Single-cell transcriptomic analysis identifies the conversion of zebrafish Etv2-deficient vascular progenitors into skeletal muscle

Brendan Chestnut1,4, Satish Casie Chetty1,4, Andrew L. Koenig1,2,4 & Saulius Sumanas1,3✉

Cell fate decisions involved in vascular and hematopoietic embryonic development are still poorly understood. An ETS transcription factor Etv2 functions as an evolutionarily conserved master regulator of vasculogenesis. Here we report a single-cell transcriptomic analysis of hematovascular development in wild-type and etv2 mutant zebrafish embryos. Distinct transcriptional signatures of different types of hematopoietic and vascular progenitors are identified using an etv2ci32Gt gene trap line, in which the Gal4 transcriptional activator is integrated into the etv2 gene locus. We observe a cell population with a skeletal muscle signature in etv2-deficient embryos. We demonstrate that multiple etv2ci32Gt; UAS:GFP cells differentiate as skeletal muscle cells instead of contributing to vasculature in etv2-deficient embryos. Wnt and FGF signaling promote the differentiation of these putative multipotent etv2 progenitor cells into skeletal muscle cells. We conclude that etv2 actively represses muscle differentiation in vascular progenitors, thus restricting these cells to a vascular endothelial fate.

1 Division of Developmental Biology, Cincinnati Children’s Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH 45229, USA. 2 Center for Cardiovascular Research, Washington University School of Medicine, 660 S. Euclid Ave, St. Louis, MO 63110, USA. 3 Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH 45229, USA. ✉These authors contributed equally: Brendan Chestnut, Satish Casie Chetty, Andrew L. Koenig. Email: saulius.sumanas@cchmc.org

https://doi.org/10.1038/s41467-020-16515-y

OPEN
During embryonic development, the lateral plate mesoderm (LPM) gives rise to multiple different cell lineages which include vascular endothelial cells, hematopoietic lineages and cardiomyocytes\(^1\). The specification of these lineages occurs in parallel to the specification of adjacent skeletal muscle progenitors which are thought to originate in the somites from different progenitors than the LPM lineages\(^3\). Although many factors involved in different steps of LPM lineage specification and differentiation have been identified, the entire timeline, transitional steps, and changes in the global transcriptional program that occur during hematopoietic and cardiovascular differentiation in vivo are still poorly understood.

While it is challenging to study early cardiovascular development in mammalian embryos, the zebrafish has emerged as an advantageous model system to study early cell fate decisions during embryogenesis. The signaling pathways and transcriptional programs that regulate the specification and differentiation of the LPM lineages are highly conserved between zebrafish and other vertebrates\(^4\).

We and others have previously identified an ETS transcription factor Etv2/Etsrp which is one of the earliest markers of vascular and hematopoietic progenitor cells and functions as a key regulator of vascular and hematopoietic development in multiple vertebrates including mouse and zebrafish\(^3\). In zebrafish embryos, etv2 is expressed in vascular endothelial progenitor cells, as well as early myeloid and erythropoid progenitors, and its expression is downregulated after cells undergo hematopoietic and vascular differentiation\(^5,6\). In the absence of Etv2 function, vascular endothelial and myeloid progenitors fail to differentiate. While some of them undergo apoptosis, others can acquire alternative cell fates and differentiate into cardiomyocytes, demonstrating fate flexibility of early progenitors\(^4\).

The relatively recent emergence of highly efficient and high-throughput single-cell transcriptomic technologies has facilitated extensive probing of cellular diversity and complex cell differentiation pathways both in vitro and in vivo. In recent years, several studies have been performed to delineate the transcriptional diversity of vascular cell types, and to uncover lineage commitment trajectories during cardiovascular development\(^11\). However, the fate decisions of LPM-derived cells are still poorly understood.

Here, we report single-cell transcriptomic profiling of etv2-expressing cells in wild-type and etv2-deficient zebrafish embryos during early stages of vascular development. We identify the transcriptional profiles and novel transitional states of cells during different steps of vascular and hematopoietic differentiation. Our results show that in the absence of Etv2 function, vascular progenitors can acquire a skeletal muscle fate, arguing that Etv2 function is required to actively repress alternative cell fates in multipotent mesodermal progenitors. These findings will be important in understanding the ontology of different cardiovascular and hematopoietic lineages and will help in designing more efficient in vitro and in vivo cell differentiation strategies to generate different types of progenitors for therapeutic purposes.

**Results**

**Single-cell RNA-seq of vascular and blood progenitors.** To analyze the diversity of hematovascular progenitors, we performed single-cell RNA-seq of GFP-positive cells sorted by fluorescence-activated cell sorting (FACS) from heterozygous and homozygous etv2\(^{12G7}\) UAS:GFP zebrafish embryos at the 20-somite stage. This reporter line, recently generated by GRISPR mediated homology-independent repair, has an insertion of the gal4 reporter within the etv2 coding sequence\(^14\). As described previously, heterozygous etv2\(^{12G7}\) embryos recapitulate the endogenous expression pattern of etv2 in vascular endothelial progenitors and differentiated vascular endothelial cells, while homozygous embryos show profound defects in vascular development due to the interruption of the etv2 coding sequence\(^14\) (Supplementary Fig. 1).

**Transcriptomes of 2049 and 588 cells were obtained from heterozygous and homozygous etv2\(^{12G7}\) embryos, respectively, using the Chromium system (10× Genomics) which employs a microdroplet technology to isolate individual cells, followed by the next-generation sequencing. The relative frequency of GFP+ cells out of the total number of cells was similar in heterozygous and homozygous embryos (1.89% and 1.98%, respectively). Transcriptomes from heterozygous and homozygous embryos were pooled and clustered using Seurat\(^5\), resulting in 12 distinct cell clusters which were visualized using the t-distributed stochastic neighbor embedding (t-SNE) approach\(^16\) (Fig. 1a–d). We subsequently assigned cell identities based on marker genes which were significantly enriched in each cluster (Supplementary Table 1, Supplementary Data 1 and 2). Two different clusters (#2 and #3) corresponded to vascular endothelial cells and were thus labeled as EC1 and EC2. The EC1 cluster showed expression of multiple known vascular endothelial markers, including cdh5, cdn3b, egfl7, sox7, ecscr and others, while the top genes expressed in EC2 cells included crip2, tpm4a, flia, and erg, all known to label vascular endothelial cells (Supplementary Table 1 shows the top marker genes for each cluster, while t-SNE and violin plots for selected marker genes are shown in Fig. 1c and Supplementary Figs. 2 and 3). There was a large overlap in marker expression between the EC1 and EC2 groups. It is currently unclear whether EC1 and EC2 cells represent distinct cellular identities or different stages of endothelial differentiation. Cluster #4 corresponded to the vascular endothelial progenitor group (EPC) which showed enriched expression of tail1, lmo2, etv2, tmem88a, egfl7 (Fig. 1c, d, Supplementary Figs. 2 and 4, Supplementary Table 1). Although some of these genes (tail1, lmo2, etv2) are known to label both vascular and hematopoietic progenitors\(^3,17,18\), other markers specific to this group (egfl7, sox7, flia1b) label vascular and not hematopoietic cells\(^19\), arguing that this population corresponds to vascular endothelial progenitors. Two groups of cells with a strictly hematopoietic gene signature were identified. Cluster #7 showed specific expression of cebpβ, spi1b, cebpα, ccr3.2, all known to be specifically enriched in macrophages (Fig. 1c, d, Supplementary Figs. 2 and 5, Supplementary Table 1, Supplementary Data 1). Cluster #11 showed expression of multiple hemoglobin genes, including hbb3, hbe1.3, hbae1.3, as well as klf17, bfl and other genes that are specific to red blood cells (Fig. 1d, Supplementary Figs. 2 and 5, Supplementary Table 1, Supplementary Data 1). Although etv2 in zebrafish does not show significant expression in zebrafish blood cells, etv2\(^{12G7}\) expression has been previously observed in myeloid and erythroid cells\(^22\), likely due to the expression of etv2 in hematopoietic progenitors, which becomes downregulated as they differentiate. Cluster #10 had very few significantly enriched genes, which included a novel protein sick211–11n16.3, and genes tubb2b and hmg2b which are likely to be ubiquitously expressed. Apoptosis and cell cycle regulators pmapi1/noxa and cong\(^23\) also showed relatively strong expression in this cell population (Fig. 1d, Supplementary Fig. 6, Supplementary Table 1, Supplementary Data 1), suggesting that this group is likely to include pro-apoptotic cells, which were observed in etv2\(^{12G7}\) embryos. Previous studies have established that the loss of etv2 function results in increased apoptosis of vascular or hematopoietic cells\(^21,23\). Cluster #12 showed strong expression of hoxa11b, hoxb7a, cdx4, aopb, which are all expressed in the posterior mesenchyme either in zebrafish or mouse embryos\(^26\–28\) (Fig. 1c, d, Supplementary Figs. 2 and 6, Supplementary Table 1, Supplementary Data 1). The tailbud progenitors have been shown to include multipotent cells which can give rise to endothelial, somitic and neural lineages\(^29\). Therefore this group may...
include early stage endothelial progenitors located in the tailbud. Indeed, many cells from the tailbud population exhibit noticeable expression of EPC markers (Fig. 1d, Supplementary Fig. 4). Another intriguing population of etv2+ cells (cluster #6) showed expression of BMP response genes id3 and id11, as well as twist1a, ptx3, and prrx1a (Fig. 1c, d, Supplementary Figs. 2 and 4, Supplementary Table 1, Supplementary Data 1). Our subsequent analysis (see below) demonstrated that all analyzed genes exhibited expression in the LPM area; therefore we annotated this cluster as LPM cells. Additional cell populations identified in our analysis included cells with neural, epidermal, cardiomyocyte, and skeletal muscle-specific gene signatures (Fig. 1c, d, Supplementary Figs. 2, 7, and 8, Supplementary Table 1, Supplementary Data 1). Observation of GFP-positive cardiomyocyte and skeletal muscle populations in both etv2<sup>C32Gf</sup> heterozygous and homozygous embryos was somewhat unexpected. Cardiomyocytes in wild-type embryos do not
express etv2, while our previous study has shown that etv2-deficient endothelial progenitors can differentiate as cardiomyocytes. The presence of GFP-positive cardiomyocytes in etv2ci32Gt homozygous embryos could be explained by the loss of one functional etv2 allele, which resulted in some etv2+ cells differentiating as cardiomyocytes. However, differentiation of hematovascular progenitors into skeletal muscle cells has not been previously reported.

scRNA seq analysis of etv2 loss-of-function embryos. We then analyzed the cell clusters present in etv2ci32Gt homozygous embryos, which exhibit severe defects in vascular differentiation. Both EC1 and EC2 populations, EPCs and macrophages were nearly completely absent in etv2ci32Gt homozygous embryos (Fig. 1b, e), consistent with the known requirement of etv2 in the formation of these lineages. In contrast, percentage of RBC cells was slightly but significantly increased in etv2ci32Gt homozygous embryos (Fig. 1b, e). To confirm this result, we analyzed expression of RBC-specific markers using data from the recently reported global transcriptome study of etv2ci32Gt heterozygous and homozygous embryos. Indeed, expression of multiple RBC-specific genes, including hbaa1, vwf, hbae3 and others was significantly upregulated in etv2ci32Gt homozygous embryos (Supplementary Table 2). Furthermore, differential expression analysis of scRNA-seq data between RBC cell populations in heterozygous and homozygous embryos revealed an upregulation of several RBC-specific genes in homozygous embryos, including alas2, hbae3, hbae1.3, and hbbe1.1, while several vascular endothelial specific genes including tmnk, escr and an EPC specific gene lmo2 were downregulated (Supplementary Data 3). For additional confirmation, we knocked down etv2 function using previously validated MO in gata1:dsRed transgenic embryos which express fluorescent reporter in RBCs. Etv2 knockdown embryos were morphologically normal and showed previously reported vascular defects that were similar to the etv2 mutant phenotype. Subsequently, cells were disaggregated at 23 hpf and analyzed by FACS sorting. A statistically significant increase in the number of dsRed-positive cells was observed in etv2 knockdown embryos (Supplementary Fig. 9). This suggests that the erythropoiesis pathway is upregulated in etv2-deficient embryos, although exact mechanism requires further investigation.

Cells in the homozygous embryos showed an increase in the LPM and myocyte populations, and a loss of EC populations. We hypothesized that the LPM group represented EPCs at an earlier stage of differentiation, and that the myocyte cells were an alternative cell fate for etv2+ progenitors when they could not differentiate into the vascular endothelial lineage.

Identification of an endocardial subcluster. In an attempt to identify additional heterogeneity within the endothelial cells, which may be missed in the global cell clustering approach, we performed subclustering of endothelial cell populations EC1 and EC2 identified in etv2ci32Gt+/− embryos using Seurat (see Methods). The subclustering of EC1 resulted in the identification of two subpopulations of cells (Fig. 2a–c), while the subclustering of EC2 did not yield meaningful subpopulations. One of the two EC1 sub-clusters showed enriched expression of cdh6, gata5, fn1a, and drl, among the other markers (Fig. 2d, e, Supplementary Data 4). Gata5 has been previously known to be enriched in endothelial cells. In situ hybridization (ISH) analysis confirmed that expression of fn1a and drl was also enriched in the endocardial cells and largely absent from other vascular endothelial cells (Fig. 2f, g). In contrast, genes enriched in the remaining EC1 subpopulation (designated as EC1a) which included etv2 and lmo2, were largely absent from endocardial progenitors at the 20-somite stage (Fig. 2h, i). This strongly suggests that the first subcluster represents a population of endocardial progenitors which possess a unique transcriptomic signature even prior to the formation of the heart tube.

Vascular progenitors convert into muscle in etv2 mutants. To confirm if GFP expression was in fact present in skeletal muscle cells, as suggested by scRNA-seq analysis, we analyzed etv2ci32Gt UAS:GFP embryos by confocal microscopy. Indeed, GFP expression was observed in several myocytes in etv2ci32Gt heterozygous embryos (3 ± 2.5 myocytes per embryo, ± refers to standard deviation, n = 8 embryos), while multiple GFP-positive cells (18.5 ± 5.4 per embryo, n = 10) were present within the somitic muscle of etv2ci32Gt homozygous embryos at 25 hpf (Fig. 3a–e). These cells displayed an elongated shape, similar to other skeletal muscle cells, and were positive for expression of skeletal muscle actin actc1b:GFP (Fig. 3f–h), GFP expression in skeletal muscle overlapped with antibody staining for fast muscle but not slow muscle cells (Supplementary Fig. 10). When crossed into the previously generated etv2:mCherry reporter line, these muscle cells were also positive for mCherry expression, indicating that the reporter fluorescence is unlikely to be an artifact of transgene misexpression in unrelated tissues. To confirm that etv2ci32Gt UAS:GFP expression in the skeletal muscle is caused by the loss of etv2 function, we crossed the etv2ci32Gt line with etv2 mutants. The etv2 allele, generated using CRISPR/Cas9 mutagenesis, contains a 13 bp insertion in exon 5 of etv2 (Supplementary Fig. 11).
Double heterozygous etv2<sup>ci32Gt/ci33</sup> embryos displayed a significant increase in GFP+ myocytes compared to etv2<sup>ci32Gt/+</sup> embryos (Fig. 3l–n, Supplementary Fig. 11e–h). In addition, GFP expression in skeletal muscle cells was also observed in etv2<sup>11+/−</sup> mutant embryos crossed into the previously established Tg(-2.3 etv2-GFP) reporter line but not in wild-type Tg(-2.3 etv2-GFP)<sup>14,32</sup> embryos (Fig. 3o). Collectively, these results argue that etv2-expressing progenitors differentiate as skeletal muscle in the absence of Etv2 function.

Fig. 2 Identification of the transcriptional signature of endocardial progenitors. Subclustering of the ECI population in etv2<sup>2;i32Gti−/−</sup> embryos identified a novel endocardial subpopulation (while the remaining endocardial cells are designated as ECIα). a A section of a heatmap showing marker gene expression in endocardial, endothelial (EC1α and EC2), endothelial progenitor (EPC), and cardiomyocyte populations. b A portion of the global t-SNE plot showing the endocardial subpopulation. c Heatmap for subclustering of ECI population showing genes enriched in endocardial and the remaining endocardial (EC1α) cells. d, e t-SNE and violin plots for selected endocardial genes. f–i In situ hybridization analysis at the 20-somite stage for selected endocardial-enriched genes fn1a and dlr (f, g, arrows point to endocardial expression) and endocardial-excluded genes etv2 and lmo2 (bilateral expression is present in cranial endothelial progenitors but is largely absent from the endocardial cluster). Flat-mounted embryos, ventral view; anterior is to the left. The number of embryos displaying the representative phenotype out of the total number of embryos obtained from two replicate experiments is shown.

To investigate the mechanism of the conversion of these etv2-deficient progenitors into skeletal muscle in greater detail, we analyzed the functional role of bHLH transcription factor scl/tal1 which functions downstream of etv2 during vascular development and myelopoiesis<sup>36</sup>. Previous research has demonstrated that the combined knockdown of scl and etv2 results in ectopic myocardial differentiation of etv2+ progenitors in zebrafish<sup>10</sup>, and that scl represses myocardial differentiation in mouse embryos<sup>35</sup>. However, its role in repressing skeletal muscle differentiation has not been investigated. The number of myocytes, positive for etv2<sup>2;i32Gti−/−</sup>; UAS:GFP expression, was significantly increased in the etv2<sup>2;i32Gti−/−</sup> embryos injected with the previously validated scl morpholino (MO)<sup>35</sup>, compared to the control etv2<sup>2;i32Gti−/−</sup> embryos (Fig. 3n, p). Furthermore, these ectopic myocytes were positive for expression of the skeletal muscle actin act1b (Fig. 3q–s), arguing that scl knockdown results in increased skeletal muscle differentiation of etv2<sup>2;i32Gti−/−</sup>; UAS:GFP cells. However, injection of scl mRNA into etv2<sup>2;i32Gti</sup>; UAS:GFP embryos did not have a significant effect on the number of ectopic GFP-positive muscle cells (Supplementary Fig. 12), suggesting that etv2 and scl genes may not function in a simple linear pathway to repress muscle differentiation in etv2+ progenitor cells.

Heat-shock inducible etv2 expression has been demonstrated to convert differentiated skeletal muscle into vascular endothelial cells<sup>34</sup>. To test if etv2 is sufficient to inhibit early specification or differentiation of myocytes, we overexpressed etv2 mRNA in zebrafish embryos. We have previously shown that such overexpression induces strong ectopic expression of vascular endothelial markers<sup>5</sup>. Conversely, expression of the early regulators of muscle differentiation myod, myf5, and myog was greatly inhibited upon etv2 overexpression (Fig. 3i, u).

Vascular endothelial progenitors are known to originate in the LPM region. However, the somitic origin of vascular progenitors has also been suggested<sup>35</sup>. In order to analyze the origin of GFP+ myocytes, we performed time-lapse imaging of etv2<sup>2;i33Gt</sup>; UAS:GFP embryos. The majority of EPCs in heterozygous embryos originated in the LPM region and subsequently migrated to the midline where they coalesced into the DA and the PCV (Fig. 4a-f, Supplementary Fig. 13a–f, Supplementary Movies 1 and 2). Occasional cells failed to migrate to the midline and elongated to form somitic muscle. In contrast, GFP+ cells in homozygous embryos failed to migrate to the midline. Some of these cells underwent apoptosis while others elongated and incorporated into the somitic muscle at approximately 14–18-somite stages (Fig. 4g–m, Supplementary Fig. 13g–l,
Supplementary Movies 3–5). Thus, etv2+/–; UAS:GFP cells that give rise to the skeletal muscle cells in the homozygous embryos appear to be derived from the initial pool of etv2+ cells in the LPM region. To analyze when GFP+ positive cells initiate muscle marker expression, we performed fluorescent ISH for myod expression combined with GFP fluorescence analysis at the 8–10-somite stages in etv2+/–; UAS:GFP embryos. GFP and myod co-expressing cells were positioned at the posterior edge of the somites next to...
GFP+ cells in the LPM (Fig. 5). In wild-type embryos, etv2-expressing vascular progenitors are known to migrate toward the midline and coalesce into the axial vasculature (arrows). Note that some cells remain in the lateral position and elongate into muscle cells (myoc). DA progenitors of the dorsal aorta, ISV intersegmental vessels. Time frames are selected from the Supplementary Movie 1. g–l In etv2c32Gt+/−; UAS:GFP embryos, cells initiate migration (blue arrows) but fail to coalesce into the axial vasculature. Instead, many cells either undergo apoptosis (apo, red arrowheads point to round apoptotic cells) or differentiate into myocytes (white arrows, myoc). Time frames are selected from the Supplementary Movie 3. m Higher magnification view showing differentiation of a GFP+ progenitor cell initially positioned in the LPM into a myocyte (arrowhead points to the same cell). Note that the cell migrates dorsally from the LPM into the somite and then elongates as it undergoes differentiation into the muscle. Time frames are selected from the Supplementary Movie 4. Representitive embryos are shown out of the total of seven heterozygous and four homozygous etv2c32Gt; UAS:GFP embryos that were imaged in two replicate experiments.

**Fig. 4 Time-lapse imaging of cell migration in etv2c32Gt embryos starting at the 9–10-somite stage.** Lateral view is shown, anterior is to the left. a–f In etv2c32Gt+/−; UAS:GFP embryos, bilaterally located vascular and hematopoietic progenitors within the lateral plate mesoderm (LPM, arrowheads) migrate toward the midline and coalesce into the axial vasculature (arrows). Note that some cells remain in the lateral position and elongate into muscle cells (myoc). DA progenitors of the dorsal aorta, ISV intersegmental vessels. Time frames are selected from the Supplementary Movie 1. g–l In etv2c32Gt+/−; UAS:GFP embryos, cells initiate migration (blue arrows) but fail to coalesce into the axial vasculature. Instead, many cells either undergo apoptosis (apo, red arrowheads point to round apoptotic cells) or differentiate into myocytes (white arrows, myoc). Time frames are selected from the Supplementary Movie 3. m Higher magnification view showing differentiation of a GFP+ progenitor cell initially positioned in the LPM into a myocyte (arrowhead points to the same cell). Note that the cell migrates dorsally from the LPM into the somite and then elongates as it undergoes differentiation into the muscle. Time frames are selected from the Supplementary Movie 4. Representitive embryos are shown out of the total of seven heterozygous and four homozygous etv2c32Gt; UAS:GFP embryos that were imaged in two replicate experiments.

Wnt and FGF pathways are required for ectopic muscle. Wnt and FGF signaling pathways have been previously implicated in somitic differentiation and myogenesis. We tested if these pathways were involved in the differentiation of etv2c32Gt; UAS:GFP+ progenitors into ectopic muscle cells. To inhibit Wnt signaling we induced expression of Wnt inhibitor dkk1 at the 8-somite stage. As it has been previously demonstrated, Wnt inhibition at this stage resulted in reduced myod expression in the tailbud region, while myod expression in differentiated somitic cells was largely unaffected (Supplementary Fig. 14a, b). Heat-shock inducible expression of Wnt inhibitor dkk1 greatly reduced the number of etv2+ myocytes observed in heterozygous etv2c32Gt; UAS:GFP embryos (Fig. 6a–c). Dkk1 overexpression did not affect the total number of cells positive for the UAS reporter expression in etv2c32Gt embryos (Supplementary Fig. 14c). This suggests that Wnt signaling promotes muscle differentiation in multipotent etv2+ progenitors.

To investigate the role of FGF signaling, we treated etv2c32Gt−/−; UAS402, a chemical inhibitor of FGF signaling. 38 SU5402-treated embryos showed a significant decrease in the number of ectopic GFP+ muscle cells (Fig. 6d–f). To confirm the requirement of FGF signaling in the ectopic muscle differentiation, etv2c32Gt+/−; UASNTR-mCherry carriers were crossed with a heat-shock inducible lshp70dnFGFR1 line, and the expression of
dnFGFR1 was induced at the 8-somite stage. A significant reduction in the number of ectopic muscle cells was observed in FGF-inhibited embryos while no significant defects in vascular development were apparent (Fig. 6g–i). The overall number of UAS reporter-positive cells was not affected in FGF-inhibited etv2+/- embryos (Supplementary Fig. 14c). Together, these results suggest that FGF signaling is required for muscle-specific differentiation of multipotent etv2+ progenitors.

**Fig. 5** Fluorescent in situ hybridization for myod expression in etv2ci32Gt+/-; UAS:GFP embryos. In situ hybridization was performed using hybridization chain reaction (HCR) at the 8–10-somite stages. a–c Maximum intensity projection is shown; d–f An area boxed in a was imaged at higher magnification; maximum intensity projection of three confocal slices is shown. Note that most GFP-expressing cells are positioned in the lateral plate mesoderm, while some cells are starting to migrate toward the midline (arrow, d, e, points to a migrating angioblast which does not have myod expression). GFP and myod co-expressing cells (arrowheads, d, e) are apparent at the posterior edge of the somite. The punctate myod expression pattern is due to the nature of the HCR probe. Dorsalateral view, anterior is to the left. The numbers in the lower left corner display the number of embryos showing the phenotype out of the total number of embryos analyzed in two replicate experiments.

**Fig. 6** Wnt and FGF signaling is required for differentiation of etv2ci32Gt; UAS:GFP cells into skeletal muscle. a–c Heat-shock inducible expression of Wnt inhibitor Dkk1 greatly reduces GFP-positive muscle cells in hsp70:dkk1; etv2ci32Gt+/-; UAS:GFP embryos. Heat-shock (HS) was performed at the 8-somite stage, and GFP fluorescence imaged at 24–26 hpf. Sixteen embryos were analyzed in each group in two replicate experiments. d–f etv2ci32Gt-/-; UAS:GFP treated with FGF inhibitor SU5402 starting at the 50% epiboly stage show a greatly reduced number of GFP-positive cells at the 22-somite stage. Nine control DMSO-treated and 15 SU5402-treated embryos were analyzed in two replicate experiments. g–i Heat-shock inducible expression of dnFGFR1 results in a significant reduction of mCherry-positive muscle cells in etv2ci32Gt+/-; UAS:NTR-mCherry; hsp70:dnFGFR1 embryos. Twenty each control and HS hsp70:dnFGFR1 embryos were analyzed in two replicate experiments. Median ± SD values are shown in all graphs, two-tailed Student’s t test was used for statistical analysis.
Identification of putative multipotent LPM progenitors. Based on scRNA-seq analysis, cell cluster #6 was significantly increased in etv2\textsuperscript{-32Gt\textsubscript{-/-}; UAS:GFP} embryos. Marker genes for this population included BMP response genes id\textsubscript{b} and id\textsubscript{1} (Supplementary Table 1, Supplementary Data 1), which have been recently shown to function upstream of etv2 during vascular differentiation of tailbud-derived neuromesodermal progenitors\textsuperscript{40}. Another marker gene for this population, twist1a, has been implicated in developmental angiogenesis\textsuperscript{41}, although its mechanistic role in this process has not been investigated. Other genes enriched in this cluster included transcription factors prrx1a, foxd1, and foxd2, which have been implicated in diverse processes, including epithelial–mesenchymal transition and craniofacial development\textsuperscript{42}. However, their function or relevance to vascular development has yet to be elucidated. We decided to investigate the identity of the cells from cluster #6 in greater detail. Expression of transcription factors twist1a, prrx1a, and foxd2 was present in the LPM region at the 18-somite stage (Fig. 7a–n). Based on ISH analysis of etv2\textsuperscript{-32Gt\textsubscript{-/-}; UAS:GFP} embryos, prrx1a and twist1a expression was absent from the midsline GFP+ cell population which correspond to vascular endothelial cells that are coalescing into axial vessels, while it partially overlapped with GFP expression in the most lateral vascular progenitors. The expression of these genes was not significantly affected in etv2\textsuperscript{-32Gt\textsubscript{-/-}; UAS:GFP} embryos (Fig. 7d–f, j–n). Similar overlap of prrx1a or twist1a expression with GFP fluorescence in the most lateral vascular progenitors was also observed in wild-type Tg(BAC(etv2:GFP) embryos (Supplementary Fig. 15). Because vascular progenitors originate within the LPM prior to their migration to the midsline, such gene expression pattern suggests that this cell cluster corresponds to the LPM cells which may include the earliest vascular progenitors (Fig. 7o, p).

Transcriptional profiling of etv2:GFP cells using Fluidigm. Although the Chromium-based scRNA-seq analysis identified novel cell populations, we were not able to distinguish endothelial subpopulations such as arterial and venous progenitors which are known to separate early during vasculogenesis. We decided to supplement our analysis by sequencing fewer cells at a higher coverage. For this approach, we used a previously established Tg(-2.3 etv2:GFP) line\textsuperscript{31} which shows specific GFP expression in vascular endothelial cells (Fig. 8a–c). Cells from embryos at 16–20-somite stages were dissociated and sorted using the Fluidigm C1 system. Totally, 96-cells were subjected to scRNA-seq analysis. After cell clustering using AltAnalyze software\textsuperscript{44}, 8 distinct cell populations were identified. They included macrophage, erythroid, tailbud progenitor, and nonspecific epidermal cell populations (Fig. 8d–g). Four different vascular endothelial cell populations were identified. The vascular endothelial progenitor population displayed high expression of etv2, lmo2 and tal1, as well as multiple genes that have not been previously associated with EPCs (Fig. 8d–g, Supplementary Datas 5 and 6). Many genes overlapped between EPC populations identified in the Chromium and Fluidigm analyses, suggesting that the same populations of cells were identified using both approaches. Cluster 5 was enriched in venous-specific vascular endothelial genes such as flt4 and stab1, while cluster 6 showed high expression of known arterial-specific markers, including dll4, efnb2a, cdhn5b (Fig. 8g, Supplementary Data 5). Cluster 7 included many genes with known expression in vascular endothelium such as col4a1, ldb2a, or hapln3. Validation by ISH analysis confirmed expected expression patterns of EPC, arterial and venous markers and showed that expression of the cluster 7 marker ldb2a was also primarily enriched in the arterial vasculature (Fig. 8i–l). The biological difference between the clusters 6 and 7 is unclear, and it will require further investigation. Notably, marker genes for cluster 7 (also labeled as EC2) are largely different from the EC2 cluster identified using the Chromium platform.

Expression of venous markers in arterial progenitors. Previous studies have suggested that arterial and venous progenitors originate at distinct spatial regions and at different times\textsuperscript{36}. However, it has been controversial as to when cells acquire distinct arterial and venous identities\textsuperscript{45–47}. We used the expression of well-established arterial and venous markers to calculate an arteriovenous index for each vascular endothelial cell (see “Methods”). While some cells showed a very clear arterial or venous signature, many cells co-expressed both arterial and venous markers (Fig. 8h). Fluorescent ISH analysis for venous dab2 and arterial marker cdhn5b expression confirmed that many arterial progenitors co-express both markers at the 20-somite stage (Fig. 8m–o). In contrast, arterial dab2 expression becomes downregulated by 24 hpf and is restricted to the PCV (Fig. 8p–r). Intriguingly, ventrally located progenitors of the PCV mostly displayed venous marker expression and showed very little, if any, arterial cdhn5b expression (Fig. 8m–o). Thus, arterial progenitors co-express both arterial and venous markers at early stages of vascular development while venous progenitors show largely venous-specific marker expression.

Discussion

In this study we used scRNA-seq analysis to identify transcriptional signatures and investigate transitional states of cell populations during the differentiation of vascular endothelial and hematopoietic lineages. Our results suggest that cells in the LPM are multipotent and that the same progenitors can give rise to hematopoietic, vascular, myocardial, and even skeletal muscle lineages. Wnt and FGF signaling pathways promote skeletal muscle development in etv2-negative multipotent progenitors (Fig. 9). Etv2 overexpression was sufficient to inhibit muscle differentiation and instead promoted hematopoietic fates. Our results argue that in addition to activating vasculogenic and hematopoietic programs, etv2 also actively represses myogenesis, possibly through the transcription factor scl/tal1. Previous work has demonstrated that etv2 can similarly repress myocardial differentiation in both zebrafish and murine embryos.\textsuperscript{8,9} Scl, similar to Etv2, also represses myocardial fates by occupying primed enhancers of myocardial specific genes and preventing their activation by cardiac factors.\textsuperscript{48} It is tempting to speculate that Etv2 also inhibits muscle specification by a similar mechanism, where Scl functions downstream of Etv2 and binds to muscle-specific promoters, thus preventing their activation by Myod. Alternatively, Scl and Etv2 may form a protein complex and/or function in parallel pathways to repress myocardial differentiation. Because scl overexpression failed to repress the formation of ectopic myocytes observed in etv2-deficient embryos, we speculate that scl may not function in a simple linear pathway downstream of etv2 to inhibit myogenesis. While the precise molecular pathways remain to be established, they will likely involve epigenetic mechanisms to repress the transcriptional activation of the myogenic program when the hematopoietic program is activated.

A recent study has demonstrated two opposing activities of FGF and BMP signaling in specifying somitic progenitors and the LPM\textsuperscript{40}. FGF signaling through bHLH factors myf5, myod and msgn1 induced medial fates such as skeletal muscle in tailbud-derived neuromesodermal progenitors while BMP signaling induced blood and endothelium marker expression, and etv2 expression in particular, through transcriptional activation of id
genes, including \textit{id1} and \textit{id3}. Loss of \textit{msgn1/myod/myf5} function or activation of \textit{hs:id3} expression resulted in a dramatic expansion of \textit{etv2} expression into the paraxial mesoderm\textsuperscript{40}. Our results show that a loss of \textit{etv2} expression results in the opposite phenotype, and endothelial progenitors differentiate as skeletal muscle cells. FGF signaling was required for this differentiation, consistent with its role in promoting muscle fate. Similar loss of endothelial cells and expansion of skeletal muscle was observed upon inhibition of BMP signaling using the \textit{HS:dnBMPR} line\textsuperscript{40}, suggesting that BMP inhibits paraxial mesoderm development through \textit{etv2} function.

Intriguingly, based on our scRNA-seq analysis, \textit{id1} and \textit{id3} genes were enriched not in the population of vascular progenitors, but in the LPM cells, together with \textit{twist1a}, \textit{prrx1a}, and...
foxd1/d2 homologs. This suggests that BMP signaling is active in the multipotent progenitors in the LPM and activates etv2 expression through id1 and id3, which are then downregulated as cells differentiate into endothelial or hematopoietic lineages (Fig. 9). Based on our results, etv2 mutant embryos showed a nearly complete absence of vascular endothelial and progenitor cells while the LPM population was increased, suggesting that many LPM cells are arrested in the progenitor stage unable to differentiate further, while others undergo differentiation into skeletal muscle.

A previous study has suggested that id1 and id3 genes function upstream of etv2 during embryogenesis.40 Here, we provide evidence that prx1a and foxd1/d2 are co-expressed with etv2 in the LPM, and are involved in regulating its expression during vasculogenesis. This further supports the model that the LPM population includes early progenitors that can give rise to the vascular endothelial lineage and suggests that additional factors which function upstream of etv2 may be present within this cell population.

While etv2−/− mutant embryos displayed UAS:GFP expression in multiple skeletal muscle cells, some muscle cells were also present in the heterozygous etv2+−/− embryos. Because no GFP expression in skeletal muscle cells was observed in wild-type Tg(-2.3etv2:GFP) embryos, this phenotype is likely caused by the loss of a single etv2 allele in heterozygous etv2+−/− embryos. Intriguingly, vascular defects observed in etv2−/− mutant embryos were more severe compared to etv2−/− and etv2−/− mutants, both of which are expected to be null alleles. A possible explanation is that the insertion of a rather large Gal4-containing hybridization for GFP expression in (Fig. 9). Based on our results, cells differentiate into endothelial or hematopoietic lineages upstream of etv2.

etv2 function results in a similar (albeit more severe) loss of vascular progenitors as seen in homozygous mutations in the mechanisms involved in the early specification of arterial and venous progenitors.

In summary, our work suggests that cells in the vascular endothelial lineage arise from multipotent progenitors in the LPM which can differentiate into skeletal muscle in the absence of Etv2 function. The role of the key transcription factor Etv2 and the molecular mechanisms governing vascular development are highly conserved between many vertebrates, and it is highly likely that our findings will also be relevant for mammalian embryos. Gaining insight into the fundamental mechanisms of cell fate commitment will not only provide us with a deeper understanding of complex developmental processes but will also be important for stem cell differentiation toward highly diverse and specific lineages for regenerative therapies.

Methods

Zebrafish lines. The following zebrafish lines were used in the study: Tg(6sUAAS:EGFP)39, Tg(6sUAAS:EGFP)39, Tg(UAS:ElIV:TRK-mCherry)31, Tg(etv2:mCherry)pf373, Tg(etv2:GFP)ce51, TgBAC(etv2:GFP)32, Tg(etv2:mCherry)zf373, Tg(bfI1aGFP)152, Tg(etv2:mCherry)37, Tg(bfI1aGFP)152, Tg(tgataisadr/Med)13, Tg(act:tcb:GFP)32, Tg(hsp70:dank1b-GFP)54, abbreviated further as hsp70:dk1b, Tg(hsp70:dnfgr1a-1a-GFP)56, abbreviated as hsp70:dnfgr1a. The etv2−/− Gal4/i1a gene trap line (further abbreviated as etv2−/−) was generated by performing a knock-in of a construct with gal4-pA sequence into exon 5 of the etv2 gene by CRISPR/Cas9 mediated homology-independent repair mechanism.4 The etv2−/− mutant line was generated by injecting Cas9 mRNA and etv2 gRNA mixture (gRNA targets the following sequence within the exon 5 of etv2: GGGAGAGCAGCAAGCTTTAGG, PAM sequence is underlined) into embryos from the kdrlCherry line. The line contains a 13 bp insertion in the region targeted by the gRNA, designated in capital letters: ggcccagtaagcttgctggag.

Tg(6sUAAS:loop:pmCerulean-jo:loop:P-NT2-2A-YFP, pm)iac16 line (abbreviated as UAAS:cm) was made using T02-mediated transposon approach. The line showed Cerulean expression in all vascular endothelial cells in non-mosaic fashion when mated to etv2+−/− line.
Throughout the study, embryos were typically incubated at 28.5 °C, except for the embryonic stages prior to 24 hpf, where embryos were incubated for part of the time at 23.5–24 °C to slow down their development. Embryos were staged using previously established criteria.

**MO injections.** Totally, 10 ng (per embryo) of previously validated splice-blocking MO was used for injections: AATGCTCTTACCATCGTTGATTTCA.

To knockdown etv2, 10 ng mixture (5 ng each MO) of two previously described etv2 MOs was used: MO1, TTGGTACATTTCCATATCTTAAAGT and MO2, CACTGAGTCCTTATTTCACTATATC.

**Skeletal muscle cell counts.** To count GFP- or mCherry-positive cells within the skeletal muscle, embryos were mounted in 0.6% low-melting-point agarose and the trunk region was imaged using a Nikon A1 confocal microscope at the CCHMC.

**Marker genes.**

- Erythrocytes
- Epithelial
- Tailbud
- Endothelial progenitors
- Venous
- Arterial
- Endothelial
- Macrophages

**Groups**

1. Erythrocytes
2. Epithelial
3. Tailbud
4. Endothelial progenitors
5. Venous
6. Arterial
7. Endothelial
8. Macrophages

**PC1 (56.0%)**

PC2 (2.8%)

PC3 (1.5%)

**Differential Expression (log2)**

-6.0 –4.5 –3.0 –1.5 0.0 1.5 3.0 4.5 6.0

**Differential Expression (log2)**

**ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-020-16515-y**

**NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-020-16515-y**

**www.nature.com/naturecommunications**
Fig. 8 Single-cell RNA-seq analysis using Fluidigm cell sorting of Tg(fli-2.3 etv2:gFP) embryos at the 16-20-somite stage. a-g GFP expression in live embryos, maximum intensity projection is shown. Arrows label vascular endothelial cells and their progenitors. Ten embryos were imaged in two independent experiments and a representative embryo is shown. a lateral view; b anterior view, c dorsal view. A anterior, P posterior. d Heatmap view of marker gene expression in different cell clusters. A complete list of differentially expressed genes is presented in Supplementary Data 5. e, f 2-D and 3-D principal component analysis plots of different cell clusters. Cluster names are the same as in (d). g Relative marker gene expression in different cell clusters. Vertical bars depict log-normalized gene expression. An Arteriovenous (A-V) index of different endothelial cells. Note that many cells are positive for both arterial and venous marker expression. i-l I SH expression analysis of key marker genes for EPC (etv2), venous (vh4), arterial (cldn5b), and EC-2 (ldba2) populations in the trunk region at the 20-somite stage. Black arrows label the DA and white arrowheads label venous progenitors which are starting to coalesce into the PCV. Note that vh4 is enriched in the PCV while cldn5b and ldba2 label the DA. m–t Two color ISH analysis for the expression of venous dab2 and arterial cldn5b at the 20-somite and 24 hpf stages. Arrows label the DA while arrowheads mark the PCV or its progenitors. Note that dab2 and cldn5b are co-expressed in the DA progenitors at the 20-somite stage but not at 24 hpf. In all panels, the numbers in the lower right corner display the number of embryos showing the expression pattern out of the total number of embryos analyzed in two replicate experiments.

Confluent Imaging Core. Fluorescent cells within the skeletal muscle were counted manually within the entire trunk region by analyzing confocal z-stacks.

In situ hybridization. To perform ISH, embryos were fixed overnight in BT-Fix (4% paraformaldehyde in 1× phosphate-buffered saline (PBS)), dehydrated in sequential ethanol series and stored at −20 °C. Embryos were rehydrated and washed 3× in PBT (1× PBS, 0.2% bovine serum albumin, 0.2% Tween 20). Embryos were incubated in the prehybridization buffer (50% formamide, 5× SSC, saline sodium citrate, 50 µg/ml heparin, 5 mM EDTA, 0.5 mg/ml RNA, 0.1% Tween 20, citric acid added to pH 6.0) for 2 h at 65 °C. DIG-labeled RNA probe solution was added at approximately 1 µg/ml concentration in prehybridization buffer (prehyb) and incubated overnight at 65 °C with shaking. Embryos were washed at 65 °C with 5× SSC buffer: 75% prehyb/25% SSC, 50% prehyb, 50% 2× SSC, 25% prehyb, 75% 2× SSC, 2× SSC; twice with 0.2× SSC. Subsequently, embryos were washed at room temperature in the following solutions: 75% 0.2× SSC, 25% PBT, 50% 0.2× SSC, 50% PBT, 25% 0.2× SSC, 75% PBT, and 100% PBT. Afterwards embryos were incubated overnight in 1× PBT with 2% lamb serum and 14000 dilution of anti-digoxigenin-AP antibody (Sigma-Aldrich, cat. No. 11093279010). Embryos were washed 6× in PBT solution and incubated in AP buffer (100 mM NaCl, 50 mM MgCl2, 100 mM Tris-Cl, pH 9.5, 0.1% Tween 20) in the presence of 4-nitrotetrazolium blue chloride (0.225 mg/ml) and 5-Bromo-4-chloro-3-indolyl phosphate disodium salt (0.175 mg/ml). Staining was stopped with PBT washes.

Microscopy imaging and image processing. Embryos were whole mounted in slides for imaging using agarose. To image stained embryos after ISH, images were captured using a 10× objective on an AxioImager Z1 (Zeiss) compound microscope with an AxioCam ICC3 color camera (Zeiss). Images in multiple focal plans were captured individually and combined using the Extended Focus module within Axiosvision software (Zeiss). For confocal imaging, embryos were mounted in 0.6% low melting point agarose stock solution. Hairpin solutions were added to embryo samples in amplification buffer and incubated overnight in the dark at room temperature. Samples were washed 5× with 5× SSCT solution, and stored at 4 °C until imaging. GFP fluorescence was still apparent after HCR. Embryos were mounted in 0.6% low melting point agarose and imaged using a Nikon A1 confocal microscope at the CCHMC Confocal Imaging Core.

Immuno-staining. F310 (1:50 dilution) and S58 (1:10) primary antibodies (Developmental Studies Hybridoma Bank, Iowa City, IA) were used for whole-mount immuno-staining of etv2 and myoD. UAS:GFP embryos at 24 hpf. No immuno-staining was performed for GFP which was visible in fixed embryos. Anti-mouse IgG, CF594 antibody (Sigma-Aldrich, SAB4600098) was used as a secondary antibody.

etv2 overexpression and quantitative RT-PCR. etv2 mRNA was synthesized as previously described11. To analyze expression of muscle markers, approximately 75 pg of zebrafish etv2 mRNA was injected into fli1a:gFP embryos at the 1-cell stage. Embryos were screened for ectopic GFP expression at the 10-somite stage and then frozen for qPCR analysis. Groups of 10 embryos were analyzed in two independent experiments. An RNAqueous 4-PCR kit (ThermoFisher) was used to extract RNA. cDNA synthesis was performed using Superscript IV cDNA synthesis kit (ThermoFisher). Quantitative real-time PCR was performed using SYBR Green Master Mix (ThermoFisher) and StepOne Software V.3.3 (Applied Biosystems). The following primers were used: ef1α (TCAACCTGGGAGTTAAGCA), (ACTTGCGAGGCCATGCTG), (GAGTTCGTACGACTGTC) and (GCGTGCAGCTGACG), myo5f (GGTGTGACTGCAAGCTGTG), and (GCGTTGCGCTGACCCTTT) TFi, (GCAATAACGGCAGAGCGA), and (CAGCCCTCTCGACTGCTT).
The results were analyzed using StepOne Software v2.3 (Applied Biosystems). Two biological and two technical replicates were obtained for each sample (four biological replicates were available for control un.injected embryos). Relative expression values were normalized for ef1α expression. Statistical analysis was performed using Prism 8 software (GraphPad Software).

**Single-cell RNA sequencing using the Chromium platform.** Approximately, 100–150 of etv2ζ2ζGt/+; UAS:GFP and 75–100 etv2ζζζGt/–; UAS:GFP embryos were collected at the 20–somite stage and dissociated into a single-cell suspension using the previously reported protocol. Briefly, the embryos were manually dechorionated, and transferred into the deyolking buffer (55 mM NaCl, 1.8 mM KCl, and 1.25 mM NaHCO3). Embryos were pipetted up and down on ice using p1000 pipettor until yolk was dissolved. Embryos were centrifugated at 300 G for 1 min, and supernatant was removed, while the pellet was resuspended in 0.5x Danieu solution (1x Danieu: 58 mM NaCl, 0.7 mM KCl, 0.4 mM MgSO4, 0.6 mM Ca(NO3)2, 5 mM HEPES, pH 7.6). Centrifugation and removal of supernatant was repeated following another centrifugation at 300 G. The pellet was resuspended in FACs max solution cell dissociation solution, which was then passed through a cell strainer. Cells were centrifugated at 300 G and suspended in a buffer of 1x PBS containing 1 mM EDTA and 2% fetal bovine serum (FBS). Fluorescence-activated cell sorting was then used to collect GFP+ cells in a solution of 50% FBS in 1x PBS (using Supplementary Fig. 16A for FACS gating strategy). Single cells were captured and processed for RNA-seq using the Chromium instrument (10x Genomics) at the CCHMC Gene Expression Core facility. RNA-seq was performed at the CCHMC DNA Sequencing core on Illumina HiSeq3000 sequencer using one flow cell of paired-end 75 bp reads, generating 240–300 million total reads.

**Single-cell RNA sequencing using Fluidigm C1.** Sequencing was performed on cells from zebrfish embryos at 16–18 somite stages and at 20 somite stages. Approximately, 300 Tg(−2.3 etv2gfp) embryos were collected at each stage, dissociated into a single-cell suspension as described above and suspended in a buffer of 1x PBS containing 1 mM EDTA and 2% FBS. Fluorescence-activated cell sorting was then used to collect GFP+ cells in a solution of 50% FBS in 1x PBS (Supplementary Fig. 16B). Cells were then counted on a hemocytometer, spun down at 300 x G and resuspended at a concentration of approximately 30,000 cells per ml. Each 5 ml resuspension was then loaded on a 96-well Fluidigm C1 chip for capture of single cells at the CCHMC Gene Expression core. Forty-eight individual cells from each stage group were selected for library synthesis, barcoding, and pooled paired-end sequencing using a single lane of a P75 flow cell at 1.5–2 million reads per cell, performed at the CCHMC DNA Sequencing core. Data analysis were performed using AltAnalyze. Reads were aligned to Ensembl release 72 of the Zv9 assembly of the zebrafish genome and gene counts were quantified using Kallisto. Iterative Clustering and Guide-gene Selection (ICGS) was used to identify de novo clusters/groups of cells. ICGS was performed using cosine similarity with an expression fold cutoff of six for a minimum of 25% of cells within a cluster and a Bonferroni adjusted q value to create a list of ordering genes. An order file was then applied to the data based on the top 425 genes created in the previous step to label whether a cell is used for ordering or not. In order to determine cell differentiation trajectories, dimensionality reduction was used to limit the number of cell clusters. The trajectories of cells were plotted on a pseudotime graph and colored to the total reads per cell and a scale factor of 10,000× applied. To obtain arteriovenous index, the average of the log2 expression was calculated across all cells for each marker per cluster. P-value was determined by a two-tailed Mann–Whitney U test comparing fold change in the cluster of interest to fold change in all other clusters. The table of average expression values for genes in each cluster was created directly from a text file generated from AltAnalyze. These values represent the average expression of all cells in that cluster calculated as reads for a specific gene that have been normalized to the total reads per cell and a scale factor of 10,000× applied.

To obtain arteriovenous index, the average expression was calculated first for the following known arterial and venous markers: dhnfb, ephb2a, dll4, hox2, dlx, notch3 (arterial), and flt4, dab2, stab11, ephb4a, mrc1a (venous).

**Fluorescent cell counting.** Percentage of fluorescent cells which express eve2 reporter was estimated in etv2ζ2ζGt/−; UAS:GFP embryos alone or crossed with hsp70:DMk1-GFP and hsp70:nDGFPR1-GFP lines. Cells from 10 to 25 embryos were collected at each stage and examined. The number of detected nucleoli (nUMI) was removed by regressing out these variables during the scaling of data. Fourteen significant principal components were selected for t-SNE dimensionality reduction, using a resolution of 1.0 for clustering. Based on these parameters, two cell types were apparent as distinct populations of endothelial cells, which was also confirmed by a heatmap of gene expression of the two groups.
Ethics approval. Zebrafish embryo experiments were performed under animal protocol: IACUC2016-0039, approved by the Institutional Animal Care and Use Committee at the Cincinnati Children’s Hospital Medical Center.

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

Data availability. All data that support the findings of this study are available from the corresponding author upon reasonable request. scRNA-seq data generated in the study using the Chromium (10× Genomics) platform have been deposited in the NCBI GEO database under the accession codeGSE143750. scRNA-seq data generated in the study using the Fluidigm platform have been deposited in the NCBI GEO database under the accession codeGSE142484.

Received: 5 May 2019; Accepted: 29 April 2020; Published online: 03 June 2020.
57. Sumanas, S., Jorniak, T. & Lin, S. Identification of novel vascular endothelial-specific genes by the microarray analysis of the zebrafish cloche mutants. *Blood* **106**, 534–541 (2005).
58. Rost, M. S. & Sumanas, S. Hyaluronic acid receptor Stabilin-2 regulates Erk phosphorylation and arterial-venous differentiation in zebrafish. *PLoS ONE* **9**, e88614 (2014).
59. Wong, K. S. et al. Hedgehog signaling is required for differentiation of endocardial progenitors in zebrafish. *Dev. Biol.* **361**, 377–391 (2012).
60. Herbol, P., Thisse, B. & Thisse, C. Ontogeny and behaviour of early macrophages in the zebrafish embryo. *Development* **126**, 3735–3745 (1999).
61. Palencia-Desai, S. et al. Myocardium and BMP signaling are required for endocardial differentiation. *Development* **142**, 2304–2315 (2015).
62. Choi, H. M. T. et al. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* **145**, dev165753 (2018).
63. Manoli, M. & Driever, W. Fluorescence-activated cell sorting (FACS) of fluorescently tagged cells from zebrafish larvae for RNA isolation. *Cold Spring Harb. Protoc.* **2012**, pdb.prot069633 (2012).
64. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
65. Macosko, E. Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
66. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
67. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
68. Olsson, A. et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* **537**, 698–702 (2016).
69. Sigdel, T. K. et al. The identification of novel potential injury mechanisms and candidate biomarkers in renal allograft rejection by quantitative proteomics. *Mol. Cell Proteom.* **13**, 621–631 (2014).

**Acknowledgements**
This research was supported by the awards from National Institute of Health: NIH R01 HL134815 and R21 AI128445 to S.S., and NIH F31 HL135986 to A.L.K. We thank Nathan Salomonis and Praneet Chaturvedi for their assistance with scRNA-seq analysis, Steve Potter for critical comments, and CCHMC Gene expression and FACS cores for their assistance with the project.

**Author contributions**
B.C. generated ev2(2002) line and performed most of experiments in the study, S.C.C. performed scRNA-seq analysis of the Chromium data and edited the paper, A.L.K. performed all experiments using the Fluidigm platform and subsequent data analysis, as well as the cell disaggregation and initial data processing using the Chromium platform, S.S. initiated and supervised the entire study, contributed to data analysis, and wrote the paper.

**Competing interests**
The authors declare no competing interests.

**Additional information**
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16515-y.

Correspondence and requests for materials should be addressed to S.S.

Peer review information *Nature Communications* thanks Peng Jiang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.