A multistep computational approach reveals a neuro-mesenchymal cell population in the embryonic hematopoietic stem cell niche

Olivera Miladinovic[1*#], Pierre-Yves Canto[1*], Claire Pouget[2], Olivier Piau[1, 3], Nevenka Radic[1], Priscilla Freschu[1], Alexandre Megherbi[1], Carla Brujas Prats[1], Sebastien Jacques[4], Estelle Hirsinger[1], Audrey Geeverding[5], Sylvie Dufour[6], Laurence Petit[1], Michele Souyri[7], Trista North[8], Hervé Isambert[9], David Traver[2], Thierry Jaffredo[1**], Pierre Charbord[1**] and Charles Durand[1**]

[1] Sorbonne Université, CNRS, Inserm U1156, Institut de Biologie Paris Seine, Laboratoire de Biologie du Développement/UMR7622, 9 Quai St-Bernard, 75005 Paris, France.

[2] Department of Cell and Developmental Biology, University of California San Diego, La Jolla, CA 92093-0380, USA.

[3] Sorbonne Université, Inserm, UMR-S 938, Centre de Recherche Saint-Antoine-Team Proliferation and Differentiation of Stem Cells, Institut Universitaire de Cancérologie, F-75012 Paris, France.

[4] Plateforme de génomique, Université de Paris, Institut Cochin, Inserm, CNRS, F-75014 Paris, France.

[5] Service de microscopie électronique, Fr3631, Institut de Biologie Paris Seine, Sorbonne Université, CNRS.

[6] Univ Paris Est Créteil, Inserm, IMRB, F94010 Créteil, France.

[7] Université de Paris, Inserm UMR 1131, Institut de Recherche Saint Louis, Hôpital Saint Louis, 1 Avenue Claude Vellefaux, 75010 Paris, France.

[8] Stem Cell Program, Department of Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA; Developmental and Regenerative Biology Program, Harvard Medical School, Boston, MA 02115, USA.

[9] Institut Curie, PSL Research University, CNRS UMR168, Paris, France.

* Co-authors
** Co-senior authors
# Present address: Stem Cell Program, Department of Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA; Developmental and Regenerative Biology Program, Harvard Medical School, Boston, MA 02115, USA.

Correspondence to Charles Durand: charles.durand@sorbonne-universite.fr
https://orcid.org/0000-0002-1970-0467

Running title: AGM HSPC niche molecular identity

Keywords: Mesenchyme, hematopoietic stem cells, niche, Aorta-Gonad-Mesonephros, laser microdissection, gene networks.

Summary statement

Identification of a unique neuro-mesenchymal cell population in the embryonic hematopoietic stem cell niche through multi-layered transcriptomics and computational analyses.

Abstract

The first hematopoietic stem and progenitor cells (HSPCs) emerge in the Aorta-Gonad-Mesonephros (AGM) region of mid-gestation mouse embryo. However, the precise nature of their supportive mesenchymal microenvironment remains largely unexplored. Here, we profiled transcriptomes of laser micro-dissected aortic tissues at three developmental stages and individual AGM cells. Computational analyses allowed identifying several cell subpopulations within the embryonic day 11.5 AGM mesenchyme, with the remarkable presence of a yet unidentified subpopulation characterized by the dual expression of genes implicated in adhesive or neuronal functions. We confirmed the identity of this cell subset as a neuro-mesenchymal population, through morphological and lineage tracing assays. Loss of function in the zebrafish confirmed that Decorin, a characteristic extracellular matrix component of the neuro-mesenchyme, is essential for HSPC development. We further demonstrated that this cell population is not merely derived from the neural crest, and hence, is a bona fide novel subpopulation of the AGM mesenchyme.

Introduction

The first adult-type hematopoietic stem cells (HSCs), defined by their capability to repopulate the hematopoietic system of adult irradiated recipients, are generated in the Aorta-Gonad-Mesonephros (AGM) region of mid-gestation mouse embryo (Medvinsky and Dzierzak, 1996, Müller et al., 1994). Intra-aortic hematopoietic clusters (IAHCs), containing hematopoietic stem and progenitor cells (HSPCs) closely attached to the ventral side of the dorsal aorta (Tavian et al., 1996, Yokomizo and Dzierzak, 2010, Jaffredo et al., 1998), originate from hemogenic endothelial cells through a process of endothelial-to-hematopoietic transition (EHT) (Jaffredo et al., 1998, Oberlin et al., 2002, de Bruijn et al., 2002, Chen et al., 2009, Zovein et al., 2008, Bertrand et al., 2010, Boisset et al., 2010, Kissa and Herbomel, 2010). In the mouse, IAHCs are also found on the dorsal side of the aorta (Taoudi and Medvinsky, 2007). However, HSCs are preferentially located in the ventral aortic domain (Taoudi and Medvinsky, 2007, Souilhol et al., 2016). Moreover, experiments in the chick embryo revealed that the subaortic tissue originated from splanchnic mesoderm and is essential for IAHC production (Richard et al., 2013). Collectively, these data indicate that the aortic hematopoietic activity is polarized along the dorsoventral axis and that the ventral domain constitutes a supportive HSPC microenvironment. Recent studies have provided novel insight on the AGM microenvironment in terms of phenotype, gene expression and lineage history (Chandrakanthan et al., 2022, Kapeni et al., 2022, Sa da Bandeira et al., 2022, Fadlullah et al., 2022, Crosse et al., 2020, Yvernogeau et al., 2020). At the molecular level, extrinsic factors including cytokines (Robin et al., 2006, Souilhol et al., 2016, Petit-Cocault et al., 2007), signaling pathways (Marshall et al., 2000, Durand et al., 2007, Peeters et al., 2009, Wilkinson et al., 2009, Robert-Moreno et al., 2005, Pouget et al., 2014, Lee et al., 2014, Souilhol et al., 2016, McGarvey et al., 2017), shear stress mediators, proinflammatory and sympathetic nervous system mediators (Adamo et al., 2009, North et al., 2009, Fitch et al., 2012, Espín-Palazón et al., 2014, Li et al., 2014, Mariani et al., 2019, Sawamiphak et al., 2014) participate in HSPC development.

The mesenchyme is primarily responsible for extracellular matrix (ECM) production, organization and remodeling. The AGM mesenchyme remains less well-defined than the hematopoietic component forming the IAHCs or the endothelial lining at the origin of the

HSPCs through EHT (Calvanese et al., 2022, Fadlullah et al., 2022, McKinney-Freeman et al., 2012, Solaimani Kartalaei et al., 2015, Baron et al., 2018). Capturing the molecular signature of the AGM mesenchyme in vivo indeed constitutes a technical and biological challenge since an array of diverse cell types (endothelial cells forming the vascular system, disseminated macrophages and neural crest cells forming the para-aortic lateral primary sympathetic ganglia and the more ventrally located adrenal primordia) are enmeshed within that ECM-embedded mesenchyme. We hypothesize that discrete mesenchymal cell subtypes may coexist in the AGM, depending on their topological association with dorsal (e.g., notochord and somites) and ventral (e.g., gut and gonads) developing tissues.

To investigate this critical issue, and in line with our previous reports aiming at extracting the molecular core of HSPC niches (Desterke et al., 2020, Charbord et al., 2014), we present here an integrated strategy combining precise laser capture, bulk and single cell transcriptomics with computational analyses including correlation- and information-based gene network analyses. Our analyses indicated the presence of a ventrally located neuro-mesenchymal subset at embryonic day (E) 11.5 within a large mesenchymal cell population. Morphological and lineage tracing studies confirmed the existence of this cell subset as bona fide yet-unexplored population and demonstrated that it was not derived from the neural crest. Moreover, loss-of-function experiments in the zebrafish validated the essential role of the characteristic ECM component Decorin for HSPC development. Altogether, this study provides an unexpected insight onto the transcriptomic landscape of the AGM mesenchyme.

Results

The dorsal and ventral aortic tissues exhibit distinct molecular signatures

We used a laser microdissection approach to isolate the ventral (VT) and dorsal (DT) aortic tissues from E10.5, E11.5 and E12.5 mouse embryos (Fig. 1A, area delineated in red, Fig. S1A, Table S1 and Video 1) and to compare their gene expression profiles.

The lists of differentially expressed genes (DEGs) between VT and DT were extracted by ANOVA and DEGs were characterized by different gene ontology (GO) profiles. Some of the VT profiles ('ECM', 'metalloproteinase activity', 'cell junction', 'synapse', 'Wnt signaling', innate immunity) were of most interest since they peaked at E11.5 (Fig. 1B), a developmental stage after which AGM HSC activity decreases (Kumaravelu et al., 2002). This profile was confirmed using Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005) (Figs. 1C, 1D). The gene sets 'Complement', 'IL6-Jak-Stat3 signaling' and 'Interferon alpha response' were significantly enriched in VT, in agreement with recent studies demonstrating the role of proinflammatory signals on AGM HSCs (Espín-Palazón et al., 2014, Sawamiphak et al., 2014, Li et al., 2014, Mariani et al., 2019). Importantly, the 'Epithelial-to-Mesenchymal Transition' gene set was also significantly enriched in VT, including numerous highly ranked genes encoding ECM proteins, such as the interstitial Collagen I (*Col1a1* and *Col1a2),* Decorin (*Dcn*), a member of the small leucine-rich proteoglycan family with essential role in collagen fibril assembly, Tenascin C (*Tnc),* an interstitial fibronectin-binding molecule implicated in neuronal guidance, Lysyl oxidase homolog 1 *(Loxl1)*, an oxidoreductase essential for collagen crosslinking, and Transforming growth factor-beta-induced protein ig-h3 (Tgfbi), a molecule implicated in cell-collagen interactions. Taken together, these analyses revealed a strong enrichment for ECM-associated molecules in the subaortic tissue at E11.5.

Correlation-based network analysis of mesenchyme-enriched genes discloses three major modules, one ventrally and two dorsally located

Our further study focused on E11.5 since the expression profiles were most contrasted at this time point.

To enrich our database in mesenchyme-specific genes, we merged our E11.5 transcriptome with two other expression sets. The first expression set corresponded to flow cytometry sorted AGM cells, endothelial cells ($CD144^{pos}CD45^{neg}$), differentiated hematopoietic cells ($CD34^{low}Kit^{low}CD45^{pos}$), and HSPCs ($CD34^{high}Kit^{high}$). Gating strategies and molecular identities of these cell fractions are presented in Fig. S2. The second expression set corresponded to AGM stromal lines derived from the aorta region (AM14-1C4, AM30-3F4) or the urogenital ridges (UG26-1B6 and UG26-3B5) (Oostendorp et al.,

2002, Charbord et al., 2014). Using ANOVA and then intersecting the resulting DT and VT DEGs led to a list of 2401 genes, 1199 upregulated in DT and 1202 in VT (Table S2). Using Principal Component Analysis (PCA) (Fig. 2A) allowed to discriminate DT from VT according to PC1, and endothelial cells (EC) from differentiated hematopoietic cells (HC) according to PC2 (34,2 and 19,1% of the variance, respectively).

To find out which genes were specific to each of the four discriminated tissue/cell populations (VT, DT, EC, HC) and to reveal their relationships, Weighed Gene Correlation Network Analysis (WGCNA) combined with PCA was applied to the set including 2401 genes (Langfelder and Horvath, 2008). PC1 and PC2 corresponded to the external traits (DT-to-VT and EC-to-HC contrasts), and the merged dynamic tree cut method was used to identify the modules (Fig. 2B, Fig. S1B). The distribution of module eigenvectors indicated that two modules (designated blue and yellow) corresponded to VT, another two (brown and turquoise) to DT, one (cyan) to EC, and two (magenta and salmon) to HC (Fig. 2C).

Gene analysis was then performed on the 7 modules (see Table S2 for gene allocation). Firstly, it confirmed the endothelial and hematopoietic identities of the populations corresponding to the EC and HC modules, respectively. Secondly, it indicated that the DT and VT modules had very distinctive features. Most genes of the DT module designated ‘brown‘ were implicated in neurogenesis; many transcription factors essential for neuronal development and differentiation such as *Sox2, Pou3f3,* and *Npas3* were among the most connected genes. Most genes of the DT module designated ‘turquoise’ were implicated in chondrogenesis; the transcription factor *Pax1* expressed in pre-chondrocyte condensations, and the ECM cartilage collagen components *Col2a1* and *Col9a1* were among the most connected genes. The VT module designated ‘yellow’ was poorly defined since including many unexplored and predicted genes (olfactory receptors, Riken, Gm genes). In line with the pairwise DT to VT comparison (as shown on Fig. 1B), the VT module designated ‘blue’ was of most interest since including genes implicated in either adhesion-related processes ('cell adhesion', 'basement membrane', ‘ECM’, ‘collagen’, ‘homophilic cell adhesion’, ‘focal adhesion’) or neuronal-related ones ('synapse', 'ion channel', 'GABA-A receptor activity', ‘neuroactive ligand-receptor’, ‘neurotrophin signaling’) (Fig. 2D). Among the most connected genes were the ECM components *Dcn* and *Tnc,* the neurotrophin-3, glutamate and netrin receptors *Ntrk3, Grid1* and *Unc5c,* and

molecules implicated in synapse functioning and/or neuronal development such as *Clstn2, Amph, Syt1, Cntn4*, and *Ntf4* (Figs. 2E, 2F). Genes encoding for soluble factors such as *Bmp6*, *Gdf6*, *Csf1, Fgf14, Igfbp1* and *Igfbp5*, *Ntf5* and *Postn* also characterized the 'blue' module (Fig. 2E).

These data suggested the presence of three cell populations within the dorsal and ventral tissues, two dorsally-located (above the aorta ceiling), either expressing neuronal genes, or chondrocytic ones, and one ventrally-located (beneath the aorta floor) expressing both adhesion- and neuronal-related genes.

To find out how those cell populations would integrate within the AGM region we performed single cell transcriptomic analysis on the entire E11.5 AGM.

Single cell analysis (scRNAseq) discloses a large and heterogeneous mesenchymal cell population

The entire AGM was dissected as shown on Figure 1A (area delineated by a green line). After Quality Control (Fig. S3A) we retained 15154 out of 15862 cells, with 2698 genes detected on average. After normalization and dimension reduction, the Louvain algorithm allowed defining clusters projected on Uniform Manifold Approximation and Projection (UMAP) embedding (Fig. 3A). Thirteen of the 22 clusters were packed together forming a large cluster, clearly discriminated from the 9 others that presented as single entities.

The large cluster consisted in mesenchymal cells as indicated by the widespread expression of the interstitial collagens *Col1a1, Col1a2* and *Col3a1*, the lysyl oxidase homolog *Loxl1,* the microfibril component *Fbn2*, the ECM-degrading metalloproteases *Mmp16* and *Bmp1*, the proteoglycan *Vcan*, several molecules implicated in matrix assembly (*Ccdc80*), matrix organization (*Fras1*) and cell-collagen interactions (*Tgfbi*), the receptors for the growth factors of mesenchymal cells (*Pdgfra* and *Pdgfrb)*, the hematopoietic chemokine expressed by stromal cells *Cxcl12*, and finally molecules implicated in EMT, the transcription factor *Snai2* and the growth factor *Ptn* (Fig. 3B, Fig. S3B). Of note, we had found in a previous work that some of these genes (*Ccdc80, Snai2* and *Tgfbi*) were regulators of HSPC development in the AGM (Charbord et al., 2014).

To find out the cell identities of the different clusters aggregated in that large mesenchymal population we first examined the first-ranked genes. Among the 8 major clusters (C0 to C7), each containing more than 600 cells, 5 corresponded to well-defined identities: cells in C0 were chondrocytic, in C1 and C5 urogenital, in C2 cycling, and in C4 smooth muscle (the characteristic genes are indicated on the upper table of Fig. 3A, and the distribution of the first-ranked gene is shown on Fig. 3C). For the 3 remaining clusters (C3, C6 and C7) the first-ranked genes did not provide any clear clue (upper table of Fig. 3A and Fig. 3C).

Our second approach was therefore to investigate in which clusters the genes of the two DT (turquoise and brown) and the one VT (blue) modules indicated by WGCNA were expressed. Genes of the VT module, but of only one of the DT modules (the 'turquoise') projected onto clusters of the mesenchymal population. The projection of the genes of the 'turquoise' DT module onto the cluster map indicated that they were expressed in the C0 cluster, with a minor contribution in C2 (Fig. S3C); these results confirmed the chondrocytic identity of C0 cells. Similar projection of the genes of the VT module onto the cluster map indicated that they were expressed in the clusters C6 and C7 (Fig. 3D). Hence, these data suggested the existence of a specific mesenchymal cell population, that we called neuro-mesenchyme since it expressed both adhesive (such as *Dcn* and *Tnc*) and neuronal genes (such as *Ntrk3*), represented in the C6 and C7 clusters (71% of the genes expressed in C6 were also expressed in C7 at about the same intensity level, suggesting that cells in C6 and C7 clusters belonged to the same population).

The remaining large cluster C3 was then examined again by taking into account the distribution of genes also expressed in other clusters. It revealed that many genes expressed in C6/C7 (such as *Ntrk3, Tnc, Hlx, Col25a1)*, in C0 (such as *Col2a1, Col9a1, Foxf1*), or in C1 (such as *Gata2)* were also expressed in C0 (Fig. 3D, Figs. S3C, S3D), indicating that the C3 cluster was a mix of several cell populations.

Finally for the 5 minor mesenchymal clusters (containing less than 600 cells), there were few (C8-C11) or no (C12) significant genes; C8 expressed some genes implicated in ECM organization or synthesis, and C8 to C11 expressed transcription factors implicated at

different developmental stages (upper table of Fig. 3A). Many genes implicated in the cell cycle were intensely expressed in C12.

The first ranked genes of the non-mesenchymal clusters revealed without ambiguity their identity (lower table of Fig. 3A). There were two neuronal clusters C13 and C14, the latter including many genes characteristic of neural crest cells (NCC); interestingly, the genes of the DT 'brown' module found using WGCNA were expressed in the non-NCC C13 cluster. There were also one endothelial cell cluster (C15) and one containing innate immunity cells (C16), onto which the genes of the endothelial ('cyan') and hematopoietic ('magenta' and 'salmon') modules, respectively, were projected (Fig. S4). Finally, there was a very small cluster (C19) of germinal cells.

Taken together, the single cell analysis confirmed both the heterogeneity of the AGM region taken as a whole, and that of the AGM mesenchyme that included 83% of all AGM cells. The ventrally-located AGM mesenchyme comprised two major populations, one of prospectively 'neuro-mesenchymal' cells expressing both adhesion and neuronal genes (C6 and C7 clusters), and the other of urogenital cells (C1 and C5). The dorsally-located AGM mesenchyme comprised one major population of chondrocytic cells (C0). Dorsal vs ventral allocation was meaningless for C4 containing smooth muscle cells located around the aorta, and for C2 containing cycling cells that could be present among all mesenchymal cells. Finally, the location was less clear for C3 that contained a mix of cells that could be ventrally- or dorsally-located, and for the minor clusters C8 to C12 in which genes were too few to allow precise spatial allocation.

Gene set scoring and information-based network analysis indicate the presence of a ventrally located neuro-mesenchymal population

The previous analyses strongly suggested that clusters C6 and C7 represented a population of mesenchymal cells expressing neuronal genes. To further demonstrate the existence of such population, we extracted from the 'blue' VT module list two gene sets representative of the neural and mesenchymal lineage (Table S3). Gene set scoring using the Seurat AddModuleScore function indicated that 63.9% of the cells in C6 and C7 were positive for both gene subsets (Fig. 4A).

Contrary to the distribution of genes representative of the neural subset that was restricted to C6 and C7, that of the mesenchyme subset was more extended, in particular to cells belonging to C1 (Fig. 4A, third panel from the left). By searching for individual genes expressing such a profile we found that it corresponded to several ECM components, *Dcn*, and collagen family members, *Col4a5,* one of the major components of basement membranes, and *Col6a3* and *Col6a2,* two chains of the cell-binding protein Collagen VI (Fig. 4B).

To get more insight on the organization of genes characterizing the neuro-mesenchyme and to identify direct gene-to-gene interactions and potential causal links, we applied the multivariate information-based inductive causation (miic) algorithm (Verny et al., 2017, Sella et al., 2018) on the VT module genes. To yield significant results the miic algorithm requires a large number of observations as compared to that of variables. We restricted therefore the analysis to the central core of the module including 98 genes and we considered all cells (N = 12516) belonging to the large mesenchymal cluster. After filtering out low confidence links, 89 nodes defining direct high-confidence paths were retained (Fig. 4C). The resulting network was pruned as compared to that provided by WGCNA. It included four major hubs (with degree ≥ 17), *Unc5c, Col1a1,* the neural cell adhesion molecule *Ncam1,* and the transcription factor *Zmiz1,* and a few less prominent ones, such as the Dopa decarboxylase *Ddc*. Of interest, *Ddc* and all its 8 links (including *Tnc, Kcnip4,* a channel interacting protein*, Nkain2,* an ion transport ATPase and *Adam23,* a disintegrin without catalytic activity), were expressed only in C7, indicating that their mutual interaction links corresponded to a specific gene distribution given by the single cell analysis (Fig. 4C, right panel). Importantly, all of the ECM components were directly linked to neuronal molecules (e.g., *Col1a1-Ntrk3-Tnc, Dcn-Ncam1-Tnc, Ntrk3-Col8a1-Ddc, Ntrk3-Tnc-Ddc, Ncam1-Tnc-Ddc),* and nine edges were confirmed by the analysis of protein-protein interactions using STRING (Szklarczyk et al., 2019) (Fig. 4D). Finally, the miic algorithm uncovered a number of 'v-structures' (e.g., the nuclear receptor *Nr1h5 --> Tnc <-- Ncam1*), suggesting cause-and-effect interactions to be investigated in future works.

Finally, to investigate whether this cell population was present only during E11.5 HSPC production, we used the WGCNA ‘blue’ module as a gene set of reference to interrogate

the global transcriptomes of ventral micro-dissected tissues at E10.5 and E12.5, as well as the transcriptomes of aortic mesenchyme and urogenital ridges stromal lines. GSEA indicate that the neuro-mesenchymal signature was enriched in E11.5 ventral tissues as compared to E10.5 and E12.5 ventral tissues, suggesting that the neuro-mesenchyme is mostly present during HSPC production in the AGM (Fig. S5, left and middle panels). In addition, we observed that the 'blue' module is statistically enriched in aorta mesenchymal as compared to urogenital ridges stromal lines, confirming the aortic mesenchymal origin of the micro-dissected tissues (Fig. S5, right panel).

Taken together the data pointed to the existence of a, yet undescribed, E11.5 neuro-mesenchymal population, in which the majority of individual cells express both neuronal and mesenchymal genes that appear to be directly linked one another in a network with a few highly connected hubs.

Morphological studies confirm the existence of a subaortic neuro-mesenchymal cell population

To validate the in-silico data we analyzed the distribution of three characteristic proteins on AGM sections. We selected three candidates, Tenascin C, Decorin and the neurotrophin-3 receptor NTRK3, which we considered as characteristic of the neuro-mesenchyme since they preeminently appeared as highly connected genes in WGCNA, as expressed in the C6 and C7 clusters in scRNAseq, and as nodes of ECM-to-neuronal links in miic. In addition, *TNC*, *DCN* and *NTRK3* expression profiles were mapped exclusively onto the 'stroma' clusters of the human week 4 AGM single cell transcriptome atlas (Calvanese et al., 2022); interestingly, *DCN* was expressed in the three stroma clusters, while *TNC* and *NTRK3* expression was confined to cells of the larger cluster, many of which expressed both genes (Fig. S6).

Decorin expression was detected by immunofluorescence around the notochord and in the subaortic tissue at E11.5, but not in the aortic endothelial and vascular smooth muscle cells, or macrophages, as shown by the absence of overlapping staining between Decorin and CD31, α-smooth muscle actin (α-SMA) or Allograft inflammatory factor 1 (Iba1) (Figs. 5A, 5C, Fig. S7A). Tenascin C was also highly expressed in cells and associated ECM in large

areas of the subaortic tissue (Fig. 5A and Video 2). Importantly, in agreement with previous report (Marshall et al., 1999), similar distribution of Decorin and Tenascin C were observed in human embryos at day 30 of gestation (Fig. 5B). NTRK3 was present, as expected in the dorsal root ganglion cells, and in cells in extensive areas around and underneath the dorsal aorta (Fig. 5C and Video 3). Importantly, the distribution of Decorin, Tenascin C and NTRK3 staining was quite similar in large areas of the subaortic tissue. The expression of these proteins was clearly distinct from that of cells expressing Tyrosine hydroxylase (Th), an enzyme involved in catecholamine biosynthesis, and of axons expressing Tuj1, a neuron-specific class III beta-tubulin expressed in neurons of the peripheral and central nervous system (Fig. 5C and Fig. S6B-D); immunostaining for Th revealed discrete foci in the dorsal and ventral domains of the aorta, corresponding to primary sympathetic ganglia and adrenal primordia (Fig. 5C); immunostaining for Tuj1 revealed a dense array of axons on the lateral sides, and very few foci underneath the aorta (Fig. 5C and Fig. S7B-D). Moreover, sagittal analyses of E11.5 3D embryo imaging revealed a strong expression of Decorin and Ntrk3 in the middle part of the AGM known to be enriched for HSC activity (Fig. S8) (Easterbrook et al., 2019, Mascarenhas et al., 2009). Finally, transmission electron microscopy revealed the presence of transverse and longitudinal ECM fibers in the intercellular space between subaortic cells at E11.5 (Fig. 5D).

These data confirmed the presence of a subaortic mesenchyme in which the ECM components Decorin and Tenascin C had distribution similar to that of the cells expressing NTRK3. Moreover, these cells were clearly distinct from neurons expressing Tubulin ß3 or Tyrosine hydroxylase. These results therefore confirmed the presence of a yet unidentified neuro-mesenchymal population.

The E11.5 subaortic neuro-mesenchymal subpopulation does not derive from neural crest cells

An obvious issue was whether this neuro-mesenchymal cell population was derived from NCCs. The computational analysis was not in favor of such assumption since genes characteristic of NCCs were expressed by cells belonging to the non-mesenchymal C13 and C14 clusters. Indeed, the major NCC marker *Sox10*, the early (*Phox2b*) and late

markers (*Phox2a* and *Th)* of the sympathoadrenal compartment were expressed in C14 and C13 (Fig. 6A). To definitely exclude this hypothesis we took advantage of a Human tissue Plasminogen Activator (HtPA)-Cre mice labelling neural crest cells and their derivatives *in vivo* (Pietri et al., 2003). After crossing those mice with GT Rosa Tomato mice, E11.5 embryos were collected and analyzed for the expression of Tomato, Decorin, Tenascin C and NTRK3 (Fig. 6B-6D). Tomato$^+$ cells were detected as expected on the dorsal, lateral and ventral sides of the aorta (Fig. 6B-6D). There was no overlap between Tomato expression and that of Decorin, Tenascin C or NTRK3 (Fig. 6B-6D). These data allowed therefore excluding the hypothesis that the neuro-mesenchymal population was NCC-derived.

Loss-of-function experiments in the zebrafish reveal the essential role for HSPC development of the characteristic ECM component Decorin

To functionally validate our computational analysis, we performed in vivo knockdown experiments in zebrafish embryos using antisense morpholinos (MOs). The zebrafish is indeed an excellent system to test candidate gene function in a robust and rapid manner (McKinney-Freeman et al., 2012, Charbord et al., 2014, Yvernogeau et al., 2020, Fadlullah et al., 2022). Dcn was selected as a potential HSPC regulator since it is an extracellular matrix component expressed in the subaortic mesenchyme both in murine and human embryos, and, to our knowledge, not reported yet as playing a role in developmental hematopoiesis.

We first investigated *dcn* expression in zebrafish embryos. In agreement with a previous study (Zoeller et al., 2009), we detected *dcn* expression in the developing head and trunk (particularly the somites) by whole mount in situ hybridization (WISH) (Fig. S9A). In line with these observations, it has been recently reported that *dcn* was preferentially expressed in somite-derived endothelial cells acting as a vascular niche for hemogenic endothelial cells (Sahai-Hernandez et al., 2023). To refine our analysis, we evaluated by RT-qPCR *dcn* expression on sorted cell populations: *flk1$^-$cmyb$^-$ (*non-hematopoietic non-endothelial mesenchymal cells), *flk1$^-$cmyb$^+$* (myeloid progenitors), *flk1$^+$cmyb$^-$*, (vascular endothelial cells) and *flk1$^+$cmyb$^+$* (HSPCs) from the trunk of 48 hours post-fertilization (hpf) embryos using the *flk1:mcherry;cmyb:gfp* zebrafish line (Fig. S9B). Our results

revealed that *dcn* was significantly more expressed in *flk1*$^{-}$*cmyb*$^{-}$ and *flk1*$^{+}$*cmyb*$^{-}$ cells in zebrafish embryo (Fig. S9C).

The formation of HSPCs in the dorsal aorta was monitored by WISH for the expression of the early (*runx1*) and late (*cmyb*) HSPC markers (Figs. 7A, 7B). The morphants *dcn* injected with MO^{START} exhibited strong hematopoietic defects in the dorsal aorta, as shown by significant reduced expression of *runx1* and *cmyb* in morphants as compared to respective controls at 30 hpf and 36 hpf, respectively (Figs. 7A, 7B). Similar phenotypes were observed with the injection of a MO^{SPLICE} (Figs. S10A, S10B).

To confirm the knockdown experiments and to exclude any 'off-the-target' effect of morpholinos we examined zebrafish mutants. Since adult *dcn*$^{-/-}$ zebrafish (*dcn*^{sa20247} line), harboring a point mutation that induces a stop codon in the *dcn* coding sequence were viable and fertile, we crossed *dcn*$^{-/-}$ females with *dcn*$^{-/-}$ males. Resulting embryos exhibited a significant decrease in the numbers of *runx1* and *cmyb* positive cells as compared to *dcn*$^{+/+}$ embryos from wild-type sibling crosses (Figs. 7C, 7D), thus confirming the phenotypes of the *dcn* morphants.

Confocal microscopy using *kdlr:mCherry;gata2b:GAL4;UAS:eGFP* transgenic animals showed a strong reduction in the number of HSPCs in the dorsal aorta in *dcn* morphants (Fig. S10C). Importantly, the vasculature of *flt1:dtomato*;*flt4:yfp* embryos was intact, indicating that the reduction in HSPCs in *dcn* morphants was not related to vascular abnormalities (Fig. S10D).

Since Dcn was known to bind to Tgfb1 and to serve as a negative-feedback regulator of the Tgfb signaling (Yamaguchi et al., 1990), we investigated by RT-qPCR the expression of *Tgfb* ligands and receptors in *dcn* morphants and control embryos. We did not detect any significant differences in the expression of *tgfb2*, *tgfb3*, *tgfr1a* and *tgfr2* (Fig. S9D). In contrast, *tgfb1a* expression was significantly reduced in *dcn* morphants as compared to controls (Fig. S9D). These data suggest that the effect of *dcn* on AGM HSPCs may be related to a deregulation of *Tgfb1* expression which is required for HSPC emergence (Monteiro et al., 2016).

Finally, since the 'blue' module contains several genes encoding for soluble factors, we also selected the Growth Differentiation Factor 6, Gdf6 (Bmp 13), for functional validation. Interestingly, *gdf6a* (but not *gdf6b*) was reported to be expressed in the trunk and around the dorsal aorta in the zebrafish embryo (Krispin et al., 2018). We observed that *gdf6a* morphants exhibited a significant decrease in the number of *runx1* positive cells in the dorsal aorta as compared to uninjected controls (Fig. S11). The effect of *gdf6a* on aortic HSPCs could be indirect since *gdf6a* is required for vascular integrity by restraining Vegf signaling (Krispin et al., 2018, Hall et al., 2002).

Discussion

In this work, we have uncovered at the molecular level the cell identities of the AGM mesenchyme at the time of HSPC formation in the mouse embryo. Our computational strategy proceeded by several steps: first, using bulk transcriptomics and additional datasets, subtracting the genetic contribution of endothelial and hematopoietic cells of the sub-dissected aortic tissues; second, using a correlation-based algorithm, unraveling the gene networks corresponding to the ventral and dorsal tissues; third, using single cell transcriptomics, projecting the gene modules onto the Louvain map to characterize the clusters that could not be defined by highly-ranked gene allocation; and fourth, using an information-based algorithm, pruning the gene network corresponding to the cell population of most interest to better discern critical gene-to-gene links. This computational analysis allowed revealing the presence within a large and heterogeneous mesenchymal cell population of a ventrally-located neuro-mesenchymal subpopulation, which was confirmed by immunofluorescence studies using antibodies against Decorin, Tenascin C and NTRK3, three proteins corresponding to genes *Dcn, Tnc,* and *Ntrk3* that appeared preeminently throughout the in-silico analysis. Moreover, lineage tracing studies indicated that the neuro-mesenchymal population was not derived from the neural crest. Finally, loss-of-function experiments in the in 30- and 36-hours post-fertilization zebrafish confirmed the essential role of the characteristic ECM component Decorin for HSPC development.

The functional association between the hematopoietic and neuronal systems has already been documented both in the adult and embryonic contexts. For example, in the adult

bone marrow, stromal cells expressing the β3-adrenergic receptor were shown to be sensitive to noradrenaline rhythmically released by the sympathetic nervous system; this intercellular connection pathway results in the circadian oscillations of Cxcl12, a major extrinsic factor involved in HSPC maintenance and homing (Méndez-Ferrer et al., 2008). In the embryo, it has been reported that Gata3 controls the number of AGM HSCs through the production of catecholamines, showing a strong influence of the sympathetic nervous system on AGM HSCs (Fitch et al., 2012). Catecholamines may exert their effect directly on AGM HSPCs since the β2-adrenergic receptor expressed by HSPCs is required for HSC production (Fitch et al., 2012, Kapeni et al., 2022). In the zebrafish, trunk NCCs were shown to interact with precursors of aortic HSPCs and to contribute to their supportive microenvironment (Damm and Clements, 2017). In mice, subaortic mesenchymal cells may have a dual origin (mesodermal and neural crest-derived) (Kapeni et al., 2022, Chandrakanthan et al., 2022) with Pdgfra^{+} mesodermal stromal cells being replaced by NCC-derived cells at E13.5, a time coinciding with the loss of EHT (Chandrakanthan et al., 2022). Remarkably, in our present work we identified at E11.5 a mesenchymal cell subpopulation, not derived from NCCs, containing both neuronal genes, implicated in synapse formation, membrane channel activity, and neurotrophin signaling, and mesenchymal genes, in particular some distinctive ECM components. This molecular signature was enriched in the subaortic tissue at E11.5 as compared to E10.5 and E12.5. The neuro-mesenchymal cell population may represent a transitory state between a mesoderm-derived mesenchymal state and a neuroectoderm-derived one (Chandrakanthan et al., 2022, Richard et al., 2013). Moreover, its establishment underneath the dorsal aorta may be connected to the concomitant development of the nervous system and the projection of neuronal fibers on the lateral and ventral sides of the dorsal aorta at mid-gestation. Collectively, our findings prompt us to hypothesize that the sympathetic nervous system and the subaortic neuro-mesenchyme may cooperate to regulate HSPC development in the dorsal aorta. The dual molecular identity of the subaortic mesenchyme suggests that these cells may eventually be responsive to signals delivered by the sympathetic nervous system.

So far, very little is known on the composition and functions of the ECM in the AGM. Tenascin C is known to be expressed underneath the dorsal aorta both in the mouse and human embryo (this paper and (Marshall et al., 1999)) and to promote the differentiation

of hematopoietic and endothelial precursors from human pluripotent stem cells (Uenishi et al., 2014). Matrix metalloproteinases, such as Mmp2 and Mmp9, have been shown to regulate EHT and the migration of aortic HSPCs towards secondary hematopoietic tissues (Travnickova et al., 2015, Theodore et al., 2017). Interestingly, in bone marrow HSPC-supportive stromal cells *Dcn* is not up- but down-regulated as compared to less or non-supportive stromal cells (Desterke et al., 2020), which is in agreement with its reported role as HSPC negative regulator in adult mice (Ichii et al., 2012). These data underscore the context-dependency of some of the mediators, activators in the embryonic niche, but inhibitors in the adult one. Decorin could exert its effects on AGM HSPCs as structural, basement membrane ECM component interacting with other ECM proteins (such as laminins, nidogens, fibronectins or collagens), and/or as substrate for cytokines such as TGFß1 or other soluble mediators, such as widely known for the bone marrow HSPC microenvironment. Our investigations in zebrafish embryos suggest that the reduction in HSPCs in *dcn* morphants was not related to vascular abnormalities. However, since *dcn* was significantly more expressed in *flk1*$^{-}$*cmyb*$^{-}$ and *flk1*$^{+}$*cmyb*$^{-}$ trunk cells, we cannot rule out that the loss-of-function phenotype for *dcn* could also be initiated in endothelial cells.

In conclusion, we identified a unique non-NCC derived mesenchymal cell population harboring both an adhesive and neuronal molecular identity within the subaortic tissue. Our data also highlight the role of a preeminent ECM component as part of the HSPC-supportive niche in the embryo.

Materials and methods

Mouse and human embryos

C57BL/6 mice were bred at Janvier (Le Genest, France) and maintained according to institutional guidelines in the animal facility of the Laboratory of Developmental Biology. Mouse models for lineage tracing were the Tg(PLAT-cre)116Sdu, referred to as (HtPA)-Cre (Pietri et al., 2003), and Cre reporter mice B6.Cg-Gt(ROSA)26Sortm9(CAG-tdTomato)Hze/J (Madisen et al., 2010). Pregnant females were killed by cervical dislocation at appropriate time of gestation considering the day of vaginal plugging represents day 0.5 (E0.5). Uteri were taken and placed in 10% Fetal Bovine Serum/1%

Penicillin/Streptomycin (Gibco)/phosphate-buffered saline (PBS) (Invitrogen). The embryonic stages were confirmed through morphological analysis.

Human embryonic tissues were obtained following voluntary abortions. Informed consent for the use of the embryo in research was obtained from the mother, and tissue collection and use were performed according to the guidelines and with the approval of the French National Ethics Committee.

Laser microdissection, microarray screening and data analysis

We used the Arcturus Laser Capture Microdissection System to isolate from mouse embryos the ventral and dorsal aortic tissues surrounding the dorsal aorta (Fig. 1A area delineated in red). The isolated dorsal tissue was localized in between the notochord and the roof of the aorta and the isolated ventral tissue was underneath the aorta. Unfixed embryos were embedded in OCT medium in presence of isopentane cooled at -65°C. Twenty μm cryostat sections were then prepared and stained before laser micro-dissection. Three to eight micro-dissected tissues, corresponding to a total surface in between 84,431 and 412,076 μm^2, depending on the developmental stages as indicated in Table S1, were collected on appropriate caps and then proceeded for RNA extraction using Acturus Picopure RNA isolation kit. RNAs were eluted in a final volume of 13 μl and their concentrations and integrity quantified on a bioanalyzer (Agilent). For each developmental stage, 3 to 4 RNA samples corresponding to DT and VT were prepared from independent embryos and proceeded for bulk transcriptomics. Between 200 and 1000 pg of total RNA was reverse transcribed following the Ovation Pico V2 system (Nugen) and the resulting double strand cDNA was used for amplification based on SPIA technology. After purification according to Nugen protocol, sense target DNA was fragmented and biotin-labelled using Encore Biotin Module kit (Nugen). After control of fragmentation using Bioanalyzer 2100, cDNA was hybridized to GeneChip® MouseGene 2.0 ST (Affymetrix). Normalization and functional analyses of the datasets were performed using Partek Genomics Suite software.

ANOVA: DT vs. VT comparisons were made using 2-way ANOVA with a low stringency p-value ($p<0.06$) with 'Polarity' and 'Embryo' as grouping variables for E11.5 and E12.5 samples, or three-way ANOVA with 'Polarity', 'Embryo' and 'Series' (corresponding to two

independent transcriptomic datasets) for E10.5 samples. After incorporation of additional transcriptomes (AGM stromal cell lines or AGM sorted cell populations) DT vs. VT comparisons were made using 2-way ANOVA with 'Cell/Tissue identity' and 'Serie' (the different batches) as grouping variables.

Gene ontology (GO) categories: they were analyzed using DAVID (https://david.ncifcrf.gov) (Huang et al., 2007).

GSEA: GSEA aims at identifying the significant genes within a gene expression set contrasting two phenotypes (DT vs. VT). The expression set is compared to pre-established datasets corresponding to a number of collection profiles, in this work the 'hallmark' database. Datasets are retained if the distribution of their genes is shifted toward either phenotype and the statistical relevance of the shift is established by random permutations (here 1000) of the phenotypes (Subramanian et al., 2005).

WGCNA: To reconstruct gene network and identify specific modules for the different cell populations, we utilized WGCNA (Langfelder and Horvath, 2008). The chosen values for essential input parameters were: 1) the power ß obtained from the scale-free fit model curve ($r^2 > 0,8$); the ß power of 18 was selected such that the connectivity distribution fitted a scale-free model (Fig. S1); 2) the Pearson's correlation; 3) the signed adjacency; 4) the average linkage hierarchical clustering; 5) the minimum module size of 30. Cytoscape (https://cytoscape.org) was used to visualize the connectivity plot of the most connected genes. Force-oriented layout was used for data representation.

Multivariate information-based inductive causation (miic) analysis: The miic algorithm combines constraint-based and information-theoretic frameworks to reconstruct causal and non-causal networks from large scale datasets (Verny et al., 2017; Sella et al., 2018). Starting from a fully connected graph, miic iteratively removes dispensable edges, by uncovering significant information contributions from indirect paths. Miic also provides an edge-specific confidence assessment of retained edges, which are oriented based on the signature of causality in observational data.

Single Cell screening and data analysis

Seven E11.5 AGMs from 3 litters were dissected (Fig. 1A area delineated in green), digested with collagenase and submitted to 10X Genomics Chromium next GEM single cell 3' kit version 3.1 (2 technical replicates). Only cells expressing between 200 and 7000 genes with a percentage of mitochondrial RNA under 15% were kept. In addition, genes expressed in less than 3 cells were excluded from the analysis. Since no batch effect was observed between the replicates, the two datasets were merged without correction. ScTransform normalization (Hafemeister and Satija, 2019) was performed with the Seurat package v3 (Stuart et al., 2019) with regression of the cell cycle. To regress the cell cycle genes, the default human cell cycle gene list given by Seurat was converted into mouse genes using the biomaRt package (Durinck et al., 2009) and was used following the default recommendation for cell cycle regression with scTransform as described in the Seurat vignette. Clustering was performed with the Louvain algorithm with a resolution of 0.7 on the first 50 PCA components and the cells were projected on a Uniform Manifold Approximation and projection (UMAP) embedding based on the first 42 PCA components. We identified then the differentially expressed genes selecting genes with a fold-change superior to 0.25 in each cluster compared to the whole dataset using the Wilcoxon's sum rank test. Only genes expressed in more than 25% of the cluster cells and with a q-value under 0.05 were kept. These genes were then ranked and used to determine each cluster identity. For the gene scoring, we defined two lists of mesenchymal and neural genes belonging to the WGCNA blue VT module and used the Seurat AddModuleScore function. This function returns positive values if the genes from the module displays a higher expression than control genes selected at random among the ones with a similar average expression over the dataset. Multivariate information-based inductive causation (miic) algorithm (Verny et al., 2017, Sella et al., 2018) was used to analyze the genes of the blue VT module taking into account all 12516 cells belonging to the large central aggregate defined by the Louvain clustering (C0-C12), and the raw counts for each gene of the VT 'blue' module core (89 genes).

Flow cytometry and cell sorting

E11.5 AGMs were dissociated in 0.125% collagenase type I (Sigma) in PBS supplemented with 10% fetal calf serum (FCS) for 20 min at 37°C. Cells were dispersed by gentle pipetting, washed twice, filtered through 70µm nylon cell strainer (BD Falcon) and resuspended in Dulbecco's modified medium (DMEM) (Invitrogen) supplemented with

10% FCS. Cells were labeled with APC anti-mouse CD34 (clone RAM34, eBiosciences), PE anti-mouse CD117 (clone 2B8) and FITC anti-mouse CD45 (clone 30F-11) or with PE anti-mouse CD144 (clone BV13) and APC anti-mouse CD45 from Biolegend (Ozyme, France). Viability was determined using 7-AAD (Beckman Coulter). Cell sorting was performed on a FacsVantage or a FacsAria (BD Biosciences) cell sorter. Data were reanalyzed with FlowJo X (Treestar) software. AGM $CD144^{+}CD45^{-}$ endothelial cells, $CD34^{low}Kit^{low}CD45^{pos}$ differentiated hematopoietic cells and $CD34^{high}Kit^{high}$ HSPCs were isolated by flow cytometry from E11.5 C57Bl/6 mouse embryos. Three and 5 independent experiments were performed for the hematopoietic fractions (in between 75 to 82 embryos per experience) and endothelial fraction (in between 29 to 80 embryos per experience), respectively.

Immunodetection

Embryos were fixed overnight in 4% paraformaldehyde (Sigma) at 4°C, transferred in a solution containing PBS/sucrose 15% for 48 hours, impregnated in PBS/15% sucrose/7.5% gelatin for 2 hours at 37°C, embedded at room temperature, frozen in isopentane at -65°C using liquid nitrogen and stored at -80°C until used for cryosectioning. Frozen sections (20 mm) of E11.5 embryos were cut on a cryostat (Leica). Antibodies used for staining the sections were: monoclonal mouse anti-mouse/human Tnc (1/50, Thermofisher, MA5-16086), polyclonal goat anti-mouse Decorin (1/200, R&D Systems, AF1060), monoclonal mouse anti-human Decorin (1/200, R&D Systems, MAB143), polyclonal goat anti-mouse Ntrk3 (1/200, R&D Systems, AF1404), polyclonal rabbit anti-mouse CD31(1/200, Abcam, Ab28364), monoclonal mouse anti-mouse α-smooth muscle actin (1/500, Sigma, A5228), polyclonal rabbit anti-mouse Iba1 (1/1000, Wako, SAB5701363), monoclonal mouse anti-mouse Tuj1 antibody (1/1000, BioLegend, MMS-435P) and polyclonal goat anti-human CD105 (1/100, R&D Systems, AF1097). These primary antibodies were visualized with bovine anti-goat IgG-Alexa647 (1/400, Jackson ImmunoResearch, 805-605-180), goat anti-mouse IgG2a-Alexa488 (1/200, Invitrogen, A-21131), donkey anti-mouse IgG-Alexa488 (1/400, Jackson ImmunoResearch, 715-545-150) and donkey anti-rabbit Cy3 (1/400, Jackson ImmunoResearch, 711-165-152). Fluorescent imaging was performed using fluorescence Leica microscope and confocal Leica TCS SP5 at the imaging facility, IBPS, Sorbonne Université.

Whole-mount immunostaining, tissue clearing and 3D imaging

Embryos were incubated at RT for 3 hours in a solution (PBSGT) containing PBS 1X 0.2% gelatin, 0.5% Triton X100 (Sigma-Aldrich) and 0.01% thimerosal (Sigma-Aldrich). Embryos were then incubated with the first antibodies as described above in PBSGT at RT for 2-3 days. After 6 washes, samples were incubated with secondary antibodies in PBSGT overnight, washed and processed for 3DISCO tissue clearing as described by Belle et al (Belle et al., 2014). Images were acquired using an ultramicroscope (LaVision BioTec). 3D imaging, stack images conversion and movies were generated using Vision 4D 3.0 software (Arivis).

Sample preparation for transmission electron microscopy analysis

The mouse embryo sections were immersed in the fixative solution 1.5% glutaraldehyde and 1% PFA in 0.1M cacodylate pH 7.4 1h at room temperature and overnight at 4°C. Samples were washed with cacodylate buffer 0.1M, incubated for 1h in 0.2% OTE in 0.1M cacodylate buffer and washed again with cacodylate buffer 0.1M. To keep them flat, the sections were embedded between two slides with 4% low melting point agarose in 0.1M cacodylate. Afterwards, the samples were incubated 1 h in 1% osmium tetroxide in 0.1M cacodylate buffer and washed with deionized water. The mouse embryo sections were dehydrated through graded concentration of ethanol (50-70-95- 100%). Samples were pre-embedded with graded concentration of anhydrous acetone and EPON resin mix (3:1 – 1:1 – 1:3) and embedded with 100% EPON resin (polymerized at 60°C for 24h) between two glass slides coated with silicone. Finally, the flat sections were mounted on EPON resin blocks (preliminarily polymerized for 72h) and cut with the Ultracut ultramicrotome (UCT, Leica microsystems) at 70 nm thickness. Ultra-fine sections deposited on 75-mesh copper grids were contrasted with 2.5% uranyl acetate and 2% lead citrate and coated with a 2 nm carbon film (ACE600, Leica microsystems). The grids were observed at 20 kV, with a STEM detector in a high-resolution Field-Emission SEM (GeminiSEM500, Zeiss) with a 20 µm aperture diameter. Images were acquired with a 1024 x 768 definition, with a pixel dwell time of 1.6 µs and a line averaging of 15.

Zebrafish studies

The following zebrafish lines were used: *Tg(kdrl*:*mCherry*) (Chi et al., 2008)*, Tg(Gata2b*:*Gal4/UAS*:*LA-GFP)* (Butko et al., 2015), *Tg(flt1:dtomato)* (Bussmann et al., 2010), *Tg(flt4:YFP)* (Hogan et al., 2009), *Tg(cmyb:GFP)* (North et al., 2007) and *dcn*^{sa20247} (from the European Zebrafish Resource Center).

Zebrafish were maintained according to the 'University of California, San Diego, International Animal Care' and 'User Committee and aquatic animal guidelines of Sorbonne Université'. Wild-type or transgenic zygotes were injected with a morpholino solution and incubated at 28.5°C until they reached the stage of interest. Embryos were collected, staged, fixed, and processed for *in situ* hybridization as previously described (Thisse et al., 1993). Antisense *dcn* morpholinos (Gene Tools, LLC) were used to target/block the 5'-untranslated region/translation start (MO^{START}) or a splice junction (MO^{SPLICE}). *Dcn* morpholino sequences were as follow: MO^{START} GACAGGCCGATTTCATGTTGCTGAC and MO^{SPLICE} GCAGACCTGGGCATTTTGACACAGA (Zoeller et al., 2009). *Gdf6a* morpholino sequence (Gene Tools, LLC) was as follow: GCAATACAAACCTTTTCCCTTGTCC (Sidi et al., 2003). To determine the phenotype of *dcn* morphants and mutants on hematopoietic production, embryos were in embedded in 80% glycerol and photographed using a Leica IC80 HD camera mounted on a Leica M60 stereoscope. For quantification of *runx1*- and *cmyb*-expressing cells in the dorsal aorta, all morphant, mutant and control embryos were considered and quantified individually.

For flow cytometry, fluorescent embryos (around 300 embryos) were tail cut at 48 hpf, and the tissue was treated with Liberase (Roche) and subsequently manually dissociated. The cell suspension was filtered through 30 μm mesh filter (CellTrics) and resuspended in 1XPBS + 1mM EDTA with SYTOX Red. Cells were sorted using FACSAria II 6 lasers. Data was processed using FlowJo software and statistics were calculated using GraphPrism 10. RT-qPCR for sorted cells was performed using TaqMan *gapdh* (#Dr03436842_m1) and *dcn* (#Dr03093424_m1) primers. For RT-qPCR to investigate *tgfβ* ligands and receptors expression, RNA (around 25 embryos per condition) was isolated using Qiagen RNeasy Micro Kit from zebrafish trunks at 30 hpf, and 200 ng of RNA was used to generate cDNA

using the Invitrogen SuperScript III. RT-qPCR was performed with SYBR™ Reagents. Samples were run in triplicates, with 4 biological replicates using the primers listed below. Statistics were calculated using GraphPad Prism 10. The sequences of the primers are indicated in Table S4.

Statistical analysis

Statistical analysis, sample sizes, p-values are described in figure legends and in the appropriate materials and methods sections for computational analysis. For the loss-of-function experiments, statistical analysis was performed using GraphPad Prism software and unpaired t-test and Mann and Whitney nonparametric test were used to evaluate the statistical significance.

Acknowledgments

We thank Claire Francastel (CNRS UMR7622), Laurent Yvernogeau (CNRS UMR7622) and Pascal Legendre for critical reading of the manuscript, and all laboratory members for helpful discussion. We thank Srinath Ramkumar and Didier Stainier (Max Planck Institute for heart and lung research, Germany) for providing us with the zebrafish *dcn* construct. We thank also Monara Angelim and Pascal Legendre (CNRS UMR8246) for providing us with the anti-Tuj1 and –Iba1 antibodies, members of the imaging (Institut de la Vision and Institut de Biologie Paris-Seine) facilities for their expertise and advices, Michael Trichet and Alexis Canette (Service de microscopie électronique, Institut de Biologie Paris Seine, Sorbonne Université, CNRS) for their support and advices on the transmission electron microscopy experiments, and Sophie Gournet (CNRS UMR7622) for drawing assistance.

Competing interests

The authors declare no competing or financial interests.

Funding

This work was supported by grants of the Fondation pour la Recherche Médicale (DEQ20100318258 and EQU201903007867), Centre National de la Recherche Scientifique (CNRS), Institut national de la santé et de la recherche médicale (Inserm) and Sorbonne Université to TJ; and the National Institutes of Health (NIH DK074482) to DT.

Author contributions

Conceptualization: TJ, PC and CD
Methodology: OM, PYC, CP, OP, SJ, EH, AG, SD, LP, MS, HI, TN, DT, TJ, PC and CD
Software: OP, HI
Validation: PC and CD
Formal analysis: OP, HI, PC and CD
Investigation: OM, PYC, CP, OP, NR, PF, AM, CBP, EH, LP and CD
Data curation: PC and CD
Writing, original writing: TJ, PC and CD
Writing, review and editing: OM, PYC, EH, SD, MS, DT, TJ, PC and CD
Visualization: OM, PYC, CP, OP, NR, PF, AM, CBP, LP, PC and CD
Supervision: TJ, PC and CD
Project administration: TJ, PC and CD
Funding acquisition: DT and TJ

Data availability

The NCBI GEO accession number for the microarrays and single cell RNA sequencing presented in this study is GSE159592.

References

ADAMO, L., NAVEIRAS, O., WENZEL, P. L., MCKINNEY-FREEMAN, S., MACK, P. J., GRACIA-SANCHO, J., SUCHY-DICEY, A., YOSHIMOTO, M., LENSCH, M. W., YODER, M. C., GARCÍA-CARDEÑA, G. & DALEY, G. Q. 2009. Biomechanical forces promote embryonic haematopoiesis. *Nature,* 459**,** 1131-5.

BARON, C. S., KESTER, L., KLAUS, A., BOISSET, J. C., THAMBYRAJAH, R., YVERNOGEAU, L., KOUSKOFF, V., LACAUD, G., VAN OUDENAARDEN, A. & ROBIN, C. 2018. Single-cell transcriptomics reveal the dynamic of haematopoietic stem cell production in the aorta. *Nat Commun,* 9**,** 2517.

BELLE, M., GODEFROY, D., DOMINICI, C., HEITZ-MARCHALAND, C., ZELINA, P., HELLAL, F., BRADKE, F. & CHÉDOTAL, A. 2014. A simple method for 3D analysis of immunolabeled axonal tracts in a transparent nervous system. *Cell Rep,* 9**,** 1191-201.

BERTRAND, J. Y., CHI, N. C., SANTOSO, B., TENG, S., STAINIER, D. Y. & TRAVER, D. 2010. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature,* 464**,** 108-11.

BOISSET, J. C., VAN CAPPELLEN, W., ANDRIEU-SOLER, C., GALJART, N., DZIERZAK, E. & ROBIN, C. 2010. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature,* 464**,** 116-20.

BUSSMANN, J., BOS, F. L., URASAKI, A., KAWAKAMI, K., DUCKERS, H. J. & SCHULTE-MERKER, S. 2010. Arteries provide essential guidance cues for lymphatic endothelial cells in the zebrafish trunk. *Development,* 137**,** 2653-7.

BUTKO, E., DISTEL, M., POUGET, C., WEIJTS, B., KOBAYASHI, I., NG, K., MOSIMANN, C., POULAIN, F. E., MCPHERSON, A., NI, C. W., STACHURA, D. L., DEL CID, N., ESPIN-PALAZON, R., LAWSON, N. D., DORSKY, R., CLEMENTS, W. K. & TRAVER, D. 2015. Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development,* 142**,** 1050-61.

CALVANESE, V., CAPELLERA-GARCIA, S., MA, F., FARES, I., LIEBSCHER, S., NG, E. S., EKSTRAND, S., AGUADE-GORGORIO, J., VAVILINA, A., LEFAUDEUX, D., NADEL, B., LI, J. Y., WANG, Y., LEE, L. K., ARDEHALI, R., IRUELA-ARISPE, M. L., PELLEGRINI, M., STANLEY, E. G., ELEFANTY, A. G., SCHENKE-LAYLAND, K. & MIKKOLA, H. K. A. 2022. Mapping human haematopoietic stem cells from haemogenic endothelium to birth. *Nature,* 604**,** 534-540.

CHANDRAKANTHAN, V., RORIMPANDEY, P., ZANINI, F., CHACON, D., OLIVIER, J., JOSHI, S., KANG, Y. C., KNEZEVIC, K., HUANG, Y., QIAO, Q., OLIVER, R. A., UNNIKRISHNAN, A., CARTER, D. R., LEE, B., BROWNLEE, C., POWER, C., BRINK, R., MENDEZ-FERRER, S., ENIKOLOPOV, G., WALSH, W., GOTTGENS, B., TAOUDI, S., BECK, D. & PIMANDA, J. E. 2022. Mesoderm-derived PDGFRA(+) cells regulate the emergence of hematopoietic stem cells in the dorsal aorta. *Nat Cell Biol,* 24**,** 1211-1225.

CHARBORD, P., POUGET, C., BINDER, H., DUMONT, F., STIK, G., LEVY, P., ALLAIN, F., MARCHAL, C., RICHTER, J., UZAN, B., PFLUMIO, F., LETOURNEUR, F., WIRTH, H., DZIERZAK, E., TRAVER, D., JAFFREDO, T. & DURAND, C. 2014. A systems biology approach for defining the molecular framework of the hematopoietic stem cell niche. *Cell Stem Cell,* 15**,** 376-391.

CHEN, M. J., YOKOMIZO, T., ZEIGLER, B. M., DZIERZAK, E. & SPECK, N. A. 2009. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature,* 457**,** 887-91.

CHI, N. C., SHAW, R. M., DE VAL, S., KANG, G., JAN, L. Y., BLACK, B. L. & STAINIER, D. Y. 2008. Foxn4 directly regulates tbx2b expression and atrioventricular canal formation. *Genes Dev,* 22**,** 734-9.

CROSSE, E. I., GORDON-KEYLOCK, S., RYBTSOV, S., BINAGUI-CASAS, A., FELCHLE, H., NNADI, N. C., KIRSCHNER, K., CHANDRA, T., TAMAGNO, S., WEBB, D. J., ROSSI, F., ANDERSON, R. A. & MEDVINSKY, A. 2020. Multi-layered Spatial Transcriptomics Identify Secretory Factors Promoting Human Hematopoietic Stem Cell Development. *Cell Stem Cell,* 27**,** 822-839.e8.

DAMM, E. W. & CLEMENTS, W. K. 2017. Pdgf signalling guides neural crest contribution to the haematopoietic stem cell specification niche. *Nat Cell Biol,* 19**,** 457-467.

DE BRUIJN, M. F., MA, X., ROBIN, C., OTTERSBACH, K., SANCHEZ, M. J. & DZIERZAK, E. 2002. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity,* 16**,** 673-83.

DESTERKE, C., PETIT, L., SELLA, N., CHEVALLIER, N., CABELI, V., COQUELIN, L., DURAND, C., OOSTENDORP, R. A. J., ISAMBERT, H., JAFFREDO, T. & CHARBORD, P. 2020. Inferring Gene Networks in Bone Marrow Hematopoietic Stem Cell-Supporting Stromal Niche Populations. *iScience,* 23**,** 101222.

DURAND, C., ROBIN, C., BOLLEROT, K., BARON, M. H., OTTERSBACH, K. & DZIERZAK, E. 2007. Embryonic stromal clones reveal developmental regulators of definitive hematopoietic stem cells. *Proc Natl Acad Sci U S A,* 104**,** 20838-43.

DURINCK, S., SPELLMAN, P. T., BIRNEY, E. & HUBER, W. 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc,* 4**,** 1184-91.

EASTERBROOK, J., RYBTSOV, S., GORDON-KEYLOCK, S., IVANOVS, A., TAOUDI, S., ANDERSON, R. A. & MEDVINSKY, A. 2019. Analysis of the Spatiotemporal Development of Hematopoietic Stem and Progenitor Cells in the Early Human Embryo. *Stem Cell Reports,* 12**,** 1056-1068.

ESPÍN-PALAZÓN, R., STACHURA, D. L., CAMPBELL, C. A., GARCÍA-MORENO, D., DEL CID, N., KIM, A. D., CANDEL, S., MESEGUER, J., MULERO, V. & TRAVER, D. 2014. Proinflammatory signaling regulates hematopoietic stem cell emergence. *Cell,* 159**,** 1070-1085.

FADLULLAH, M. Z. H., NEO, W. H., LIE, A. L. M., THAMBYRAJAH, R., PATEL, R., MEVEL, R., AKSOY, I., DO KHOA, N., SAVATIER, P., FONTENILLE, L., BAKER, S. M., RATTRAY, M., KOUSKOFF, V. & LACAUD, G. 2022. Murine AGM single-cell profiling identifies a continuum of hemogenic endothelium differentiation marked by ACE. *Blood,* 139**,** 343-356.

FITCH, S. R., KIMBER, G. M., WILSON, N. K., PARKER, A., MIRSHEKAR-SYAHKAL, B., GÖTTGENS, B., MEDVINSKY, A., DZIERZAK, E. & OTTERSBACH, K. 2012. Signaling from the sympathetic nervous system regulates hematopoietic stem cell emergence during embryogenesis. *Cell Stem Cell,* 11**,** 554-66.

HAFEMEISTER, C. & SATIJA, R. 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol,* 20**,** 296.

HALL, C. J., FLORES, M. V., DAVIDSON, A. J., CROSIER, K. E. & CROSIER, P. S. 2002. Radar is required for the establishment of vascular integrity in the zebrafish. *Dev Biol,* 251**,** 105-17.

HOGAN, B. M., HERPERS, R., WITTE, M., HELOTERA, H., ALITALO, K., DUCKERS, H. J. & SCHULTE-MERKER, S. 2009. Vegfc/Flt4 signalling is suppressed by Dll4 in developing zebrafish intersegmental arteries. *Development,* 136**,** 4001-9.

HUANG, D. W., SHERMAN, B. T., TAN, Q., KIR, J., LIU, D., BRYANT, D., GUO, Y., STEPHENS, R., BASELER, M. W., LANE, H. C. & LEMPICKI, R. A. 2007. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res,* 35**,** W169-75.

ICHII, M., FRANK, M. B., IOZZO, R. V. & KINCADE, P. W. 2012. The canonical Wnt pathway shapes niches supportive of hematopoietic stem/progenitor cells. *Blood,* 119**,** 1683-92.

JAFFREDO, T., GAUTIER, R., EICHMANN, A. & DIETERLEN-LIÈVRE, F. 1998. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development,* 125**,** 4575-83.

KAPENI, C., NITSCHE, L., KILPATRICK, A. M., WILSON, N. K., XIA, K., MIRSHEKAR-SYAHKAL, B., CHANDRAKANTHAN, V., MALOUF, C., PIMANDA, J. E., GOTTGENS, B., KIRSCHNER, K., TOMLINSON, S. R. & OTTERSBACH, K. 2022. p57Kip2 regulates embryonic blood stem cells by controlling sympathoadrenal progenitor expansion. *Blood,* 140**,** 464-477.

KISSA, K. & HERBOMEL, P. 2010. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature,* 464**,** 112-5.

KRISPIN, S., STRATMAN, A. N., MELICK, C. H., STAN, R. V., MALINVERNO, M., GLEKLEN, J., CASTRANOVA, D., DEJANA, E. & WEINSTEIN, B. M. 2018. Growth Differentiation Factor 6 Promotes Vascular Stability by Restraining Vascular Endothelial Growth Factor Signaling. *Arterioscler Thromb Vasc Biol,* 38**,** 353-362.

KUMARAVELU, P., HOOK, L., MORRISON, A. M., URE, J., ZHAO, S., ZUYEV, S., ANSELL, J. & MEDVINSKY, A. 2002. Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. *Development,* 129**,** 4891-9.

LANGFELDER, P. & HORVATH, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics,* 9**,** 559.

LEE, Y., MANEGOLD, J. E., KIM, A. D., POUGET, C., STACHURA, D. L., CLEMENTS, W. K. & TRAVER, D. 2014. FGF signalling specifies haematopoietic stem cells through its regulation of somitic Notch signalling. *Nat Commun,* 5**,** 5583.

LI, Y., ESAIN, V., TENG, L., XU, J., KWAN, W., FROST, I. M., YZAGUIRRE, A. D., CAI, X., CORTES, M., MAIJENBURG, M. W., TOBER, J., DZIERZAK, E., ORKIN, S. H., TAN, K., NORTH, T. E. & SPECK, N. A. 2014. Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes Dev,* 28**,** 2597-612.

MADISEN, L., ZWINGMAN, T. A., SUNKIN, S. M., OH, S. W., ZARIWALA, H. A., GU, H., NG, L. L., PALMITER, R. D., HAWRYLYCZ, M. J., JONES, A. R., LEIN, E. S. & ZENG, H. 2010. A robust and

high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci,* 13**,** 133-40.

MARIANI, S. A., LI, Z., RICE, S., KRIEG, C., FRAGKOGIANNI, S., ROBINSON, M., VINK, C. S., POLLARD, J. W. & DZIERZAK, E. 2019. Pro-inflammatory Aorta-Associated Macrophages Are Involved in Embryonic Development of Hematopoietic Stem Cells. *Immunity,* 50**,** 1439-1452.e5.

MARSHALL, C. J., KINNON, C. & THRASHER, A. J. 2000. Polarized expression of bone morphogenetic protein-4 in the human aorta-gonad-mesonephros region. *Blood,* 96**,** 1591-3.

MARSHALL, C. J., MOORE, R. L., THOROGOOD, P., BRICKELL, P. M., KINNON, C. & THRASHER, A. J. 1999. Detailed characterization of the human aorta-gonad-mesonephros region reveals morphological polarity resembling a hematopoietic stromal layer. *Dev Dyn,* 215**,** 139-47.

MASCARENHAS, M. I., PARKER, A., DZIERZAK, E. & OTTERSBACH, K. 2009. Identification of novel regulators of hematopoietic stem cell development through refinement of stem cell localization and expression profiling. *Blood,* 114**,** 4645-53.

MCGARVEY, A. C., RYBTSOV, S., SOUILHOL, C., TAMAGNO, S., RICE, R., HILLS, D., GODWIN, D., RICE, D., TOMLINSON, S. R. & MEDVINSKY, A. 2017. A molecular roadmap of the AGM region reveals BMPER as a novel regulator of HSC maturation. *J Exp Med,* 214**,** 3731-3751.

MCKINNEY-FREEMAN, S., CAHAN, P., LI, H., LACADIE, S. A., HUANG, H. T., CURRAN, M., LOEWER, S., NAVEIRAS, O., KATHREIN, K. L., KONANTZ, M., LANGDON, E. M., LENGERKE, C., ZON, L. I., COLLINS, J. J. & DALEY, G. Q. 2012. The transcriptional landscape of hematopoietic stem cell ontogeny. *Cell Stem Cell,* 11**,** 701-14.

MEDVINSKY, A. & DZIERZAK, E. 1996. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell,* 86**,** 897-906.

MÉNDEZ-FERRER, S., LUCAS, D., BATTISTA, M. & FRENETTE, P. S. 2008. Haematopoietic stem cell release is regulated by circadian oscillations. *Nature,* 452**,** 442-7.

MONTEIRO, R., PINHEIRO, P., JOSEPH, N., PETERKIN, T., KOTH, J., REPAPI, E., BONKHOFER, F., KIRMIZITAS, A. & PATIENT, R. 2016. Transforming Growth Factor beta Drives Hemogenic Endothelium Programming and the Transition to Hematopoietic Stem Cells. *Dev Cell,* 38**,** 358-70.

MÜLLER, A. M., MEDVINSKY, A., STROUBOULIS, J., GROSVELD, F. & DZIERZAK, E. 1994. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity,* 1**,** 291-301.

NORTH, T. E., GOESSLING, W., PEETERS, M., LI, P., CEOL, C., LORD, A. M., WEBER, G. J., HARRIS, J., CUTTING, C. C., HUANG, P., DZIERZAK, E. & ZON, L. I. 2009. Hematopoietic stem cell development is dependent on blood flow. *Cell,* 137**,** 736-48.

NORTH, T. E., GOESSLING, W., WALKLEY, C. R., LENGERKE, C., KOPANI, K. R., LORD, A. M., WEBER, G. J., BOWMAN, T. V., JANG, I. H., GROSSER, T., FITZGERALD, G. A., DALEY, G. Q., ORKIN, S. H. & ZON, L. I. 2007. Prostaglandin E2 regulates vertebrate haematopoietic stem cell homeostasis. *Nature,* 447**,** 1007-11.

OBERLIN, E., TAVIAN, M., BLAZSEK, I. & PÉAULT, B. 2002. Blood-forming potential of vascular endothelium in the human embryo. *Development,* 129**,** 4147-57.

OOSTENDORP, R. A., HARVEY, K. N., KUSADASI, N., DE BRUIJN, M. F., SARIS, C., PLOEMACHER, R. E., MEDVINSKY, A. L. & DZIERZAK, E. A. 2002. Stromal cell lines from mouse aorta-gonads-mesonephros subregions are potent supporters of hematopoietic stem cell activity. *Blood,* 99**,** 1183-9.

PEETERS, M., OTTERSBACH, K., BOLLEROT, K., ORELIO, C., DE BRUIJN, M., WIJGERDE, M. & DZIERZAK, E. 2009. Ventral embryonic tissues and Hedgehog proteins induce early AGM hematopoietic stem cell development. *Development,* 136**,** 2613-21.

PETIT-COCAULT, L., VOLLE-CHALLIER, C., FLEURY, M., PÉAULT, B. & SOUYRI, M. 2007. Dual role of Mpl receptor during the establishment of definitive hematopoiesis. *Development,* 134**,** 3031-40.

PIETRI, T., EDER, O., BLANCHE, M., THIERY, J. P. & DUFOUR, S. 2003. The human tissue plasminogen activator-Cre mouse: a new tool for targeting specifically neural crest cells and their derivatives in vivo. *Dev Biol,* 259**,** 176-87.

POUGET, C., PETERKIN, T., SIMÕES, F. C., LEE, Y., TRAVER, D. & PATIENT, R. 2014. FGF signalling restricts haematopoietic stem cell specification via modulation of the BMP pathway. *Nat Commun,* 5**,** 5588.

RICHARD, C., DREVON, C., CANTO, P. Y., VILLAIN, G., BOLLÉROT, K., LEMPEREUR, A., TEILLET, M. A., VINCENT, C., ROSSELLÓ CASTILLO, C., TORRES, M., PIWARZYK, E., SPECK, N. A., SOUYRI, M. & JAFFREDO, T. 2013. Endothelio-mesenchymal interaction controls runx1 expression and modulates the notch pathway to initiate aortic hematopoiesis. *Dev Cell,* 24**,** 600-11.

ROBERT-MORENO, A., ESPINOSA, L., DE LA POMPA, J. L. & BIGAS, A. 2005. RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells. *Development,* 132**,** 1117-26.

ROBIN, C., OTTERSBACH, K., DURAND, C., PEETERS, M., VANES, L., TYBULEWICZ, V. & DZIERZAK, E. 2006. An unexpected role for IL-3 in the embryonic development of hematopoietic stem cells. *Dev Cell,* 11**,** 171-80.

SA DA BANDEIRA, D., KILPATRICK, A. M., MARQUES, M., GOMEZ-SALAZAR, M., VENTURA, T., GONZALEZ, Z. N., STEFANCOVA, D., ROSSI, F., VERMEREN, M., VINK, C. S., BELTRAN, M., HENDERSON, N. C., JUNG, B., VAN DER LINDEN, R., VAN DE WERKEN, H. J. G., VAN IJCKEN, W. F. J., BETSHOLTZ, C., FORBES, S. J., CUERVO, H. & CRISAN, M. 2022. PDGFRbeta(+) cells play a dual role as hematopoietic precursors and niche cells during mouse ontogeny. *Cell Rep,* 40**,** 111114.

SAHAI-HERNANDEZ, P., POUGET, C., EYAL, S., SVOBODA, O., CHACON, J., GRIMM, L., GJOEN, T. & TRAVER, D. 2023. Dermomyotome-derived endothelial cells migrate to the dorsal aorta to support hematopoietic stem cell emergence. *Elife,* 12.

SAWAMIPHAK, S., KONTARAKIS, Z. & STAINIER, D. Y. 2014. Interferon gamma signaling positively regulates hematopoietic stem cell emergence. *Dev Cell,* 31**,** 640-53.

SELLA, N., VERNY, L., UGUZZONI, G., AFFELDT, S. & ISAMBERT, H. 2018. MIIC online: a web server to reconstruct causal or non-causal networks from non-perturbative data. *Bioinformatics,* 34**,** 2311-2313.

SIDI, S., GOUTEL, C., PEYRIERAS, N. & ROSA, F. M. 2003. Maternal induction of ventral fate by zebrafish radar. *Proc Natl Acad Sci U S A,* 100**,** 3315-20.

SOLAIMANI KARTALAEI, P., YAMADA-INAGAWA, T., VINK, C. S., DE PATER, E., VAN DER LINDEN, R., MARKS-BLUTH, J., VAN DER SLOOT, A., VAN DEN HOUT, M., YOKOMIZO, T., VAN SCHAICK-SOLERNO, M. L., DELWEL, R., PIMANDA, J. E., VAN, I. W. F. & DZIERZAK, E. 2015. Whole-transcriptome analysis of endothelial to hematopoietic stem cell transition reveals a requirement for Gpr56 in HSC generation. *J Exp Med,* 212**,** 93-106.

SOUILHOL, C., GONNEAU, C., LENDINEZ, J. G., BATSIVARI, A., RYBTSOV, S., WILSON, H., MORGADO-PALACIN, L., HILLS, D., TAOUDI, S., ANTONCHUK, J., ZHAO, S. & MEDVINSKY, A. 2016. Inductive interactions mediated by interplay of asymmetric signalling underlie development of adult haematopoietic stem cells. *Nat Commun,* 7**,** 10784.

STUART, T., BUTLER, A., HOFFMAN, P., HAFEMEISTER, C., PAPALEXI, E., MAUCK, W. M., 3RD, HAO, Y., STOECKIUS, M., SMIBERT, P. & SATIJA, R. 2019. Comprehensive Integration of Single-Cell Data. *Cell,* 177**,** 1888-1902.e21.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102**,** 15545-50.

SZKLARCZYK, D., GABLE, A. L., LYON, D., JUNGE, A., WYDER, S., HUERTA-CEPAS, J., SIMONOVIC, M., DONCHEVA, N. T., MORRIS, J. H., BORK, P., JENSEN, L. J. & MERING, C. V. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res,* 47**,** D607-d613.

TAOUDI, S. & MEDVINSKY, A. 2007. Functional identification of the hematopoietic stem cell niche in the ventral domain of the embryonic dorsal aorta. *Proc Natl Acad Sci U S A,* 104**,** 9399-403.

TAVIAN, M., COULOMBEL, L., LUTON, D., CLEMENTE, H. S., DIETERLEN-LIÈVRE, F. & PÉAULT, B. 1996. Aorta-associated CD34+ hematopoietic cells in the early human embryo. *Blood,* 87**,** 67-72.

THEODORE, L. N., HAGEDORN, E. J., CORTES, M., NATSUHARA, K., LIU, S. Y., PERLIN, J. R., YANG, S., DAILY, M. L., ZON, L. I. & NORTH, T. E. 2017. Distinct Roles for Matrix Metalloproteinases 2 and 9 in Embryonic Hematopoietic Stem Cell Emergence, Migration, and Niche Colonization. *Stem Cell Reports,* 8**,** 1226-1241.

THISSE, C., THISSE, B., SCHILLING, T. F. & POSTLETHWAIT, J. H. 1993. Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development,* 119**,** 1203-15.

TRAVNICKOVA, J., TRAN CHAU, V., JULIEN, E., MATEOS-LANGERAK, J., GONZALEZ, C., LELIEVRE, E., LUTFALLA, G., TAVIAN, M. & KISSA, K. 2015. Primitive macrophages control HSPC mobilization and definitive haematopoiesis. *Nat Commun,* 6**,** 6227.

UENISHI, G., THEISEN, D., LEE, J. H., KUMAR, A., RAYMOND, M., VODYANIK, M., SWANSON, S., STEWART, R., THOMSON, J. & SLUKVIN, I. 2014. Tenascin C promotes hematoendothelial development and T lymphoid commitment from human pluripotent stem cells in chemically defined conditions. *Stem Cell Reports,* 3**,** 1073-84.

VERNY, L., SELLA, N., AFFELDT, S., SINGH, P. P. & ISAMBERT, H. 2017. Learning causal networks with latent variables from multivariate information in genomic data. *PLoS Comput Biol,* 13**,** e1005662.

WILKINSON, R. N., POUGET, C., GERING, M., RUSSELL, A. J., DAVIES, S. G., KIMELMAN, D. & PATIENT, R. 2009. Hedgehog and Bmp polarize hematopoietic stem cell emergence in the zebrafish dorsal aorta. *Dev Cell,* 16**,** 909-16.

YAMAGUCHI, Y., MANN, D. M. & RUOSLAHTI, E. 1990. Negative regulation of transforming growth factor-beta by the proteoglycan decorin. *Nature,* 346**,** 281-4.

YOKOMIZO, T. & DZIERZAK, E. 2010. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development,* 137**,** 3651-61.

YVERNOGEAU, L., KLAUS, A., MAAS, J., MORIN-POULARD, I., WEIJTS, B., SCHULTE-MERKER, S., BEREZIKOV, E., JUNKER, J. P. & ROBIN, C. 2020. Multispecies RNA tomography reveals regulators of hematopoietic stem cell birth in the embryonic aorta. *Blood*.

ZOELLER, J. J., PIMTONG, W., CORBY, H., GOLDONI, S., IOZZO, A. E., OWENS, R. T., HO, S. Y. & IOZZO, R. V. 2009. A central role for decorin during vertebrate convergent extension. *J Biol Chem,* 284**,** 11728-37.

ZOVEIN, A. C., HOFMANN, J. J., LYNCH, M., FRENCH, W. J., TURLO, K. A., YANG, Y., BECKER, M. S., ZANETTA, L., DEJANA, E., GASSON, J. C., TALLQUIST, M. D. & IRUELA-ARISPE, M. L. 2008. Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell,* 3**,** 625-36.

Figure legends

Fig. 1. The dorsal and ventral aortic tissues exhibit distinct molecular signatures

(A) In a first experiment, E11.5 dorsal and ventral tissues (DT and VT) surrounding the aorta were isolated by laser microdissection (red lines) and utilized for bulk transcriptomics (microarrays). In a second experiment, E11.5 whole AGM regions were dissected (green line) and proceeded for single cell transcriptomics (scRNASeq). Nc: notochord, VT: ventral tissue, DT: dorsal tissue, Ao: Aorta, CV: cardinal vein. Scale bar length corresponds to 100 μm.

(B) GO analysis using DAVID of DEGs comparing DT and VT at E10.5, E11.5 and E12.5. The graph indicates the main GO categories enriched in VT at E11.5.

(C) GSEA showing 4 gene sets significantly enriched in VT at E11.5 (profiles are shifted to the right corresponding to the VT genes). NES: normalized enrichment score, FDR: false discovery rate.

(D) Rank metric score of the 23 first genes from the 'Epithelial_Mesenchymal_Transition' gene set enriched in VT shown on the lower right panel of Fig. 1C.

Fig. 2. Correlation-based network analysis of mesenchyme-enriched genes discloses three major modules, one ventrally and two dorsally located

(A) PCA (using 2423 genes and 15 samples) highlighting the different clusters. The score plot of PC1 vs. PC2 is shown.

(B) Module detection using WGCNA. Modules named by colors correspond to squares along the diagonal of the topological overlap matrix plot shown below the clustering tree. Modules are defined in a first step according to the dynamic tree cut method, and in a second step according to the merged dynamic one. DT and VT modules are highlighted.

(C) Module eigenvectors corresponding to the DT to VT (PC1 trait) and to the EC to HC contrasts (PC2 trait). The vectors corresponding to the major DT and VT modules ('turquoise' and 'blue', respectively) are in thicker lines.

(D) Significant GO categories corresponding to genes belonging to the major VT blue module.

(E) Connectivity plot of the 119 most connected genes (out of a total of 295) belonging to the VT blue module.

(F) Connectivity bar plot of the top-30 genes of the VT blue module.

Fig. 3. Single cell analysis (scRNAseq) discloses a large and heterogeneous mesenchymal cell population

(A) Louvain clusters. Left panel: UMAP embedding. Right panels: cluster identities. DEG: differentially expressed gene.

(B) UMAP expression pattern of genes defining the mesenchymal cell population.

(C) UMAP expression of the first-ranked differentially expressed gene defining each of the 8 major cell clusters of the mesenchymal population.

(D) Distribution of the genes of the VT blue WGCNA module among the different clusters. Extreme left panel: projection of the VT module genes onto the Louvain clusters; most genes project onto the C6 and C7 clusters. Other three panels: example of UMAP expression patterns of genes belonging to that module.

Fig. 4. Gene set scoring and information-based network analysis indicate the presence of a ventrally located neuro-mesenchymal population

(A) Gene set scoring. Extreme left panel: UMAP expression pattern of the genes of the VT blue WGCNA module. Middle panels: UMAP expression patterns of the genes belonging to the neuronal and mesenchymal subsets extracted from that module. Extreme right panel: Superimposition of the genes of the neuronal and mesenchymal subsets; the cells expressing both neuronal and mesenchymal genes are represented in purple.

(B) UMAP expression pattern of individual ECM genes belonging to the VT blue mesenchymal subset.

(C) Information-based network. Left panel: Miic plot of the 89 central core genes of the VT blue module. Right panel: UMAP expression patterns for *Ddc* (encircled on the miic plot) and of three of its connected genes.

(D) Analysis of protein-protein interactions using STRING. Direct STRING interactions are indicated by continuous black thick lines and indirect ones by black dashed thick lines. The thin purple lines represent the remaining miic interactions not found using STRING. The table indicates the STRING interactions and the intermediate nodes if any.

Fig. 5. Morphological studies confirm the existence of a subaortic neuro-mesenchymal population

(A) Immunodetection of Dcn, Tnc and CD31 on E11.5 mouse embryo sections.

(B) Immunodetection of Dcn, Tnc and CD105 on day 30 human embryo sections.

(C) Immunodetection of Ntrk3, CD31, Th, Tuj1, Iba1 and Dcn on E11.5 mouse embryo sections.

(D) Transmission electron microscopy of E11.5 subaortic tissue. The image on the right is a magnification of the square indicated on the left panel.

Nc: notochord, Ao: dorsal aorta. Scale bars: Fig. 5A: 100 μm, Fig. 5B: 50 μm, Fig. 5C: 100 μm, Fig. 5D left: 10 μm and Fig. 5D right: 200 nm.

Fig. 6. The E11.5 subaortic neuro-mesenchyme does not derive from NCCs.

(A) UMAP expression pattern of NCC markers *Sox10, Phox2b, Phox2a* and *Th.*

(B) Immunodetection of Dcn on E11.5 sections of HtPA-Cre; R26R-Tomato mouse embryos. Panels from right to left: Dapi, Tomato, Dcn, merge.

(C) Immunodetection of Tnc on E11.5 sections of HtPA-Cre; R26R-Tomato mouse embryos.

(D) Immunodetection for Ntrk3 on E11.5 sections of HtPA-Cre; R26R-Tomato mouse embryos.

Nc: notochord, Ao: dorsal aorta, WD: wolffian duct, C: coelom. Scale bars are 100 μm.

Fig. 7. Loss-of-function experiments in the zebrafish confirm the essential role for HSPC development of the characteristic ECM component Decorin

(A) WISH for *runx1* in 30 hpf *dcn* morphants (MO^{START}) and their respective uninjected controls. Right panel indicates the numbers of $runx1^{+}$ cells in the dorsal aorta of control and *dcn* morphant embryos.

(B) WISH for *cmyb* in 36 hpf *dcn* morphants (MO^{START}) and their respective uninjected controls. Right panel indicates the numbers of $cmyb^{+}$ cells in the dorsal aorta of control and *dcn* morphant embryos.

(C) WISH for *runx1* in 30 hpf $dcn^{+/+}$ *and* $dcn^{-/-}$ embryos. Right panel indicates the numbers of $runx1^{+}$ cells in the dorsal aorta of $dcn^{+/+}$ *and* $dcn^{-/-}$ embryos.

(D) WISH for *cmyb* in 36 hpf $dcn^{+/+}$ *and* $dcn^{-/-}$ embryos. Right panel indicates the numbers of $cmyb^{+}$ cells in the dorsal aorta of $dcn^{+/+}$ *and* $dcn^{-/-}$ embryos.

**** indicates p value < 0.0001 using either unpaired t-test and Mann and Whitney test. 24 to 57 embryos were analyzed for each condition from 1 to 3 independent experiments. Arrowheads indicate HSPCs.

Fig. 1

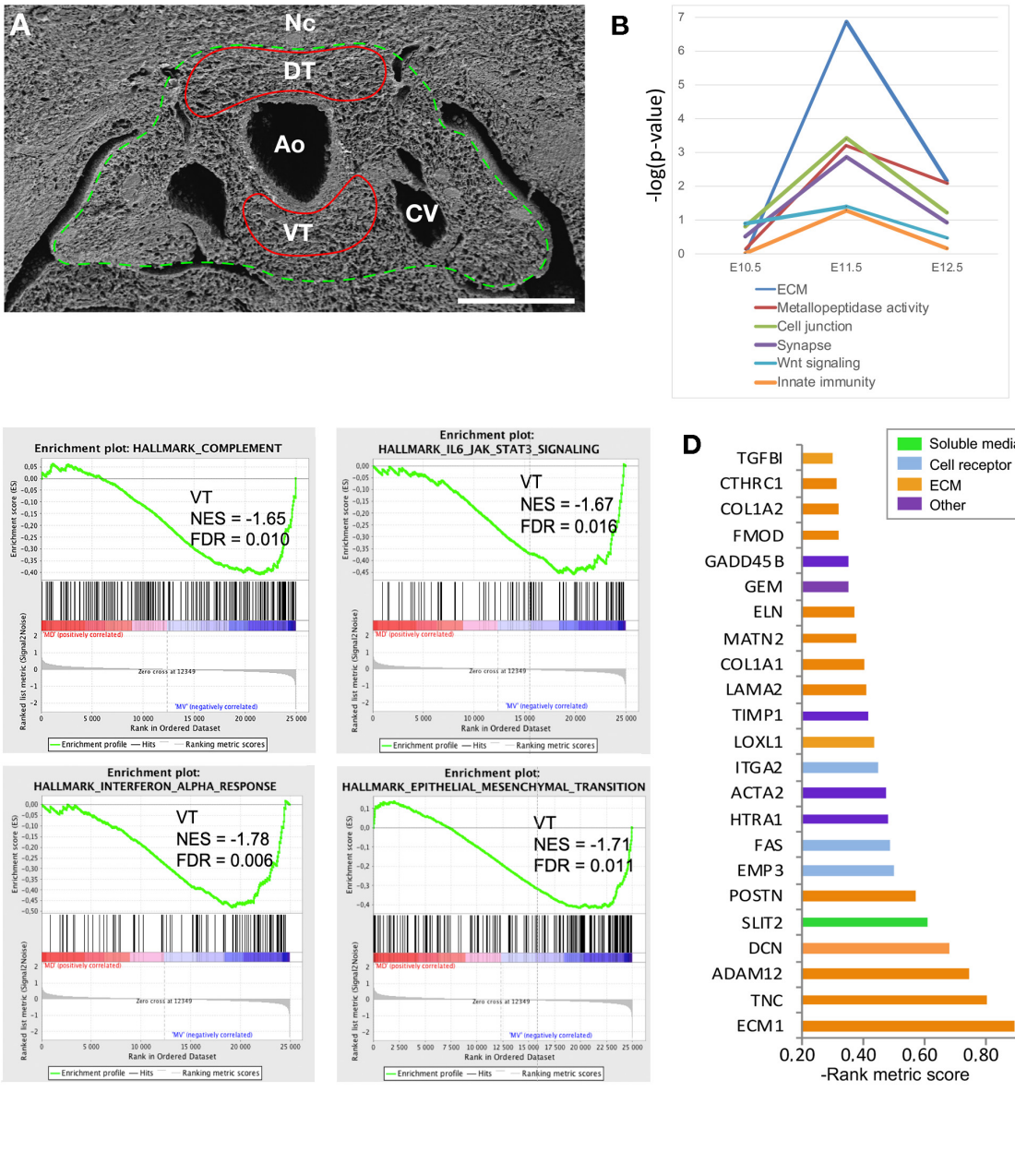
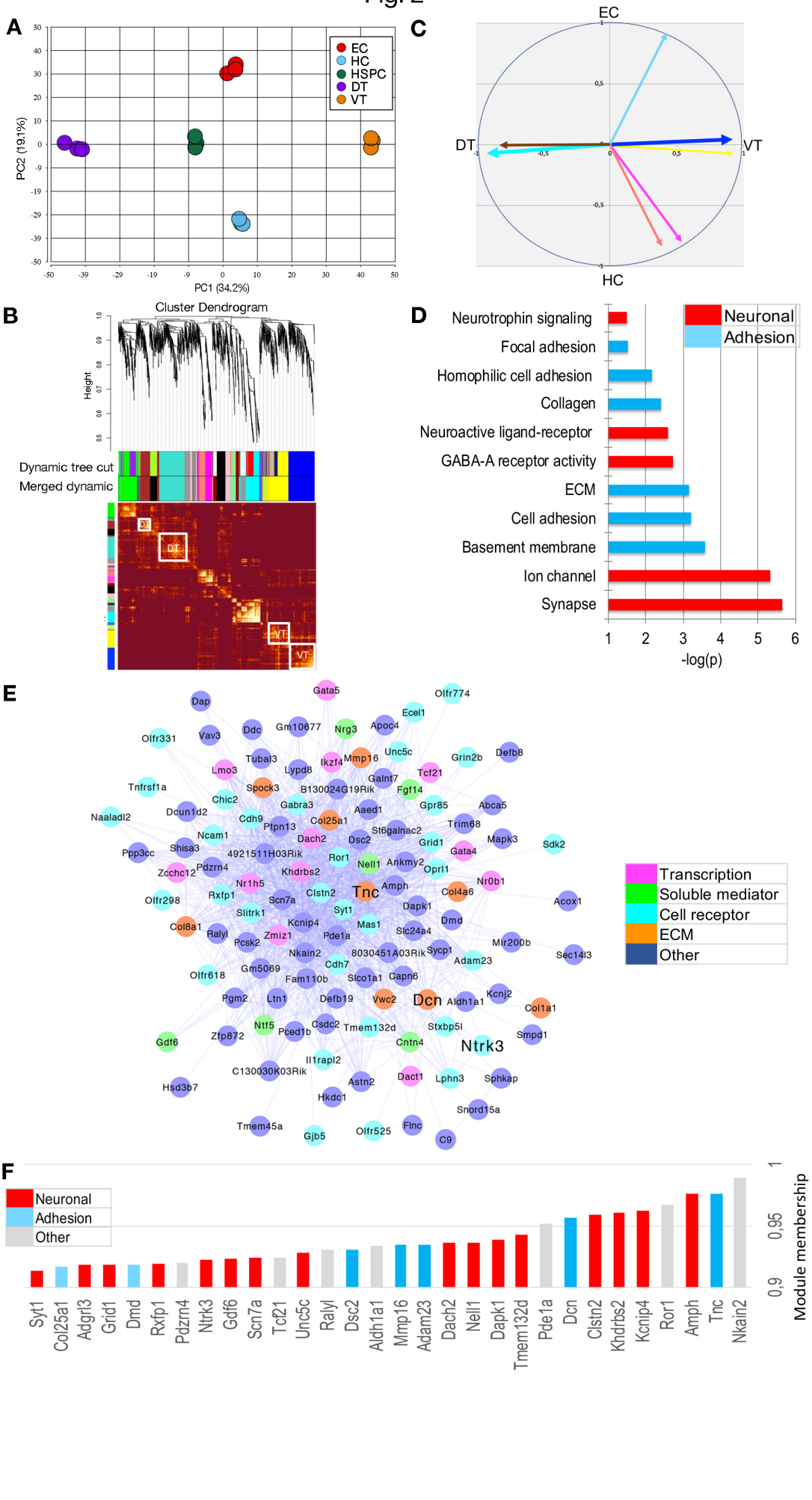
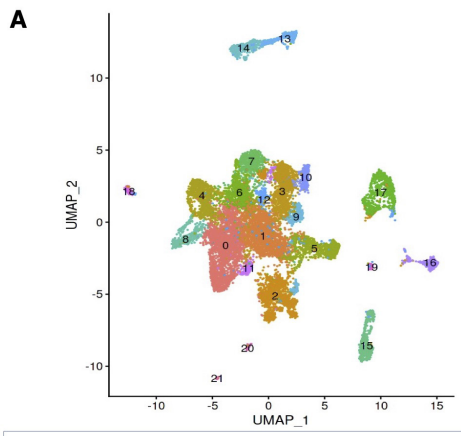

A
Nc
DT
Ao
CV
VT
B
-log(p-value)
E10.5
E11.5
E12.5
ECM
Metallopeptidase activity
Cell junction
Synapse
Wnt signaling
Innate immunity
C
Enrichment plot: HALLMARK_COMPLEMENT
VT
NES = -1.65
FDR = 0.010
Enrichment plot: HALLMARK_IL6_JAK_STAT3_SIGNALING
VT
NES = -1.67
FDR = 0.016
Enrichment plot: HALLMARK_INTERFERON_ALPHA_RESPONSE
VT
NES = -1.78
FDR = 0.006
Enrichment plot: HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
VT
NES = -1.71
FDR = 0.011
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores
D
Soluble mediator
Cell receptor
ECM
Other
TGFBI
CTHRC1
COL1A2
FMOD
GADD45B
GEM
ELN
MATN2
COL1A1
LAMA2
TIMP1
LOXL1
ITGA2
ACTA2
HTRA1
FAS
EMP3
POSTN
SLIT2
DCN
ADAM12
TNC
ECM1
-Rank metric score

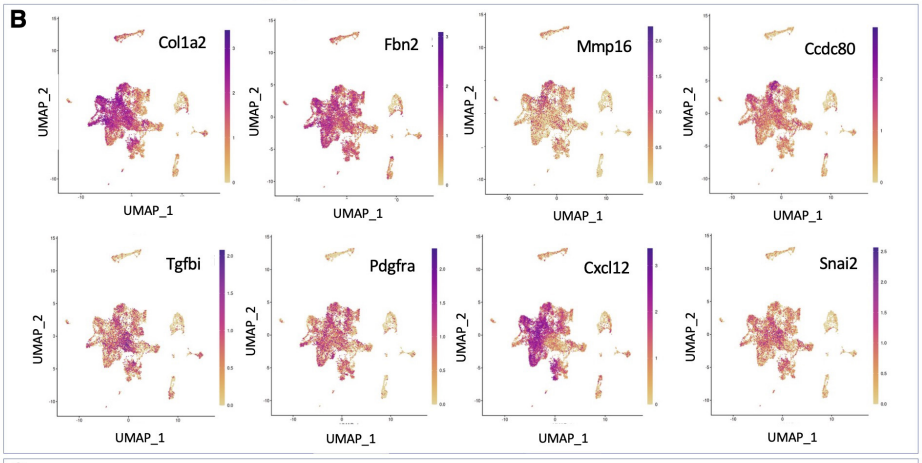
Fig. 2

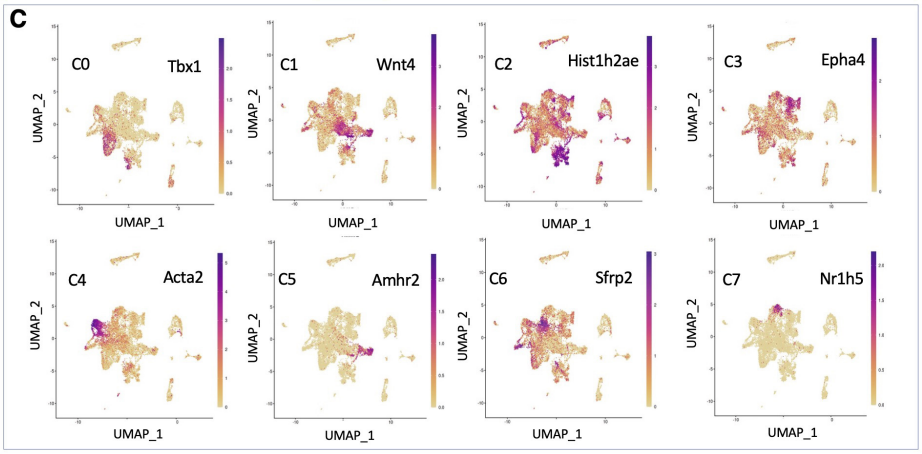

A
EC
HC
HSPC
DT
VT
PC2 (19.1%)
PC1 (34.2%)
B
Cluster Dendrogram
Height
Dynamic tree cut
Merged dynamic
C
EC
DT
VT
HC
D
Neurotrophin signaling
Focal adhesion
Homophilic cell adhesion
Collagen
Neuroactive ligand-receptor
GABA-A receptor activity
ECM
Cell adhesion
Basement membrane
Ion channel
Synapse
Neuronal
Adhesion
-log(p)
E
Tnc
Dcn
Ntrk3
Transcription
Soluble mediator
Cell receptor
ECM
Other
F
Neuronal
Adhesion
Other
Module membership
Syt1
Col25a1
Adgrl3
Grid1
Dmd
Rxfp1
Pdzrn4
Ntrk3
Gdf6
Scn7a
Tcf21
Unc5c
Ralyl
Dsc2
Aldh1a1
Mmp16
Adam23
Dach2
Nell1
Dapk1
Tmem132d
Pde1a
Dcn
Clstn2
Khdrbs2
Kcnip4
Ror1
Amph
Tnc
Nkain2

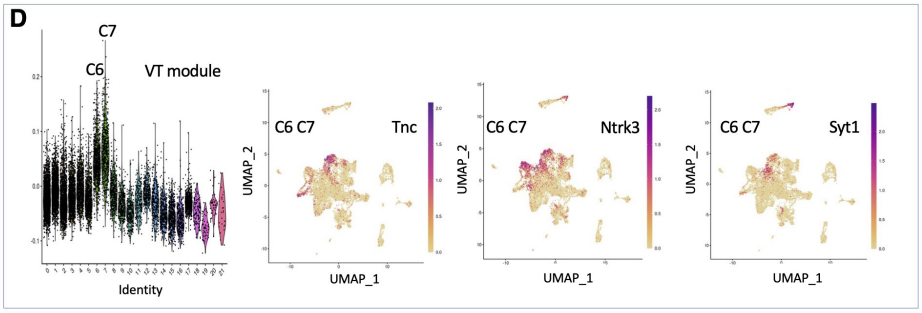


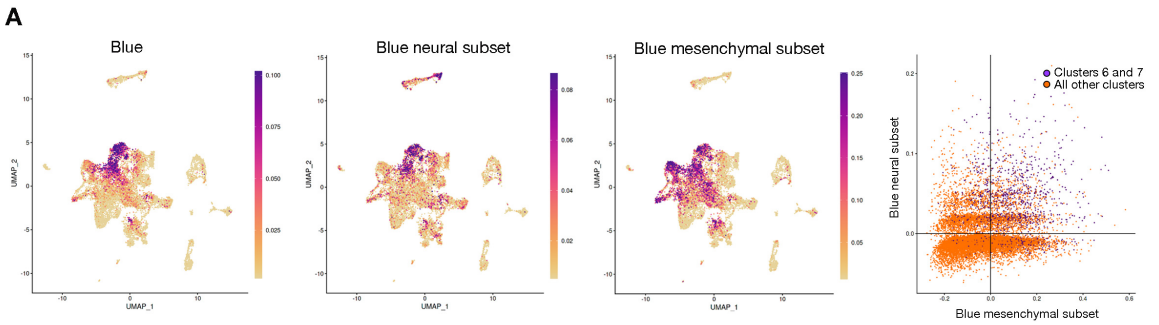
Mesenchymal			
Cluster	Cell number	Representative DEG	Identity
C0	2502	Tbx1, Col2a1	Chondrocytic (Ch)
C1	2103	Wnt4, Wt1	Urogenital (UG)
C2	1420	Hist1h2ae, Fbxo5	Cell cycle
C3	1211	Epha4, Hoxd10	mix
C4	995	Acta2, Cnn1	Smooth muscle (SM)
C5	960	Amhr2, Podxl	Urogenital (UG)
C6	932	Sfrp2, Shisa3	Neuro-mesenchyme (N-M)
C7	626	Nr1h5, Tnc	Neuro-mesenchyme (N-M)
C8	506	Pitx2, Foxf1	Mesenchyme
C9	384	Six2, Uncx	Development
C10	325	Six1, Alx3	Development
C11	221	Pitx1, Meox1	Development
C12	331	*NS*	-

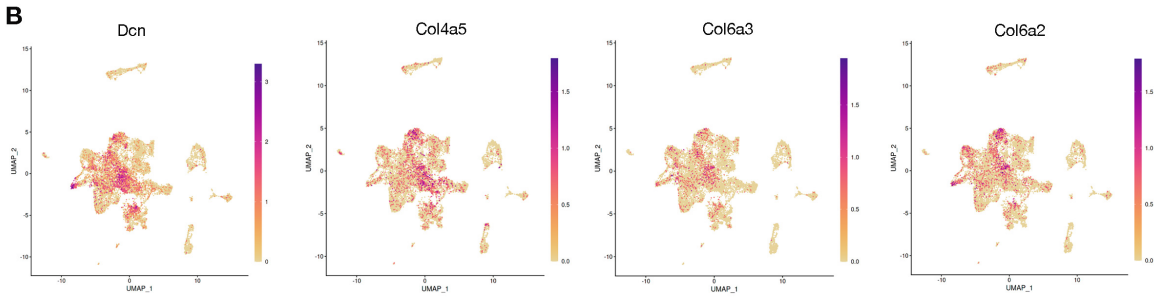
Other			
Cluster	Cell number	Representative DEG	Identity
C13	343	Nefl, Tubb3	Neurons (N)
C14	433	Foxd3, Sox10	Neural crest cells (NCC)
C15	526	Kdr, Cdh5	Endothelial cells
C16	322	Lyz2, Ccl3	Innate immunity
C17	761	Hba-a2, Gpya	Red cells
C18	129	Lhx1, Hnf1b	Kidney, liver
C19	54	Dppa3, Pou5f1	Germinal cells
C20	44	Ttn, Myod1	Muscle
C21	26	Ppbp, Mpl	Hemostasis

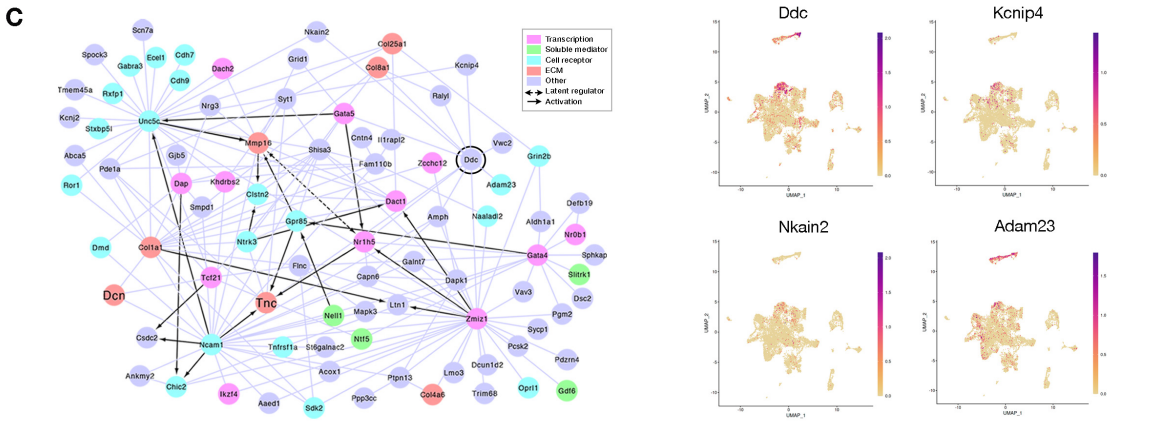




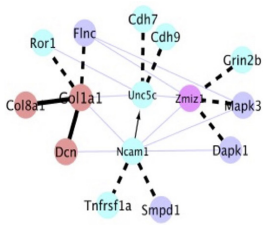








D



Node 1	Node 2	Intermediate
Col1a1	Col8a1	No
Col1a1	Dcn	No
Cdh7	Cdh9	via Unc5c
Dapk1	Grin2b	via Zmiz
Dapk1	Unc5c	via Zmiz
Grin2b	Mapk3	via Zmiz
Smpd1	Tnfrsf1a	via Ncam1
Flnc	Ror1	via Col1a1

Fig. 5

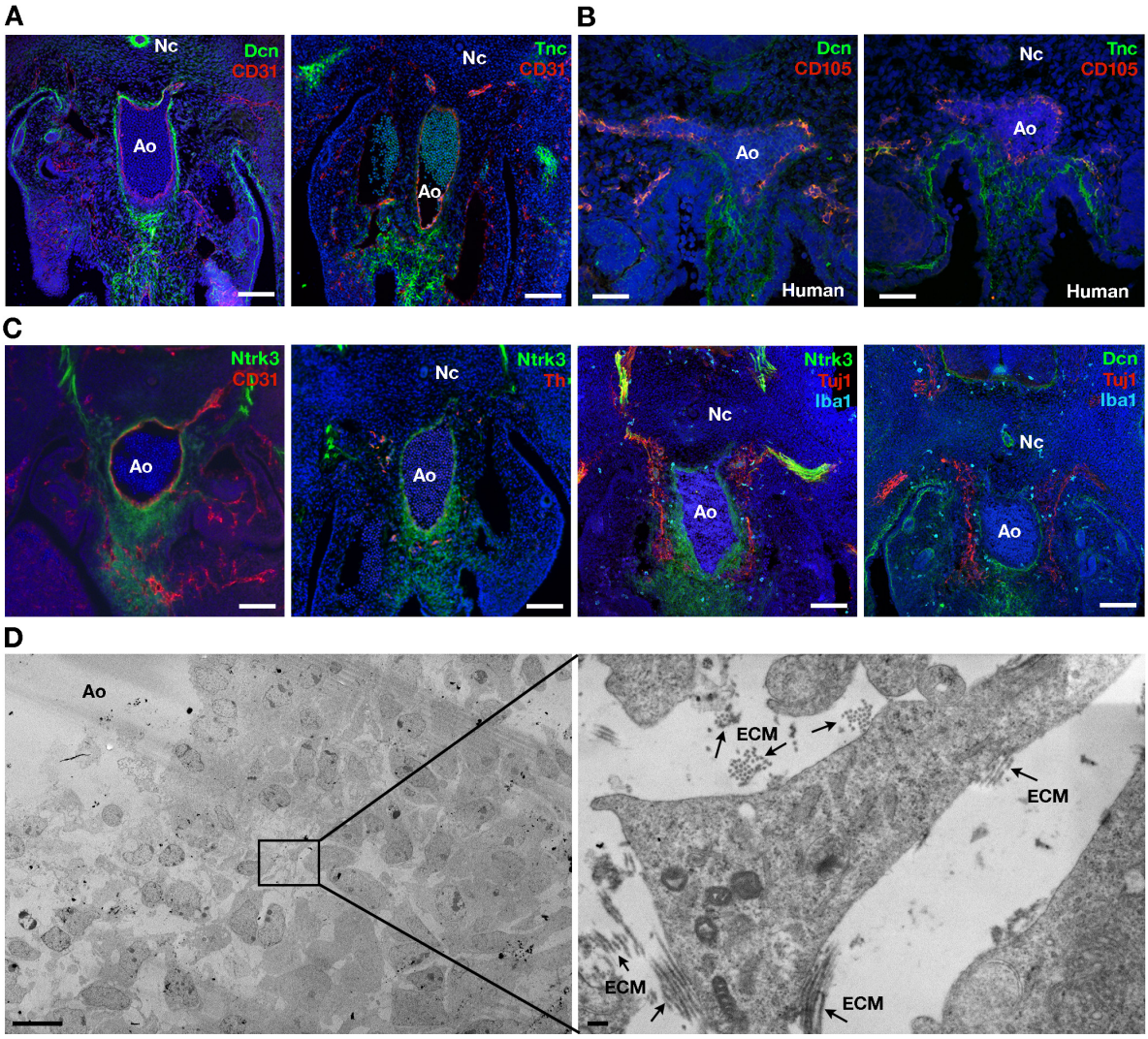

A
Nc
Dcn
CD31
Ao
Nc
Tnc
CD31
Ao
B
Dcn
CD105
Ao
Human
Nc
Tnc
CD105
Ao
Human
C
Ntrk3
CD31
Ao
Nc
Ntrk3
Th
Ao
Nc
Ntrk3
Tuj1
Iba1
Ao
Nc
Dcn
Tuj1
Iba1
Ao
D
Ao
ECM
ECM
ECM
ECM

Fig. 6

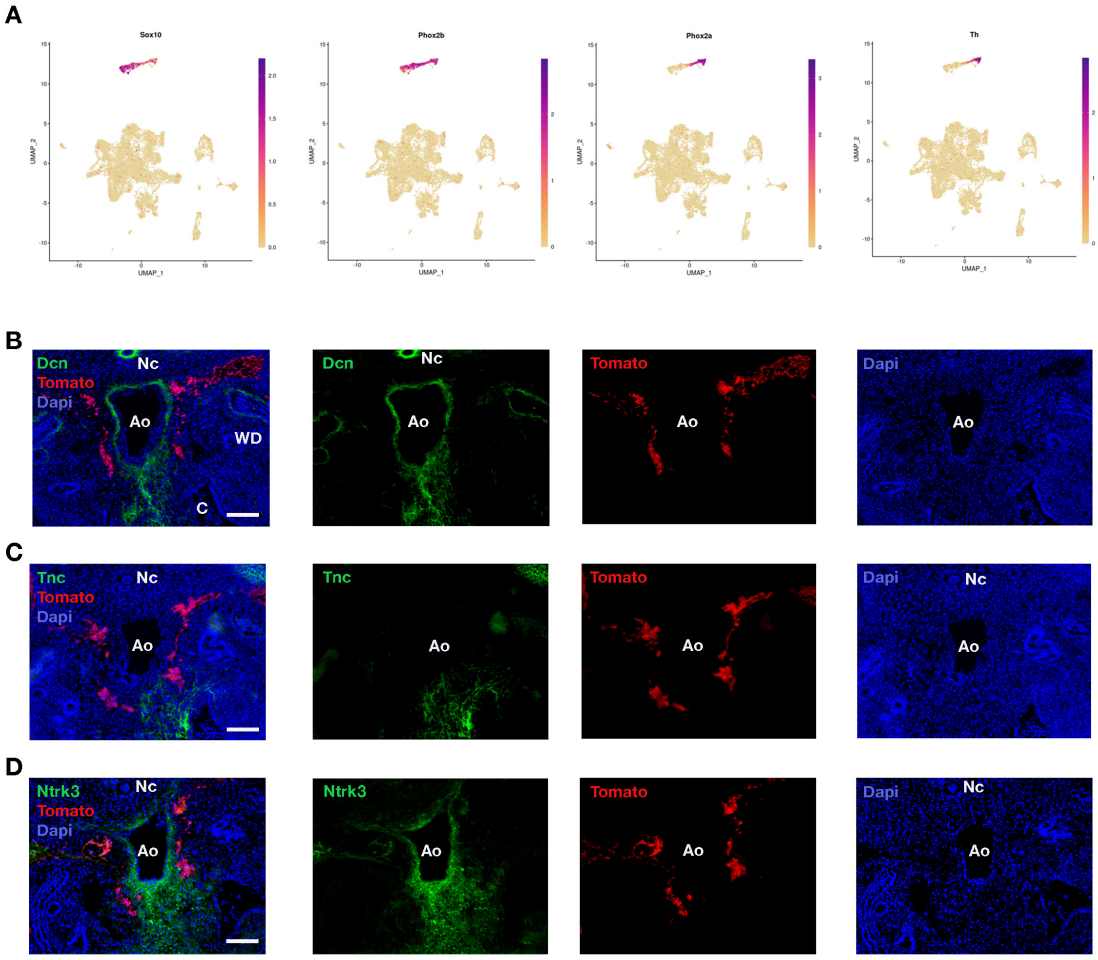
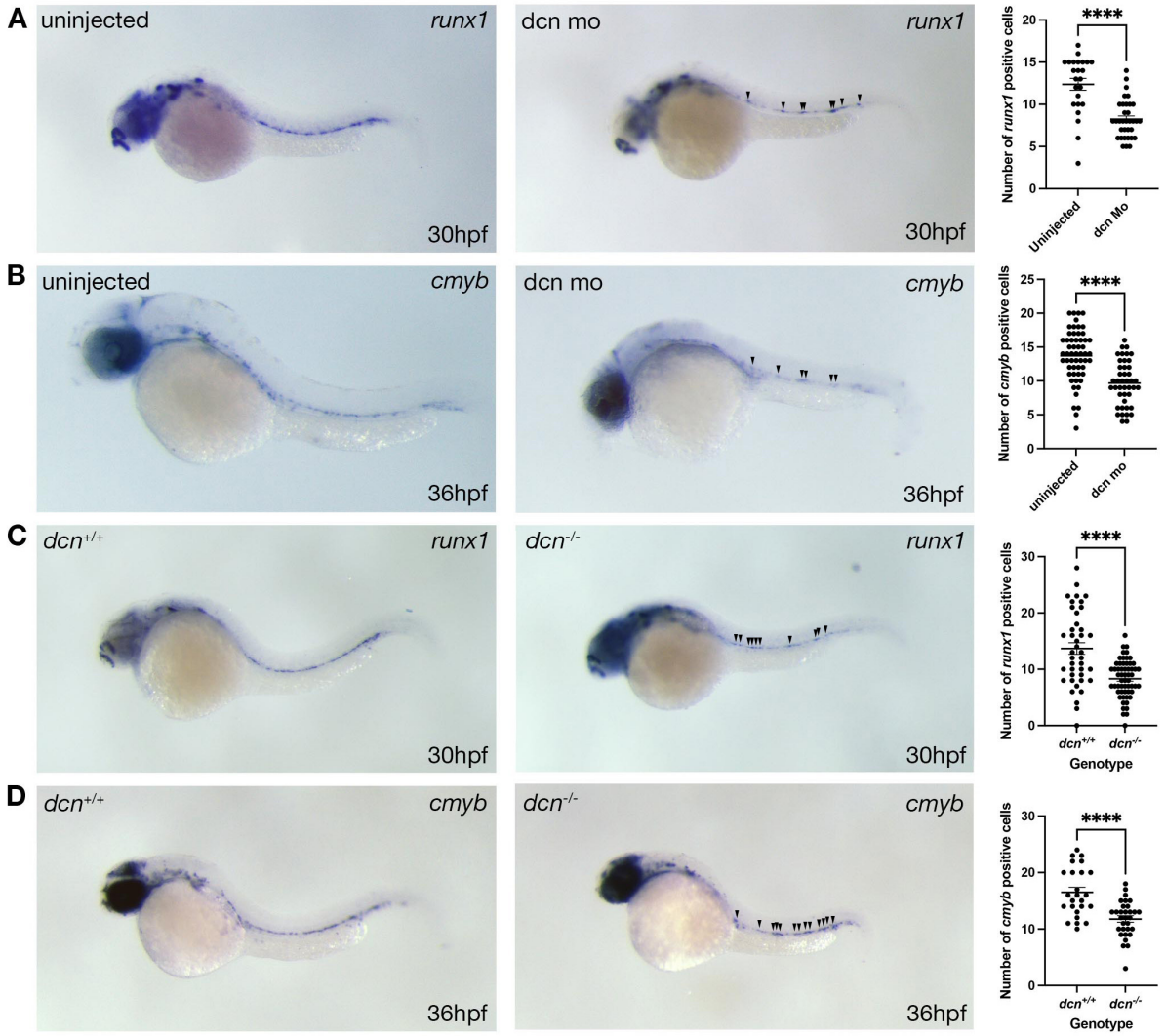

A
Sox10
Phox2b
Phox2a
Th
UMAP_1
UMAP_2
B
Dcn
Tomato
Dapi
Nc
Ao
WD
C
Tnc
D
Ntrk3

Fig. 7


A
uninjected
runx1
30hpf
dcn mo
runx1
30hpf
****
Number of runx1 positive cells
Uninjected
dcn Mo
B
uninjected
cmyb
36hpf
dcn mo
cmyb
36hpf
****
Number of cmyb positive cells
uninjected
dcn mo
C
dcn+/+
runx1
30hpf
dcn-/-
runx1
30hpf
****
Number of runx1 positive cells
dcn+/+
dcn-/-
Genotype
D
dcn+/+
cmyb
36hpf
dcn-/-
cmyb
36hpf
****
Number of cmyb positive cells
dcn+/+
dcn-/-
Genotype