et al.*, 2017), several other notable genes that drive EHT in HSC-competent HE (egs *Gfi1*, *Hlf*, *Myb*), as reported by our lab and others (Dignum *et al.*, 2021; Hadland *et al.*, 2022; Hou *et al.*, 2020), are distinctly lacking in our EB dataset (Figure S4B). This underscores the notion that acquisition of key transcriptional programs prior to undergoing EHT are necessary to endow HSC potential and highlights the need for future studies into driving expression of these critical programs to achieve HSC fate from mESC *in vitro*.

sci-Plex highlights a role for retinoic acid signaling in HE fate

Retinoic acid (RA) signaling has previously been shown to inhibit yolk sac-like HE and promote intraembryonic AGM-like type HE from human PSC *in vitro*, and is required for development of long-term HSC in the AGM *in vivo* (Chanda *et al.*, 2013; Luff *et al.*, 2022). RA precursors are not present in serum-free media used in this protocol. However, a subset of EB-derived HE/EC do express RA receptors *Rarg* and *Rara*, indicating competence to respond to exogenous retinoic acid (Figure S5A). Hence, we tested whether addition of all trans retinoic acid (ATRA) to EB cultures on day 5 altered HE/EC cell fates. We found that while exogenous ATRA did not impact total cell numbers or generation of VE-Cadherin⁺DLL4⁺ arterial EC/HE (Figure S5B and C), it significantly enhanced frequency of VE-Cadherin⁺DLL4⁺ cells expressing CXCR4⁺ while reducing those expressing LYVE1⁺, in a dose-dependent manner (Figure 4D). ATRA also significantly reduced numbers of VE-cadherin⁺CD41⁺ cells (Figure 4E), suggesting inhibition of yolk sac-like hematopoiesis that predominates in the cultures at this stage. These results are consistent with huiPSC studies showing ATRA inhibited

primitive hematopoiesis but improved definitive (Luff *et al.*, 2022). However, we have not observed reproducible HSC activity from ATRA-treated mESC-derived EC/HE based on transplantation assays following AGM-EC co-culture (not shown), suggesting that additional exogenous factors necessary to promote efficient induction of HSCs from HE at this stage in the protocol remain to be determined in future studies.

Discussion

In this study, we present several advances towards elucidating the complex, combinatorial signaling pathways involved in driving early cell fate decisions and manipulating those pathways to optimize in vitro hematopoietic differentiation of mESCs. First, we provide a proof of principle for utilizing sci-Plex scRNA-seq to dissect the complexities of dose and temporal integration of interacting signal pathways to determine cell fates. Next, we created a single cell atlas of early mesoderm patterning and various downstream lineages that can be used for discovery of additional cell fate choices. Finally, we used sci-Plex to optimize the production of intra-embryonic mesodermal fates, including LPM and HE, which serve as the precursors to HSC.

Here, we employed sci-Plex to decipher the combinatorial dose and time-dependent effects of an Activin/BMP4 gradient on mESC-derived mesoderm and hematopoietic lineages. High-throughput single cell methodologies such as sci-Plex are essential to dissect the complex nature of cell signaling pathways during early embryonic cell fate decisions, notably the subtle changes in timing and/or dosage of morphogens/cytokines which can have profound ramifications on downstream cell types. While sci-Plex has previously been applied to identify heterogeneity in drug response (Srivatsan *et al.*, 2020) or understand cell fate determinants in retinal organoids (Tresenrider *et al.*, 2023), here we utilized sci-Plex to deconvolute the complex, heterogeneous cellular landscape of an actively differentiating pluripotent stem cell population in response to combinations of morphogens. To this end, we were able to define the developmental trajectory of hematopoiesis accurately and confidently within our EBs, based on published markers and gene sets (Figure 1, 2 and Figure S2).

This underscores the notion that sci-Plex allows for high throughput single-cell screening of a population with adequate resolution to distinguish subtle cell types and transitions in developmental states. Furthermore, it suggests that sci-Plex could have profound applications toward modeling both normal development and disease processes in hPSC and human organoid models. For example, sci-Plex could be employed to better understand the heterogeneity within the tumor microenvironment of various cancers, notably the interactions/signaling between cancer stem cells and the niche, in order to highlight novel therapeutic targets, and subsequently observe the cellular/transcriptional response to such perturbations.

The data generated in our study provides a robust single cell atlas that accurately recapitulates expected cell fates in early embryonic patterning in response to Activin/BMP4 gradients. Of note, our transcriptional analysis identified lineage pathways undertaken by mESCs differentiating primarily into various mesodermal lineages including cardiomyocytes, allantois, mesenchyme, and hematovascular cell types (Figure 1C and D). With each of these lineage pathways, there are key cell fate choices being made at the transcriptional level that are the primary drivers of subsequent differentiation, many of which remain poorly characterized. Here, we focused on defining some of those cell fate decisions important for generation of intra-embryonic HE, which serve as precursors to HSCs. However, this dataset provides a robust resource for analysis of other novel cell fate decisions during mesodermal patterning and differentiation resulting from combinatorial Activin/nodal and BMP4 signals.

We utilized sci-Plex to assess intra-embryonic vs extra-embryonic fates and optimize production of LPM, including HE, within our mESC differentiation protocol.

Notably, we observed that higher Activin concentrations coupled with lower BMP4 favored IEC output and that LPM production peaked at day 5 of our protocol in 20 ng/mL Activin and 5 ng/mL BMP4 conditions (Figures 2D-E and 3A). Moreover, the transition from LPM to hematovascular fates, specifically HE, requires acquisition of endothelial programs (Chestnut *et al.*, 2020; Kataoka *et al.*, 2011), which were inhibited by high Activin (Figure 3B-C), leading to a revised protocol with a lower Activin concentration after day 5. This optimization led to increased production of both yolk sac-like and intraembryonic-like HE, which we further defined through expression of LYVE1 and subsequent lineage assays (Figure 4 and FigureS3). Moreover, we observed these two HE populations are differentially responsive to RA signaling consistent with findings in human PSC (Chanda *et al.*, 2013; Luff *et al.*, 2022).

Ultimately, while we were able to utilize sci-Plex to enhance production of HE, we did not observe long-term engraftment following transplantation into mice, suggesting there are additional barriers to efficient generation of HSC-competent HE. Of note, we observed minimal or absent expression of several key genes identified by our lab and others that define AGM arterial HE *in vivo* (Dignum *et al.*, 2021; Hou *et al.*, 2020). This may be, in part, due to the lack of key signaling pathways in the EBs that are typically present during development. Thus, future studies exploring additional exogenous factors important for HSC-competence that are currently absent from our EB system are warranted.

Taken together, the results from this study underscore the importance of deciphering the complex, combinatorial signaling that occurs during early cell fate decisions and how that knowledge can be applied to optimize *in vitro* differentiation of

pluripotent stem cells to a desired cell type or lineage. High-throughput tools, like sci-Plex utilized here, which allow researchers to measure transcriptional responses to highly multiplexed perturbations at single-cell resolution in a time-efficient and cost-effective manner will be a key aspect in any subsequent investigations. Thus, we present this study as a resource for the dataset generated on early mesodermal patterning, as well as a roadmap for the design of future studies to optimize specific cell fates in dynamic stem cell differentiation protocols.

References

- Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* *566*, 496-502. [10.1038/s41586-019-0969-x](https://doi.org/10.1038/s41586-019-0969-x).
- Chanda, B., Ditadi, A., Iscove, N.N., and Keller, G. (2013). Retinoic acid signaling is essential for embryonic hematopoietic stem cell development. *Cell* *155*, 215-227. [10.1016/j.cell.2013.08.055](https://doi.org/10.1016/j.cell.2013.08.055).
- Chestnut, B., Casie Chetty, S., Koenig, A.L., and Sumanas, S. (2020). Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle. *Nat Commun* *11*, 2796. [10.1038/s41467-020-16515-y](https://doi.org/10.1038/s41467-020-16515-y).
- Clarke, R.L., Yzaguirre, A.D., Yashiro-Ohtani, Y., Bondue, A., Blanpain, C., Pear, W.S., Speck, N.A., and Keller, G. (2013). The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat Cell Biol* *15*, 502-510. [10.1038/ncb2724](https://doi.org/10.1038/ncb2724).
- Dignum, T., Varnum-Finney, B., Srivatsan, S.R., Dozono, S., Waltner, O., Heck, A.M., Ishida, T., Nourigat-McKay, C., Jackson, D.L., Rafii, S., et al. (2021). Multipotent progenitors and hematopoietic stem cells arise independently from hemogenic endothelium in the mouse embryo. *Cell Rep* *36*, 109675. [10.1016/j.celrep.2021.109675](https://doi.org/10.1016/j.celrep.2021.109675).
- Ditadi, A., Sturgeon, C.M., and Keller, G. (2017). A view of human haematopoietic development from the Petri dish. *Nat Rev Mol Cell Biol* *18*, 56-67. [10.1038/nrm.2016.127](https://doi.org/10.1038/nrm.2016.127).

Duarte, A., Hirashima, M., Benedito, R., Trindade, A., Diniz, P., Bekman, E., Costa, L., Henrique, D., and Rossant, J. (2004). Dosage-sensitive requirement for mouse Dll4 in artery development. *Genes Dev* 18, 2474-2478. 10.1101/gad.1239004.

Dudley, B.M., Runyan, C., Takeuchi, Y., Schaible, K., and Molyneaux, K. (2007). BMP signaling regulates PGC numbers and motility in organ culture. *Mech Dev* 124, 68-77. 10.1016/j.mod.2006.09.005.

Frame, J.M., Fegan, K.H., Conway, S.J., McGrath, K.E., and Palis, J. (2016). Definitive Hematopoiesis in the Yolk Sac Emerges from Wnt-Responsive Hemogenic Endothelium Independently of Circulation and Arterial Identity. *Stem Cells* 34, 431-444. 10.1002/stem.2213.

Freire, A.G., and Butler, J.M. (2020). Blood making: learning what to put into the dish. *F1000Res* 9. 10.12688/f1000research.21245.1.

Giles, P.B., Candy, C.L., Fleming, P.A., Owens, R.W., Argraves, W.S., and Drake, C.J. (2005). VEGF directs newly gastrulated mesoderm to the endothelial lineage. *Dev Biol* 279, 169-178. 10.1016/j.ydbio.2004.12.011.

Hadland, B., Varnum-Finney, B., Dozono, S., Dignum, T., Nourigat-McKay, C., Heck, A.M., Ishida, T., Jackson, D.L., Itkin, T., Butler, J.M., et al. (2022). Engineering a niche supporting hematopoietic stem cell development using integrated single-cell transcriptomics. *Nat Commun* 13, 1584. 10.1038/s41467-022-28781-z.

Hadland, B., and Yoshimoto, M. (2018). Many layers of embryonic hematopoiesis: new insights into B-cell ontogeny and the origin of hematopoietic stem cells. *Exp Hematol* 60, 1-9. 10.1016/j.exphem.2017.12.008.

Hadland, B.K., Varnum-Finney, B., Mandal, P.K., Rossi, D.J., Poulos, M.G., Butler, J.M., Rafii, S., Yoder, M.C., Yoshimoto, M., and Bernstein, I.D. (2017). A Common Origin for B-1a and B-2 Lymphocytes in Clonal Pre- Hematopoietic Stem Cells. *Stem Cell Reports* 8, 1563-1572. [10.1016/j.stemcr.2017.04.007](https://doi.org/10.1016/j.stemcr.2017.04.007).

Hadland, B.K., Varnum-Finney, B., Poulos, M.G., Moon, R.T., Butler, J.M., Rafii, S., and Bernstein, I.D. (2015). Endothelium and NOTCH specify and amplify aorta-gonad-mesonephros-derived hematopoietic stem cells. *J Clin Invest* 125, 2032-2045. [10.1172/JCI80137](https://doi.org/10.1172/JCI80137).

Hou, S., Li, Z., Zheng, X., Gao, Y., Dong, J., Ni, Y., Wang, X., Li, Y., Ding, X., Chang, Z., et al. (2020). Embryonic endothelial evolution towards first hematopoietic stem cells revealed by single-cell transcriptomic and functional analyses. *Cell Res* 30, 376-392. [10.1038/s41422-020-0300-2](https://doi.org/10.1038/s41422-020-0300-2).

Irion, S., Clarke, R.L., Luche, H., Kim, I., Morrison, S.J., Fehling, H.J., and Keller, G.M. (2010). Temporal specification of blood progenitors from mouse embryonic stem cells and induced pluripotent stem cells. *Development* 137, 2829-2839. [10.1242/dev.042119](https://doi.org/10.1242/dev.042119).

Jackson, S.A., Schiesser, J., Stanley, E.G., and Elefanty, A.G. (2010). Differentiating embryonic stem cells pass through 'temporal windows' that mark responsiveness to exogenous and paracrine mesendoderm inducing signals. *PLoS One* 5, e10706. [10.1371/journal.pone.0010706](https://doi.org/10.1371/journal.pone.0010706).

Kalucka, J., de Rooij, L.P.M.H., Goveia, J., Rohlenova, K., Dumas, S.J., Meta, E., Conchinha, N.V., Taverna, F., Teuwen, L.A., Veys, K., et al. (2020). Single-Cell Transcriptome Atlas of Murine Endothelial Cells. *Cell* 180, 764-779.e720. [10.1016/j.cell.2020.01.015](https://doi.org/10.1016/j.cell.2020.01.015).

Kataoka, H., Hayashi, M., Nakagawa, R., Tanaka, Y., Izumi, N., Nishikawa, S., Jakt, M.L., and Tarui, H. (2011). Etv2/ER71 induces vascular mesoderm from Flk1+PDGFR α + primitive mesoderm. *Blood* 118, 6975-6986. 10.1182/blood-2011-05-352658.

Kattman, S.J., Witty, A.D., Gagliardi, M., Dubois, N.C., Niapour, M., Hotta, A., Ellis, J., and Keller, G. (2011). Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* 8, 228-240. 10.1016/j.stem.2010.12.008.

Lee, L.K., Ghorbanian, Y., Wang, W., Wang, Y., Kim, Y.J., Weissman, I.L., Inlay, M.A., and Mikkola, H.K. (2016). LYVE1 Marks the Divergence of Yolk Sac Definitive Hemogenic Endothelium from the Primitive Erythroid Lineage. *Cell Rep* 17, 2286-2298. 10.1016/j.celrep.2016.10.080.

Loh, K.M., Ang, L.T., Zhang, J., Kumar, V., Ang, J., Auyeong, J.Q., Lee, K.L., Choo, S.H., Lim, C.Y., Nichane, M., et al. (2014). Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell* 14, 237-252. 10.1016/j.stem.2013.12.007.

Luff, S.A., Creamer, J.P., Valsoni, S., Dege, C., Scarfò, R., Dacunto, A., Cascione, S., Randolph, L.N., Cavalca, E., Merelli, I., et al. (2022). Identification of a retinoic acid-dependent haemogenic endothelial progenitor from human pluripotent stem cells. *Nat Cell Biol.* 10.1038/s41556-022-00898-9.

Müller, A.M., Medvinsky, A., Strouboulis, J., Grosveld, F., and Dzierzak, E. (1994). Development of hematopoietic stem cell activity in the mouse embryo. *Immunity* 1, 291-301.

Nostro, M.C., Cheng, X., Keller, G.M., and Gadue, P. (2008). Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood. *Cell Stem Cell* 2, 60-71. [10.1016/j.stem.2007.10.011](https://doi.org/10.1016/j.stem.2007.10.011).

Palis, J., Malik, J., McGrath, K.E., and Kingsley, P.D. (2010). Primitive erythropoiesis in the mammalian embryo. *Int J Dev Biol* 54, 1011-1018. [10.1387/ijdb.093056jp](https://doi.org/10.1387/ijdb.093056jp).

Palpant, N.J., Pabon, L., Roberts, M., Hadland, B., Jones, D., Jones, C., Moon, R.T., Ruzzo, W.L., Bernstein, I., Zheng, Y., and Murry, C.E. (2015). Inhibition of β -catenin signaling respecifies anterior-like endothelium into beating human cardiomyocytes. *Development* 142, 3198-3209. [10.1242/dev.117010](https://doi.org/10.1242/dev.117010).

Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* 566, 490-495. [10.1038/s41586-019-0933-9](https://doi.org/10.1038/s41586-019-0933-9).

Prummel, K.D., Nieuwenhuize, S., and Mosimann, C. (2020). The lateral plate mesoderm. *Development* 147. [10.1242/dev.175059](https://doi.org/10.1242/dev.175059).

Rivera-Pérez, J.A., and Hadjantonakis, A.K. (2014). The Dynamics of Morphogenesis in the Early Mouse Embryo. *Cold Spring Harb Perspect Biol* 7. [10.1101/cshperspect.a015867](https://doi.org/10.1101/cshperspect.a015867).

Saykali, B., Mathiah, N., Nahaboo, W., Racu, M.L., Hammou, L., Defrance, M., and Migeotte, I. (2019). Distinct mesoderm migration phenotypes in extra-embryonic and embryonic regions of the early mouse embryo. *Elife* 8. [10.7554/eLife.42434](https://doi.org/10.7554/eLife.42434).

Simic, M., Manosalva, I., Spinelli, L., Gentek, R., Shayan, R.R., Siret, C., Girard-Madoux, M., Wang, S., de Fabritus, L., Verschoor, J., et al. (2020). Distinct Waves from

the Hemogenic Endothelium Give Rise to Layered Lymphoid Tissue Inducer Cell Ontogeny. *Cell Rep* 32, 108004. [10.1016/j.celrep.2020.108004](https://doi.org/10.1016/j.celrep.2020.108004).

Srivatsan, S.R., McFaline-Figueroa, J.L., Ramani, V., Saunders, L., Cao, J., Packer, J., Pliner, H.A., Jackson, D.L., Daza, R.M., Christiansen, L., et al. (2020). Massively multiplex chemical transcriptomics at single-cell resolution. *Science* 367, 45-51. [10.1126/science.aax6234](https://doi.org/10.1126/science.aax6234).

Tober, J., Koniski, A., McGrath, K.E., Vemishetti, R., Emerson, R., de Mesy-Bentley, K.K., Waugh, R., and Palis, J. (2007). The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. *Blood* 109, 1433-1441. [10.1182/blood-2006-06-031898](https://doi.org/10.1182/blood-2006-06-031898).

Tresenrider, A., Sridhar, A., Eldred, K.C., Cuschieri, S., Hoffer, D., Trapnell, C., and Reh, T.A. (2023). Single-cell sequencing of individual retinal organoids reveals determinants of cell-fate heterogeneity. *Cell Rep Methods* 3, 100548. [10.1016/j.crmeth.2023.100548](https://doi.org/10.1016/j.crmeth.2023.100548).

Wilkinson, D.G., Bhatt, S., and Herrmann, B.G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* 343, 657-659. [10.1038/343657a0](https://doi.org/10.1038/343657a0).

Xu, C., Gao, X., Wei, Q., Nakahara, F., Zimmerman, S.E., Mar, J., and Frenette, P.S. (2018). Stem cell factor is selectively secreted by arterial endothelial cells in bone marrow. *Nat Commun* 9, 2449. [10.1038/s41467-018-04726-3](https://doi.org/10.1038/s41467-018-04726-3).

Yamamoto, M., Beppu, H., Takaoka, K., Meno, C., Li, E., Miyazono, K., and Hamada, H. (2009). Antagonism between Smad1 and Smad2 signaling determines the site of distal visceral endoderm formation in the mouse embryo. *J Cell Biol* 184, 323-334. [10.1083/jcb.200808044](https://doi.org/10.1083/jcb.200808044).

Yoshimoto, M., Montecino-Rodriguez, E., Ferkowicz, M.J., Porayette, P., Shelley, W.C., Conway, S.J., Dorshkind, K., and Yoder, M.C. (2011). Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. *Proc Natl Acad Sci U S A* *108*, 1468-1473. [10.1073/pnas.1015841108](https://doi.org/10.1073/pnas.1015841108).

Zhu, Q., Gao, P., Tober, J., Bennett, L., Chen, C., Uzun, Y., Li, Y., Howell, E.D., Mumau, M., Yu, W., et al. (2020). Developmental trajectory of prehematopoietic stem cell formation from endothelium. *Blood* *136*, 845-856. [10.1182/blood.2020004801](https://doi.org/10.1182/blood.2020004801).

Figure Legends

Figure 1: sci-Plex analysis reveals cell types generated in embryoid body

differentiation. A: Diagram of sci-Plex experimental setup. B: UMAP and cluster analysis of scRNA-seq data from EB differentiation sci-Plex experiment (see STAR Methods for details). C: Cell type classification based on diagnostic markers (see STAR Methods for details). D: (Top) Cell type classification for each differentiation time point collected. (Bottom) Heatmaps for number of cells generated for notable cell types per condition across differentiation time points collected.

Figure S1: Phenotypic analysis of sci-Plex conditions. Expression of Flk1 (Kdr) vs Pdgfr α (A) or VECadherin (Cdh5) (B). Gates are drawn based off isotype controls. Red boxes indicate chosen conditions for optimal lateral plate mesoderm production and subsequent hematopoietic output.

Figure S2: Cell type classification of sci-Plex dataset. A: Heatmap of diagnostic marker expression for cell type based on cluster. B: Cell type classification based on diagnostic markers subset for Activin/BMP conditions. C: Gene expression heatmap for primitive erythrocyte marker hemoglobin Z, beta-like embryonic chain (*Hbb-bh1*) within the full EB sci-Plex dataset.

Figure 2: Activin and BMP4 exposure dictates early cell fate patterning. A: Gene expression of early mesodermal cell fate markers across differentiation time points collected. B: Heatmaps for number of cells generated per cell fate markers from (A), across differentiation time points collected. C: Subset of clusters representing primitive streak and mesodermal patterning in UMAP space. D: Heatmaps of gene-set scores for signature gene sets defining primitive streak (Pujian-Sala et al, 2019), extra-embryonic cells and intra-embryonic cells (Saykali et al, 2019) within the primitive streak and mesodermal subset from (C). E: Heatmaps for number of cells above the gene set midpoint from (D) across differentiation time points collected.

Figure 3: sci-Plex reveals optimal Activin/BMP conditions for generation of LPM and HE. A: (Top) Heatmap for gene-set scores for lateral plate mesoderm signature gene set (Zhao and Choi 2019, Development, Prummel et al 2019) within the primitive streak and mesodermal subset from (2C). (Bottom) Heatmaps for number of cells above the gene set midpoint across differentiation time points collected. B: Gene expression of hematovascular cell fate markers across differentiation time points collected. C: Heatmaps for number of cells generated per cell fate markers from (B), across differentiation time points collected. D: Heatmaps of gene-set scores for signature gene sets defining arterial endothelial cells (Kalucka et al, Xu et al), across the entire sci-Plex dataset.

Figure 4: sci-Plex highlights production of intra- and extra-embryonic HE and a

role for retinoic acid signaling. A: Expression of Lyve1 from embryo proper (left) and yolk sac (right) vascular tissues dissected from early E9 embryos. Cells were gated on $Dll4^+VECadherin^+$. Gates are drawn based on isotype controls. B: Gene expression heatmap for *Lyve1* expression for embryo proper (left) and yolk sac (right) tissues within the HE/EC subset of embryo dissections collected for sci-Plex. C: Gene expression of $Dll4^+VECadherin^+Lyve1^{+/-}$ within the HE/EC subset of EB sci-Plex data. D: Percent expression of Lyve1 (blue) and Cxcr4 (red) within the $Dll4^+VECadherin^+$ population across ATRA concentrations (see Figure S5B for associated dot plots). Asterisks designate significant difference compared to 0 μ M. Error bars represent standard deviation. E: Number of VECadherin⁺CD41⁺ cells across ATRA concentrations (see Figure S5B for associated dot plots). P-values listed are comparison to 0 μ M.

Figure S3: Functional assessment of hematopoietic activity for Lyve1^{+/-}

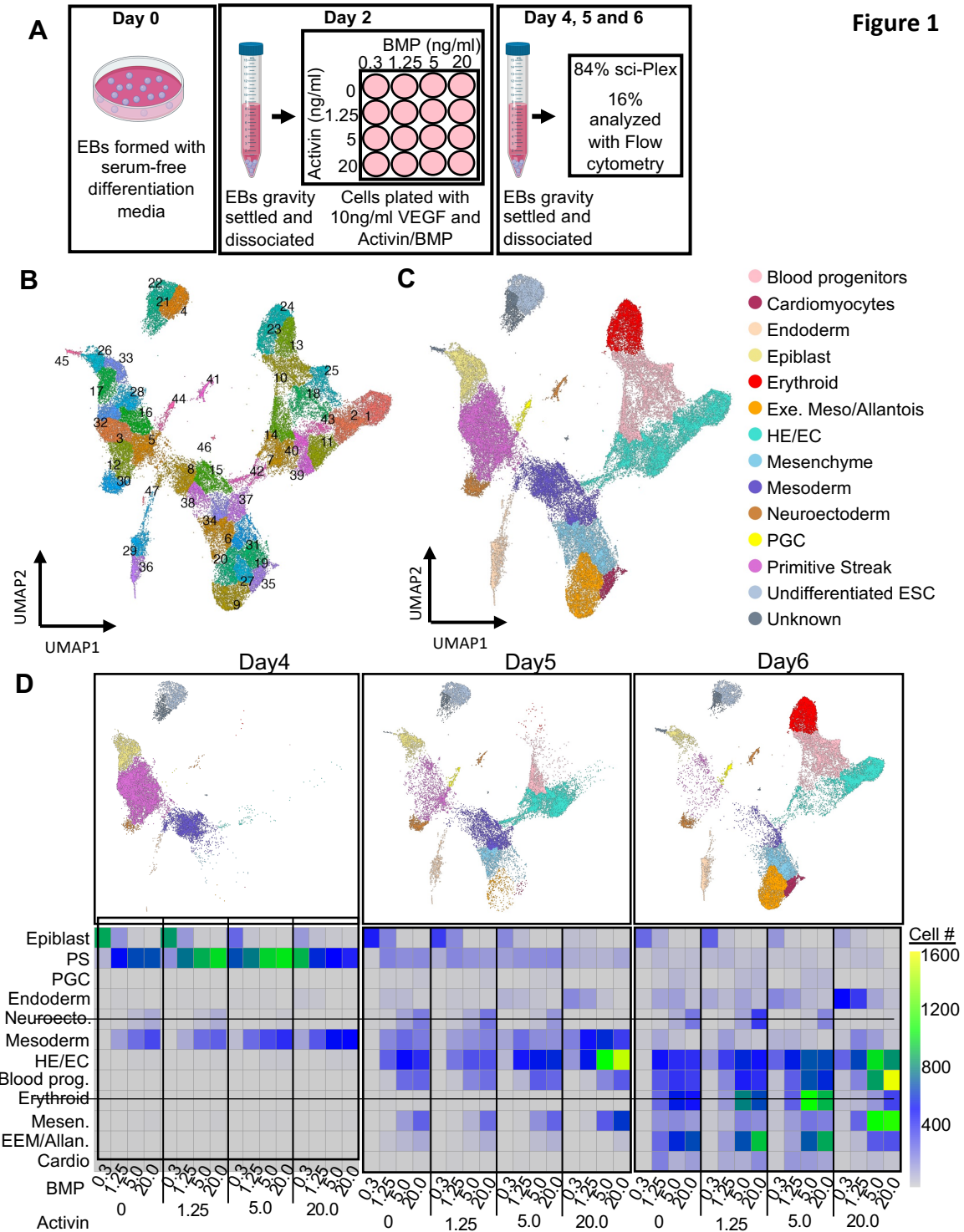
populations generated from EB differentiation. A: Gene expression heatmap for VECadherin (*Cdh5*) within embryo dissection sci-Plex dataset. B: Gene expression of $Dll4^+VECadherin^+Cxcr4^+$ within the HE/EC subset of EB sci-Plex data. C: Expression of VECadherin and Lyve1 from EBs collected for sorting on Day 7 of differentiation. Green and orange boxes indicate the sorted populations. D: Phenotypic expression of Lyve1⁻ (left) and Lyve1⁺ (right) cells sorted from (C) for the pan-hematopoietic marker CD45 and myeloid markers Gr1 and F480 within the VECadherin⁻ population of cells after 13 days of co-culture. E: Images of CFU-E (left) and CFU-GEMM (right) colonies generated in CFU assay for cells collected from (D). F: Phenotypic expression of Tcell

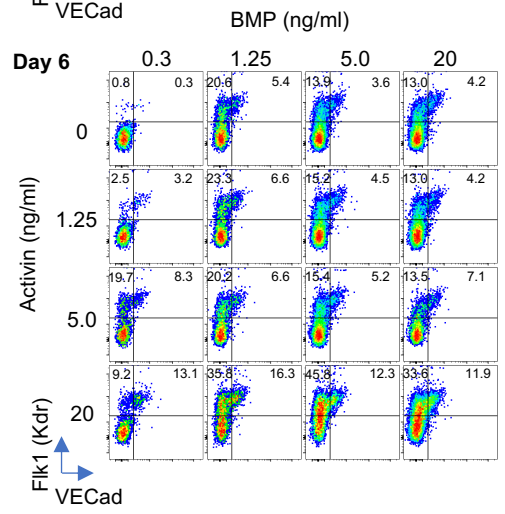
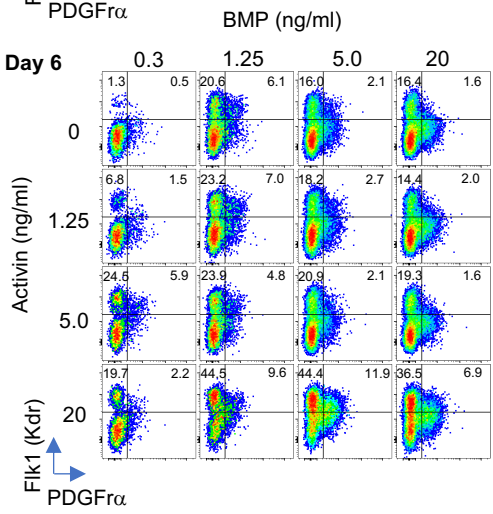
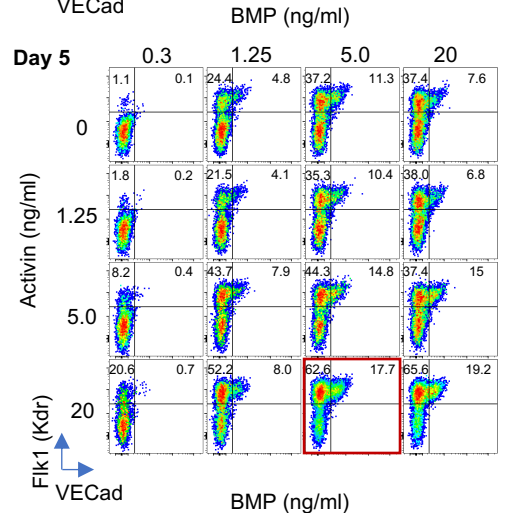
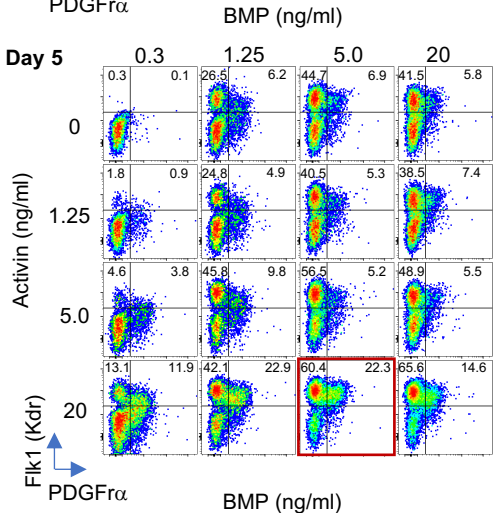
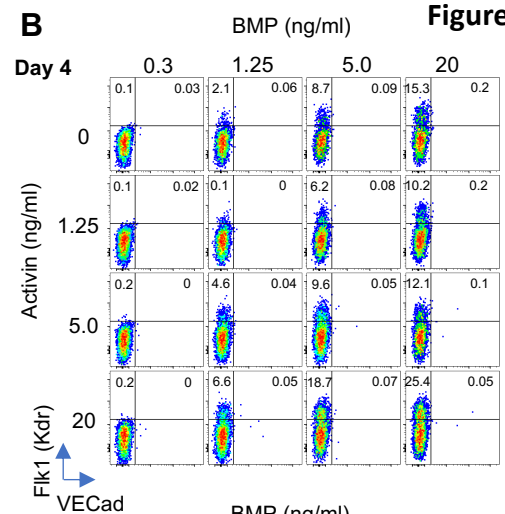
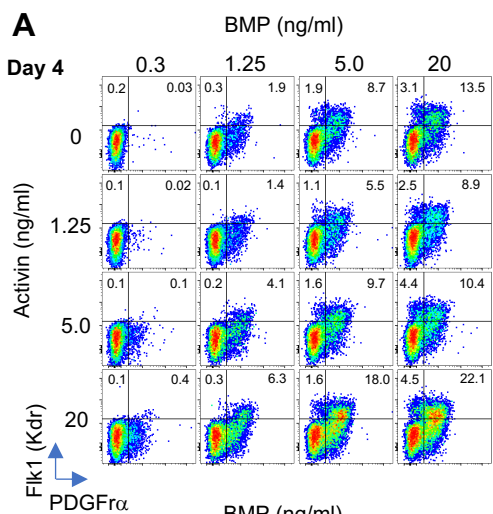
markers CD44⁻CD25⁺CD8⁺CD4⁺ for Lyve1⁻ (top) and Lyve1⁺ (bottom) cells collected from (D) and subjected to co-culture on OP9-Dll4 stromal cells. G: Phenotypic expression of Bcell markers CD19⁻AA4.1⁺B220⁻CD11b⁺ for Lyve1⁺ cells collected from (D) and subjected to co-culture on OP9 stromal cells. For all dot plots in C-G, gates are drawn based off isotype controls.

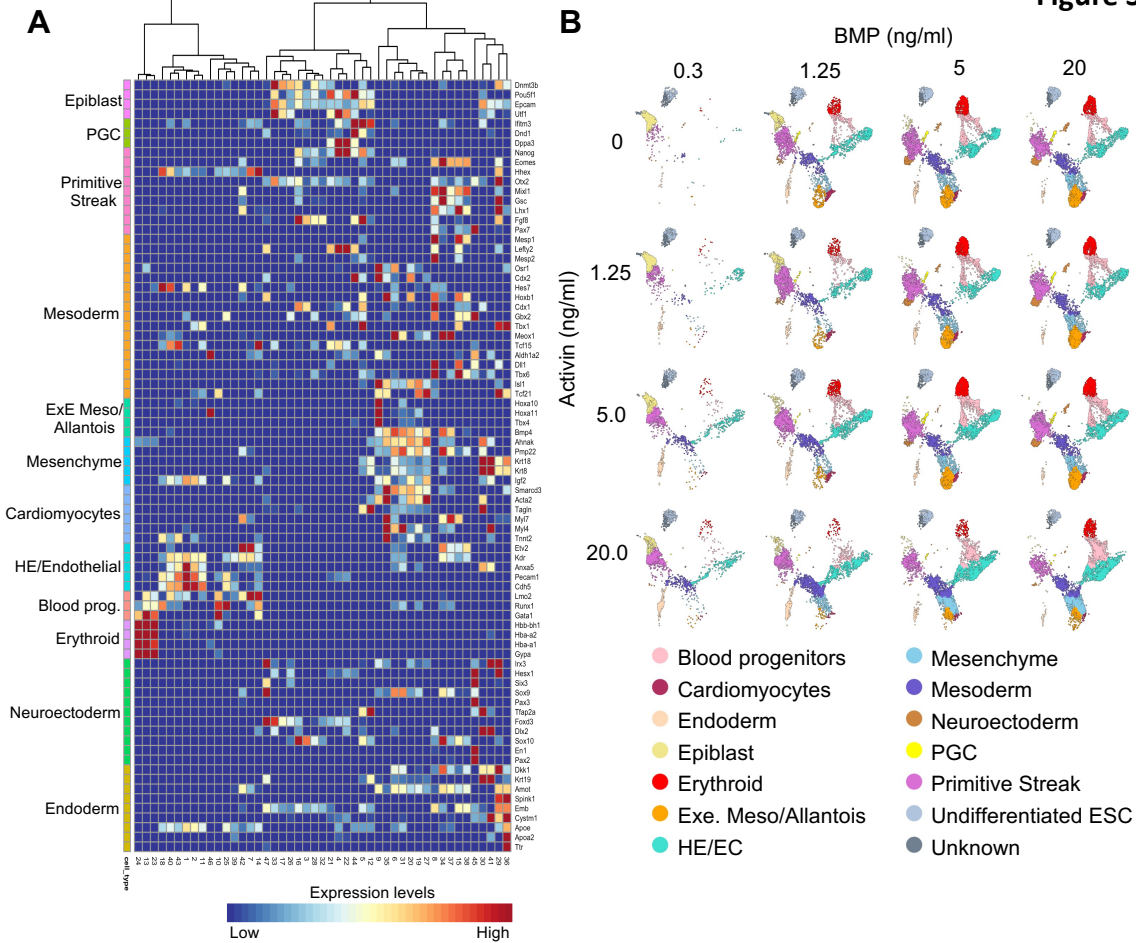
Figure S4: Expression of key hematopoietic development markers from sci-Plex data. Gene expression heatmaps for hematovascular (A) and EHT (B) markers within the HE/EC subset of EB sci-Plex data.

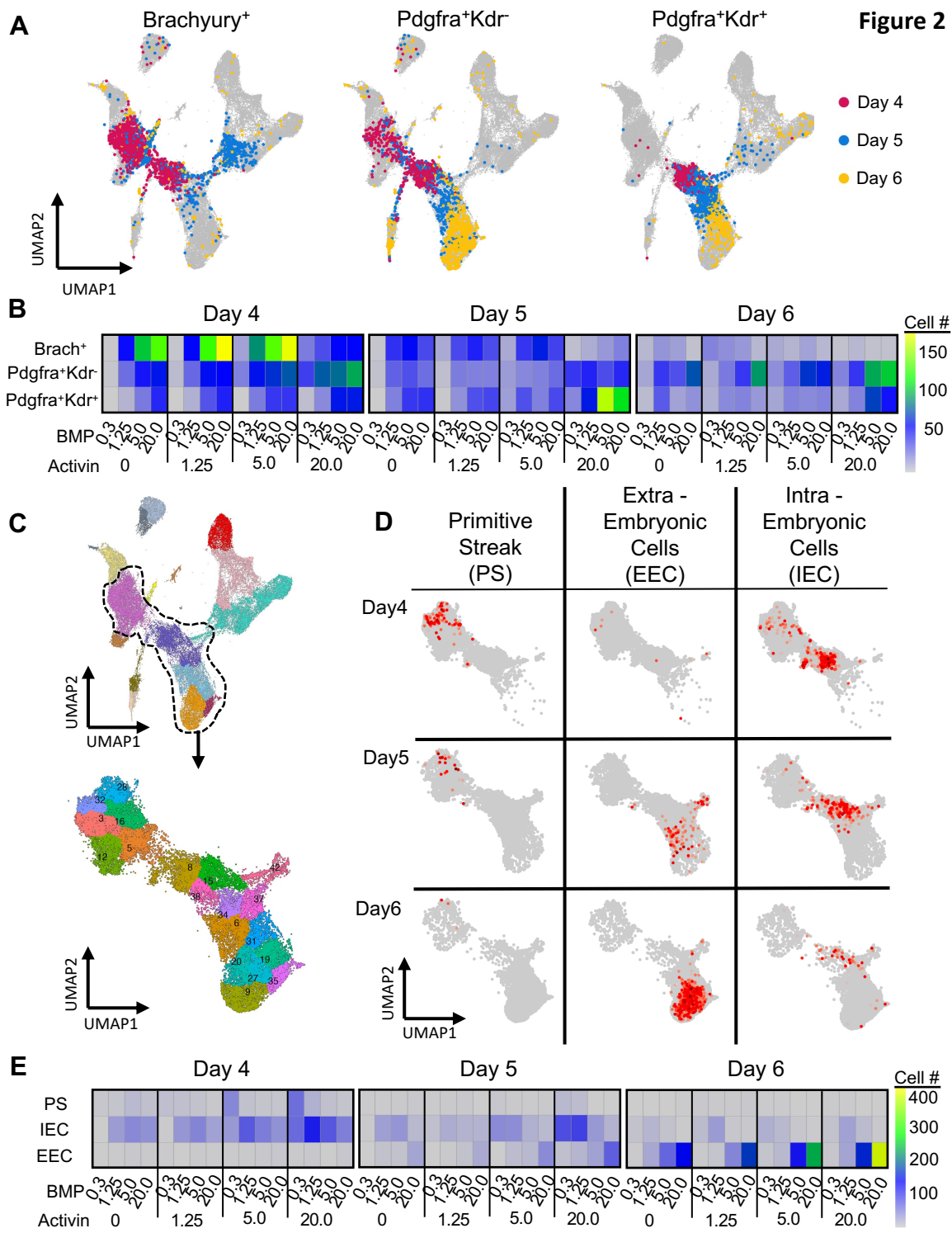
Figure S5: Retinoic acid signaling influences HE fate in EB differentiation protocol. A: Gene expression heatmaps retinoic acid receptors within the HE/EC subset of EB sci-Plex data. B: Phenotypic expression of Lyve1 (left) and Cxcr4 (middle) within the Dll4⁺VECadherin⁺ population, and CD41 x VECadherin (right). Gates are drawn based off isotype controls. C: Total (left) or VECadherin⁺Dll4⁺ number of cells across ATRA concentrations. Error bars represent standard deviation. Comparisons to 0 μ M were non-significant, as listed.

Figure 1









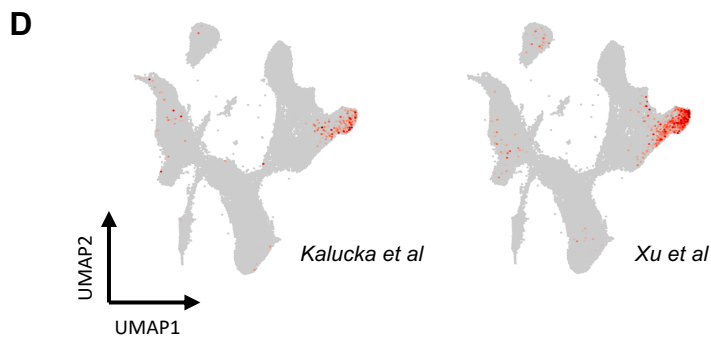
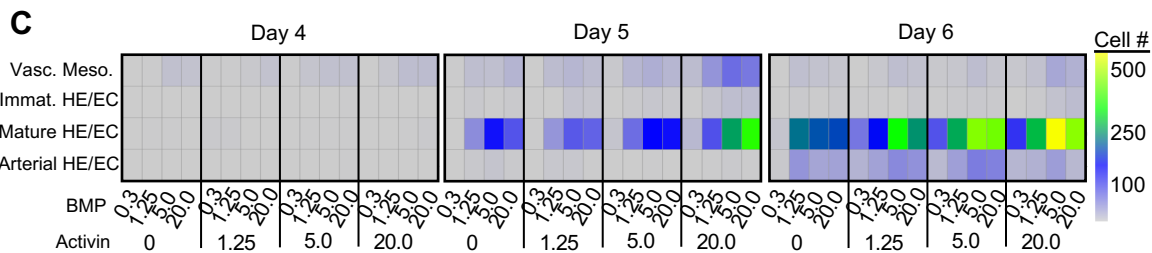
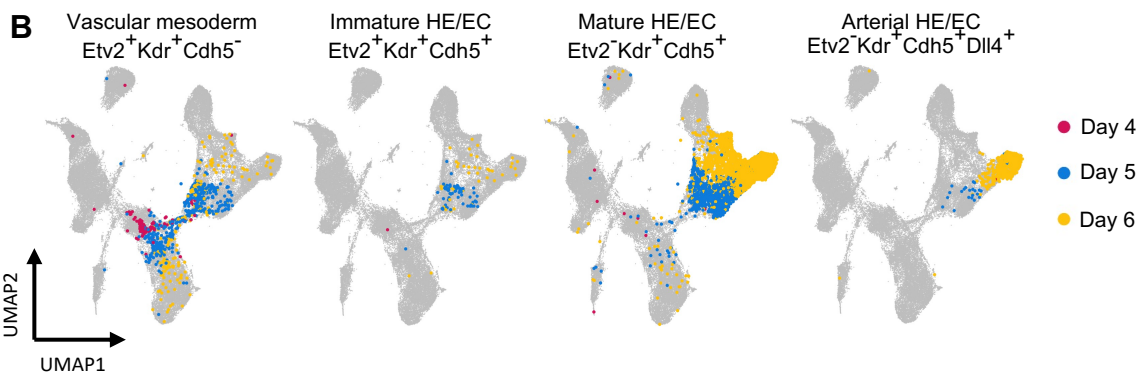
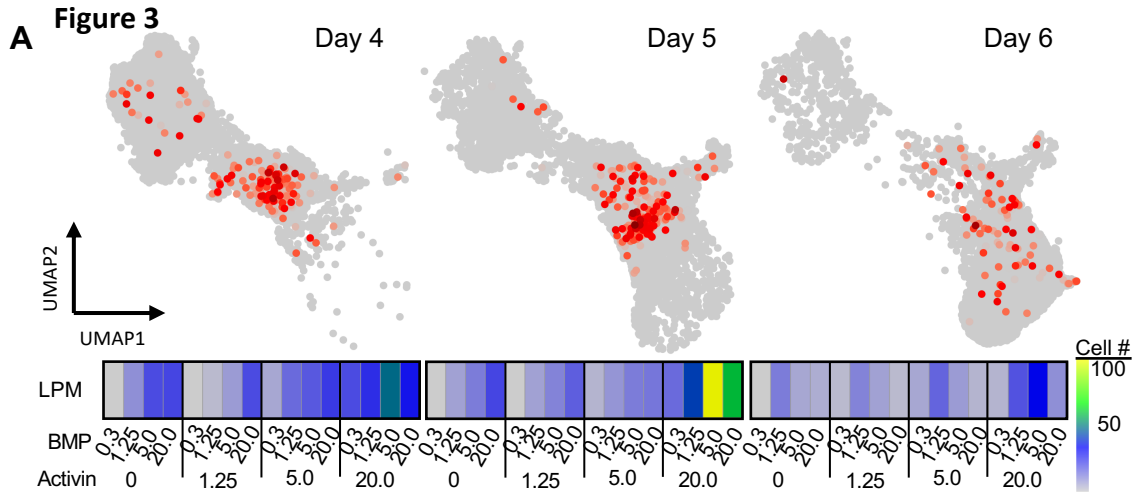
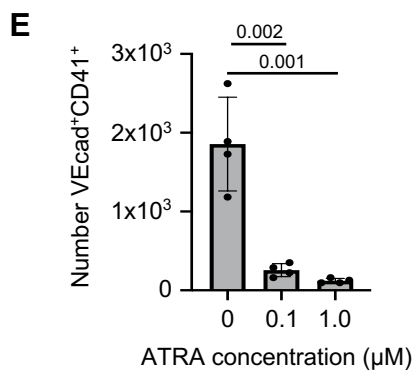
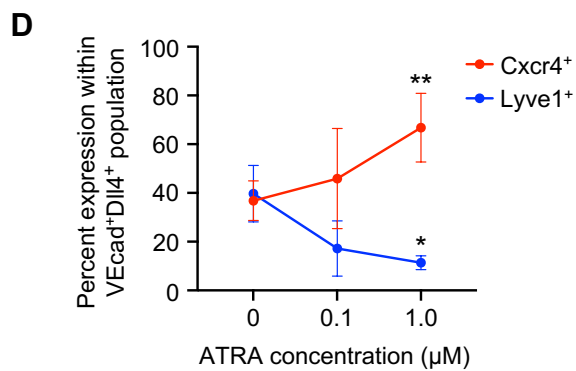
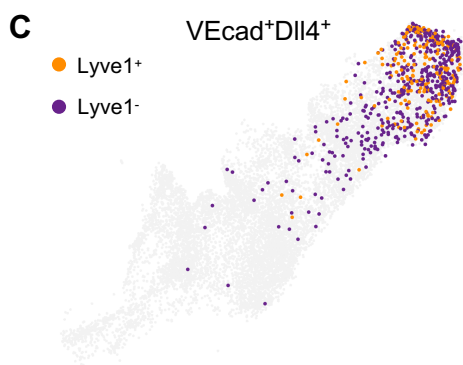
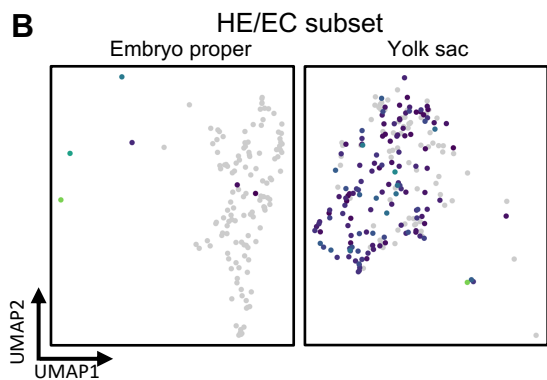
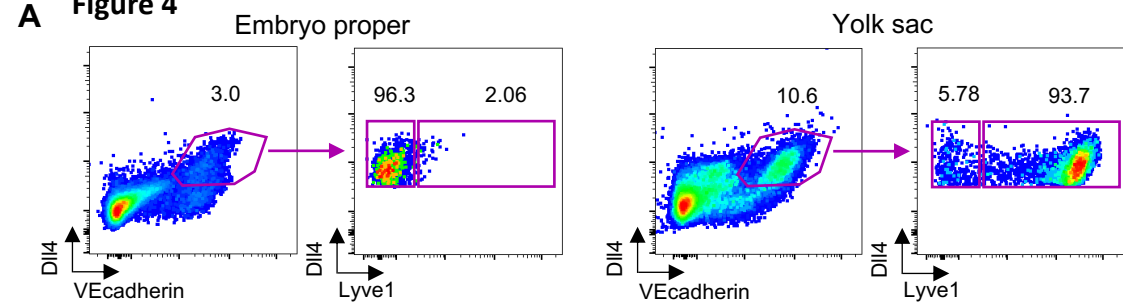


Figure 4

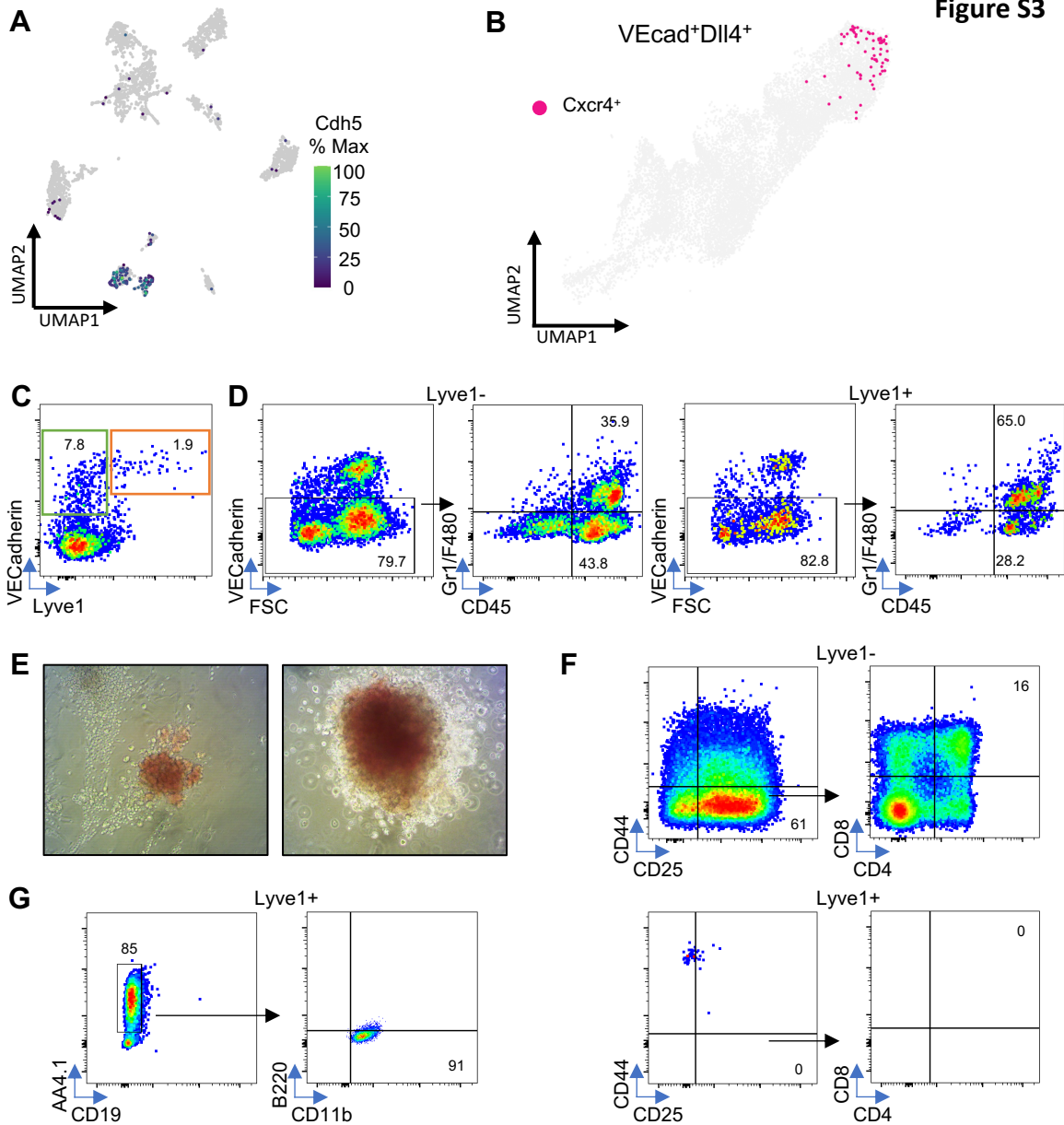
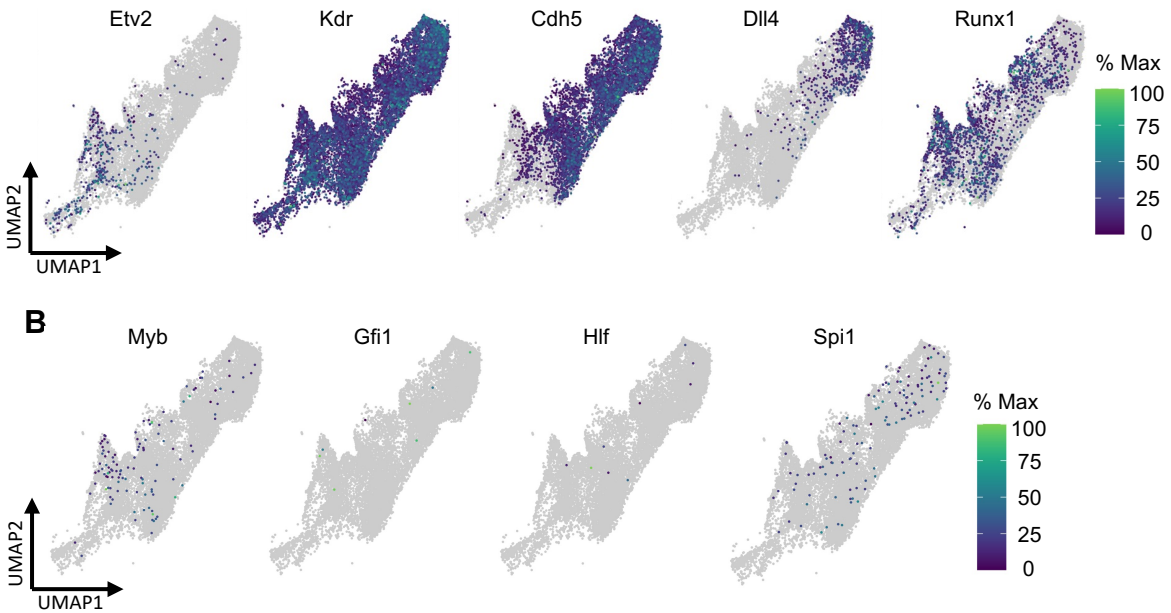
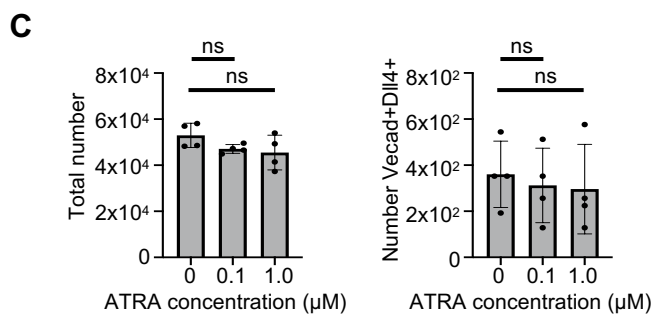
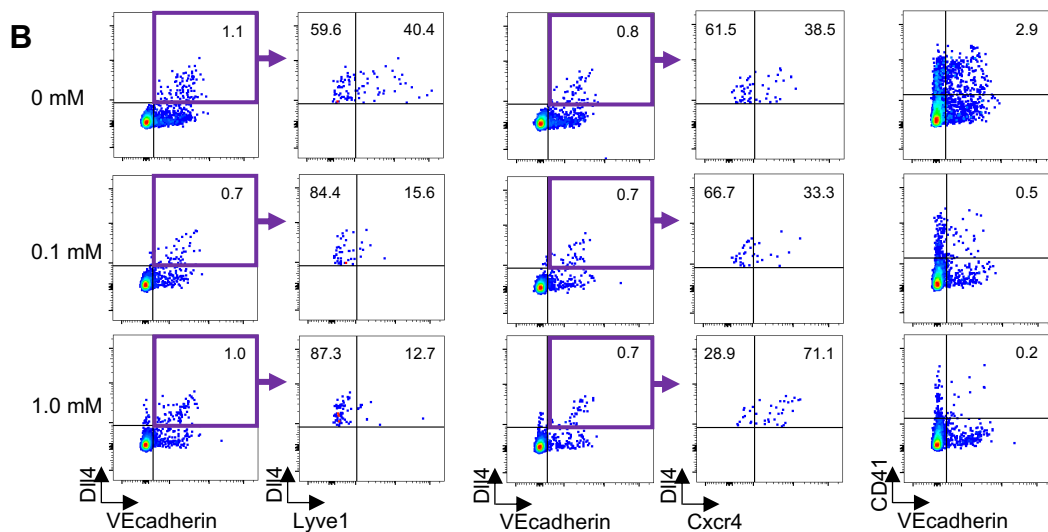
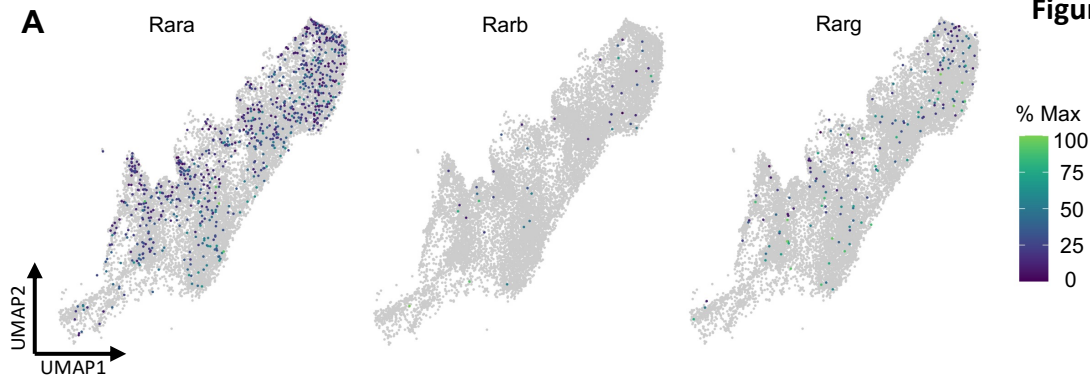


Figure S4



Supplemental Table 1: Cell type gene listsCurated from Pijuan-Sala et al. 2019, <https://www.nature.com/articles/s41586-019-0933-9>

<u>Epiblast</u>	<u>PGC</u>	<u>PS</u>	<u>Mesoderm</u>	<u>ExMeso</u>	<u>Alla</u>	<u>Mesenchyme</u>
Dnmt3b	Ifitm3	Nanog	Mesp1	Hoxa10		Ahnak
Pou5f1	Dnd1	Eomes	Lefty2	Hoxa11		Pmp22
Epcam	Dppa3	Hhex	Mesp2	Tbx4		Krt18
Utf1		Otx2	Osr1	Bmp4		Krt8
		Mixl1	Cdx2			Igf2
		Gsc	Hes7			
		Lhx1	Hoxb1			
		Fgf8	Cdx1			
		Pax7	Gbx2			
			Tbx1			
			Meox1			
			Tcf15			
			Aldh1a2			
			Dll1			
			Tbx6			
			Isl1			
			Tcf21			
<u>Cardiomyocytes</u>	<u>HE/Endothelial</u>	<u>Blood Progenitors</u>	<u>Erythroid</u>	<u>Neuroectoderm</u>		<u>Endoderm</u>
Smarcd3	Etv2	Lmo2	Hbb-bh1	Irx3		Dkk1
Acta2	Kdr	Runx1	Hba-a2	Hesx1		Krt19
Tagln	Anxa5	Gata1	Hba-a1	Six3		Amot
Myl7	Pecam1		Gypa	Sox9		Spink1
Myl4	Cdh5			Pax3		Emb
Tnnt2				Tfap2a		Cystm1
				Foxd3		Apoe
				Dlx2		Apoa2
				Sox10		Ttr
				En1		
				Pax2		

Supplemental Table 2: Cell type gene lists

<u>Primitive Streak (PS)</u>	<u>Extra-Embryonic Cells</u>	<u>Intra-Embryonic Cells</u>	
Wnt8a	Runx1	Lefty2	
Trh	Hapln1	Lrig3	
Pim2	Fli1	Pcdh8	
	Ddah1	Zic3	
	Tal1	Lhx1	
	Sox7	Dll1	
	Gata2	Pcdh19	
	Lmo2	Sema6d	
	Col1a1	Jag1	
	Morc4	Zic2	
	Lgr5	Lfng	
	Lama4	Pmaip1	
	Dlk1	Dll3	
		Rimbp2	
<u>Lateral Plate Mesoderm</u>	<u>Arterial EC (Kalucka et al)</u>	<u>Arterial EC (Xu et al)</u>	
Mixl1	8430408G22Rik	Efnb2	Ednrb
Cxcr4	Clu	Bmx	Dusp26
Myl4	Crip1	Gja5	Ace
Asb4	Fbln2	Gja4	Car7
Dkk1	Gja4	Cd200	Pcp4l1
Lefty2	Hey1	Cxcl12	Edn1
	Mecom	Kitl	Sema3g
	Sat1	Gkn3	Efhd1
	Sema3g	Fbln2	Jam2
	Sox17	Stmn2	Cd34
	Tm4sf1	Gper1	Ramp3
	Tsc22d1	Tmem255b	Plat
		C1qtnf9	Ltbp4
		Timp4	Ackr3
		Tmem100	Col18a1
		Sparcl1	Serpinf1
		Slc26a10	Acta2
		Rbp7	Podx1
		Alpl	Gprc5a
		Apq7	Lmcd1
		Hey1	Vegfc
		Emp2	Nxpe4
		Lims2	Igfbp3
		Cav1	

Methods

Mice

Wild-type C57Bl6/J7 (CD45.2) and congenic C57BL/6.SJL-Ly5.1-Pep3b (CD45.1) mice were bred at the Fred Hutchinson Cancer Research Center. Male and female C57Bl6/J7 CD45.2 mice at 8–12 weeks of age were used for transplantation experiments. All animal studies were conducted in accordance with the NIH guidelines for humane treatment of animals and were approved by the Institutional Animal Care and Use Committee at the Fred Hutchinson Cancer Research Center.

Mouse embryonic stem cells (mESCs)

mESCs from frozen aliquots of the same derivation were used for all differentiation experiments at less than passage 10. Briefly, cells were thawed fresh for each differentiation experiment and cultured in gelatinized, tissue culture 6-well plates or 60 mm dishes in ESGRO Complete PLUS Clonal Grad Medium (MilliporeSigma). Cells were passaged every other day using TrypLE to dissociate into a single cell suspension for 2-4 passages before being seeded for embryoid body differentiation.

AGM-derived Akt-ECs (AGM-ECs)

AGM-EC were generated as previously described (Hadland et al., 2015) and further detailed in a protocol available at Nature Protocol Exchange (<https://protocolexchange.researchsquare.com/>). For reproducibility, AGM-EC from

frozen aliquots of the same derivation were used for all co-culture experiments at less than passage 15 for experiments in this manuscript. AGM-EC were plated at a density of 5×10^5 cells/flask in gelatinized, tissue culture T75 flasks in EC media (consisting of IMDM with 20% FBS, Penicillin/streptomycin, Heparin, L-glutamine, and EC mitogen), and passaged every 3–4 days when confluent.

OP9/OP9-DLL4 stromal cells (OP9 cells)

GFP-expressing OP9 bone marrow stromal cells and GFP-expressing OP9 bone marrow stromal cells ectopically expressing Delta-like 4 (OP9-DLL4 cells) generated by the Zúñiga-Pflücker lab (Schmitt and Zúñiga-Pflücker, 2006) were used to assay the B- and T-lymphoid potential of hematopoietic colony cells. OP9 cell lines were plated at a density of 1.6×10^5 cells/flask in gelatinized, tissue culture T25 flasks in OP9 media (consisting of aMEM with 20% FBS, and Penicillin/streptomycin). Cells were passaged every 3 days (or when confluent) and used at passage numbers below 13 for co-culture assays.

Embryoid body differentiation and cell sorting

mESCs were collected and dissociated into single cell suspension by incubation in Accutase. Cells were washed in serum-free media, resuspended in StemPro-34 serum-free media and plated onto 100 mm petri dishes at a density of 2×10^6 cells per plate to allow for EB formation. EBs were cultured in StemPro-34 for 51 hours, EBs were then collected into 15 mL conical tubes (no more than 2x 100 mm dishes per tube) and

allowed to gravity settle for 10 minutes at 37°C. Once settled, media was gently removed, 1 mL TrypLE per dish was added, and cells were resuspended by gently flicking the tube 2-3 times. Cells were incubated in 37°C water bath for 3 minutes with gentle swirling and 1-2 flicks halfway through the incubation to keep the EBs suspended. An equivalent volume of StemPro-34 was added and cells were gently triterated 10x times with a 1000 uL pipette to fully dissociate EBs. Cells were then pooled, if applicable, washed in serum-free media, and resuspended in 300 uL of StemPro-34. 30 uL of DNase was added and cells were incubated for 10 minutes at room temperature. Following DNase treatment, cells were washed in serum-free median and resuspended in StemPro-34 containing recombinant vascular endothelial growth factor (VEGF) at 10 ng/mL and a combinatorial gradient of Activin and bone morphogenetic protein 4 (BMP4). It was determined a combination of 20 ng/mL Activin and 5 ng/mL BMP4 yielded the most robust downstream hematopoietic output, and was used in subsequent experiments unless otherwise noted. Cells were plated onto 24-well low-adherence plates at 4×10^5 cells per well to allow for EBs reformation. Following 72 hours of incubation (Day 5 of the differentiation protocol), the Activin concentration in the media was lowered. To do this, EBs were collected, gravity settled, and media was replaced with StemPro-34 containing 10 ng/mL VEGF, 5 ng/mL Activin and BMP4. When noted in certain experiments, all-trans retinoic acid (ATRA) was added at this point as well. Following media swap, EBs were gently replated onto new 24-well low-adherence plates.

After another 24-48 hours of culture (Day 6 or 7 of differentiation), dependent on the experiment, cells were sorted to isolate hematopoietic precursors and/or endothelium.

Briefly, EBs were collected and dissociated to single-cell suspension as described above. Cells were then washed with PBS containing 10% FBS. Cells were incubated with anti-mouse CD16/CD32 (FcγRII block) and stained with the following monoclonal antibodies as indicated. Relevant isotype control antibodies were used to set gates. DAPI staining was used to gate out dead cells. All reagents for cell staining were diluted in PBS with FBS, and staining was carried out on ice or at 4°C. Cells were sorted on a BD FACS Aria II. Sorted cells were subjected to downstream co-culture, lineage or transplantation assays as described below.

As noted for Sci-Plex experiments, EBs were collected and dissociated to single-cell suspension as described above over the course of differentiation (day 4, 5, and 6) to assess cell surface phenotype by flow cytometry or transcriptional expression by Sci-plex. Dissociated cells were processed for Sci-plex library prep and sequencing as previously described (Srivatsan et al., 2020).

Flow cytometry analysis

At various timepoints during differentiation or following co-culture, a fraction of the cell population (approximately 10-20% unless otherwise stated) was harvested for analysis of surface phenotype by flow cytometry. For differentiation timepoints, EBs were dissociated by incubation in TrypLE. Following co-culture, hematopoietic progeny were harvested from the EC layer by vigorous pipetting. Following collection, cells were spun and re-suspended in PBS with 2% FBS, pre-incubated with anti-mouse CD16/CD32 (FcRII block) and then stained with anti-mouse monoclonal antibodies as indicated.

DAPI was used to exclude dead cells. Flow cytometry was performed on a Becton Dickinson Canto 2 and data analyzed using FlowJo Software.

AGM-EC co-culture

Co-culture experiments were carried out with AGM-EC at passage 15 or less plated 24-48 hours prior to initiation of co-culture onto gelatin-treated 96-well or 24-well tissue culture plates at a density of 1×10^4 or 5×10^4 cells per well, respectively. Sorted mESCs were re-suspended in serum-free X-vivo 20 culture media with recombinant cytokines: stem cell factor (SCF) and FMS-like tyrosine kinase 3 ligand (FLT3L) each at 100 ng/ml, and interleukin-3 (IL-3) and thrombopoietin (TPO) each at 20 ng/ml. Following 5 to 7 days of co-culture hematopoietic progeny were harvested by vigorous pipetting for subsequent analysis by flow cytometry and/or lineage assays, as described below.

Colony forming unit (CFU) assays

In preparation for CFU assays, methylcellulose-based medium with recombinant cytokines was plated to 96-well non-tissue culture plates. Hematopoietic progeny were harvested by vigorous pipetting from the EC layer, washed in serum-free media, and resuspended in serum-free media. 10% of the cells from each well/replicate were added to individual 96-wells containing methylcellulose/cytokines (the remaining 90% was used for OP9 co-culture as described below). Visual characterization of resultant CFU colonies—CFU-Erythroid (CFU-E), CFU-granulocyte/macrophage (CFU-GM), or

multilineage colonies (CFU-GEMM) was performed under a light microscope after 10–12 days of culture.

OP9 co-culture assays

OP9 and OP9 DLL4-expressing (OP9-DLL4) cells were plated 24–48 hours prior to OP9 co-culture initiation at a density of 1×10^4 cells/well to gelatin-treated 24-well tissue culture plates in OP9 media. On the day of co-culture, 1mL/well of OP9 co-culture media containing FLT3L and interleukin-7 (IL-7) at 5 ng/ml each was added to OP9 24-wells in place of the OP9 media used for plating. The remaining 90% of hematopoietic co-culture cells processed as described above were divided equally and plated to OP9 or OP9-DLL4 coated wells.

After 10–12 days of co-culture, 50% of cells from the OP9 wells were harvested for flow cytometry analysis. Cells were stained in PBS with 2% FBS containing anti-mouse CD16/32 (FcRII block) and monoclonal antibodies AA4.1 (PE), B220 (Per-CP), CD19 (APC), and CD11b (APC-eFluor780). DAPI was used to exclude dead cells and contaminating GFP-expressing OP9 cells were excluded by gating on the FITC-negative population. Flow cytometry was performed on a Becton Dickinson Canto 2 and data analyzed using FlowJo Software. Isotype control-stained cells were used to set gates for analysis. B-lymphoid cells were identified phenotypically as CD11b-AA4.1+CD19+B220+.

Every 5–8 days of co-culture (or when cells began to overgrow), 50% of cells from OP9-DLL4 co-culture were transferred to newly plated OP9-DLL4 wells. After 12–21 days

total, 50% of cells from OP9-DLL4 wells were harvested for flow cytometry analysis. Cells were stained in PBS with 2% FBS containing anti-mouse CD16/32 (FcRII block) and monoclonal antibodies CD25 (PE-Cy7), CD44 (APC), CD4 (PerCP), and CD8 (PE). DAPI was used to exclude dead cells, and contaminating GFP-expressing OP9 stromal cells were excluded by gating on the FITC-negative population. Flow cytometry was performed on a Becton Dickinson Canto 2 and data analyzed using FlowJo Software. Isotype control-stained cells were used to set gates for analysis. T cells were identified phenotypically as CD25⁺CD44⁻CD4⁺CD8⁺.

Transplantation assays

Following co-culture, a fraction of the hematopoietic cells generated were harvested from the EC layer by vigorous pipetting, washed with 2% FBS and re-suspended in 100 uL PBS/2% FBS per mouse transplanted. Samples were supplemented with 5×10^4 whole marrow cells from adult congenic C57BL/6.SJL-Ly5.1-Pep3b (CD45.1) mice in 100 ul PBS/2% FBS to provide hematopoietic rescue. Cell suspensions (200 ul total volume/mouse) were injected into lethally irradiated (1,000 cGy using a Cesium source) congenic CD45.1 adult recipients via the tail vein. Flow cytometry analysis of peripheral blood obtained by retro-orbital bleeds was performed at various intervals between 2 and 24 weeks following transplantation. Lineage-specific staining for donor (CD45.2) and recipient/rescue (CD45.1) cells from peripheral blood was performed as previously described (Hadland et al., 2015), using anti-mouse monoclonal antibodies indicated: CD45.1, CD45.2, CD3, CD19, Gr1, and F4/80.

Embryo dissections for Sci-Plex

Embryos were harvested from pregnant females, as previously described (Hadland et al., 2015). Embryo age was determined by morphology (E6.5-7) or by counting somite pairs (E8.5-9). For E6.5-7 and E8.5-9, whole embryo/extra-embryonic tissue, or yolk sac and P-Sp region of the embryo proper were collected, respectively. Dissected tissues were treated with 0.25% collagenase for 30 minutes at 37°C, pipetted to single cell suspension, and washed with PBS containing 10% FBS. The resulting single cell suspension was process for Sci-plex library prep and sequencing as previously described (Srivatsan et al., 2020).

Sci-Plex data and analysis

All raw and processed sequencing data is being uploaded to GEO and relevant accession numbers will be made publicly available. All scripts for single cell transcriptomic analysis will be uploaded and publicly available on the FHCC GitHub.

Dimensionality reduction, batch correction, and cluster analysis

The preprocess_cds() function was used to project the data onto the top principal components. EB and Embryo datasets were analyzed separately with the num_dim set to 98 and 50, respectively. Uniform Manifold Approximation (UMAP) was used for dimensionality reduction with the reduce_dimension() function, with umap.n_neighbors

= 4L. `set.seed` was used to specify the number of seeds to avoid variability in output due to a random number generator in the function. Clustering was performed with the `cluster_cells` function with a resolution of $3.6e-4$ and $3e-4$ for the EB and Embryo datasets, respectively.

Cell type classification

Cell type classification was performed using diagnostic marker genes curated from (insert Pujin-Sala et al citation) (marker genes listed in Supplementary Table 1). Cell clusters were classified as a given cell type based aggregate expression of diagnostic markers (see heatmap in Figure S4A). Using this method, a small population of cells remains unclassified rather than inputting cell types for all cells in the scRNA-seq dataset. Cell type classifications were used for all downstream analyses, as indicated.

Gene-set scores

Gene set scores were calculated for each single cell as the log-transformed sum of the size factor-normalized expression for each gene in published signature gene sets as follows; primitive streak (Pijuan-Sala et al 2019), extra-embryonic (EEC) and intra-embryonic (IEC) (Saykali et al Elife '19), lateral plate mesoderm (LPM) (Zhao and Choi, 2019 Development, Prummel et al 2019), and arterial endothelial cells (Kalucka et al, Xu et al). A list of the genes and associated references can be found in Table 2.

Sub-setting and cell counting

Cells were subset or groups of cells were highlighted based on cell classification and/or specific gene expression as labeled. Cell counts for a given group were generated from cell dataset metadata for Activin/BMP4 concentration and/or collection timepoint. All heatmaps and cell dot plots were generated in the ggplot2 package.

Quantification and statistical analysis

Statistical analyses of colony frequencies were conducted with Prism. Data are expressed as mean \pm standard deviation (s.d.), and n indicates biological repeats, unless indicated otherwise. P values calculated by Student's t test unless otherwise indicated.