nature communications

Article

https://doi.org/10.1038/s41467-025-66646-3

3D spatial organization of heterogeneous *nkx2.5*+ progenitors in the zebrafish heart field pre-patterns cardiovascular development

Received: 28 May 2025

Accepted: 12 November 2025

Published online: 25 November 2025

Check for updates

Hakan Coskun[1,2,4], Yunxia Wang[1,2,4], Xinlei Gao[1,2], Kaifu Chen[1,2,5] ✉, C. Geoffrey Burns[1,2,5] ✉ & Caroline E. Burns[1,2,3,5] ✉

Nkx2.5+ progenitor cells are specified in the bilateral zebrafish heart field, which contributes to multiple cardiovascular derivatives. After specification, the field physically separates into a cardiac fraction that migrates to the midline to form ventricular and atrial myocardium, a pharyngeal fraction that is sequestered in the pharyngeal arches before generating head muscles, the outflow tract, and pharyngeal arch arteries, and a lateral fraction that becomes inflow atrial and pacemaker cardiomyocytes as well as pericardium. Whether these progenitors are transcriptionally distinct and regionally organized has remained unclear. Here, we demonstrate that the heart field is composed of heterogeneous *nkx2.5*+ progenitor populations that are arranged according to their migratory paths and end fates. Through prospective fate mapping and mutant analyses, we define the derivatives of each subgroup. Our data provide a 3D blueprint of heart field organization, offering insights into cardiovascular development and identifying vulnerabilities contributing to congenital heart disease.

The vertebrate cardiovascular system is built by progressive differentiation of *nkx2.5*+ progenitor cells specified in the bilateral anterior lateral plate mesoderm (ALPM), a region classically defined as the heart field[1–3] In zebrafish, these bilateral domains physically separate into three major subpopulations, including cardiac, pharyngeal, and lateral, during the early stages of heart morphogenesis[4–7].

The cardiac fraction, consisting of first heart field (FHF) cardiomyocytes and arterial pole second heart field (SHF) progenitors, dives deep into the embryo towards the midline to form the cardiac disc, which later becomes the myocardium of the single ventricular chamber[5,8–12] and much of the atrium[11,13]. The pharyngeal fraction remains superficial and becomes sequestered in the cores of pharyngeal arches 2-6 (PA2-6)[4,5]. Cells in PA2 give rise to a complex mix of cell fates and end structures, including five striated head muscles (HMs), the hypobranchial artery endothelium, and three lineages in the outflow tract (OFT), including the myocardium, smooth muscle, and endothelium[5]. By contrast, progenitors in PAs 3-6 only produce the endothelial lining of their respective pharyngeal arch artery (PAA)[4]. The lateral fraction contributes to the inflow portion of the atrium through late differentiation of myocardial and pacemaker cardiomyocytes[7] and to the pericardium[6], a mesothelial sac surrounding the heart[14]. While FHF progenitors differentiate in the ALPM, the remaining cells stay undifferentiated with respect to their downstream fates until after migration or PA sequestration[4–7,11,15]. As such, it

[1]Division of Basic and Translational Cardiovascular Research, Department of Cardiology, Boston Children's Hospital, Boston, MA, USA. [2]Harvard Medical School, Boston, MA, USA. [3]Harvard Stem Cell Institute; Cambridge, Boston, MA, USA. [4]These authors contributed equally: Hakan Coskun, Yunxia Wang. [5]These authors jointly supervised this work Kaifu Chen, C. Geoffrey Burns, Caroline E. Burns. ✉e-mail: Kaifu.Chen@childrens.harvard.edu; Geoff.Burns@childrens.harvard.edu; Caroline.Burns@childrens.harvard.edu

has been unclear whether pre-migratory *nkx2.5*+ progenitors in the heart field are transcriptionally distinct based on their end fates, and if so, how they are spatially arranged.

In this study, we combine single-cell RNA sequencing (scRNA-seq), spatial transcriptomics, prospective lineage tracing, and mutant analyses to demonstrate that the *nkx2.5*+ heart field consists of molecularly distinct subpopulations that are spatially organized according to their future migratory paths and structural derivatives. Beyond myocardial progenitors, we identify precursors for the HMs, OFT, PAAs, and pericardium. Together, our findings provide a high-resolution, single-cell transcriptional and 3D spatial atlas of the ALPM, offering a comprehensive framework for understanding cardiovascular morphogenesis.

Results

Nkx2.5+ progenitors in the ALPM are molecularly heterogeneous

To reveal potential heterogeneity within the *nkx2.5*-expressing lineage in zebrafish, we used fluorescence-activated cell sorting (FACS) to isolate ZsYellow+ cells from homozygous *Tg(nkx2.5:ZsYellow)* embryos before [14-somite stage; 16 hours post-fertilization (hpf)] and after (20-somite stage; 19 hpf) the physical separation of the cardiac, pharyngeal, and lateral subpopulations[5] (Fig. 1a; and Supplementary Fig. 1a). Using the inDrops platform[16], we performed scRNA-seq of 6,000 cells from two biological replicates (3,000 cells/replicate). Transcriptomes were analyzed using the python package celseq2[17] with 3,761 and 3,170 cells passing filtering and quality control at the 14- and 20-ss, respectively (Supplementary Fig. 1b, c; and Supplementary Data 1). Using targeted cross-time integration and graph-based clustering, we identified 6 cell populations at each timepoint (Fig. 1b, c) that were completely overlapping on independent and integrated UMAP plots (Supplementary Fig. 2a; and Fig. 1d). Based on marker gene expression, we observed an undifferentiated lateral plate mesoderm (LPM) cluster, a FHF cardiomyocyte cluster, and several unanticipated clusters including neural, endothelial, epidermal, and tailbud (Fig. 1b,c). Feature plots revealed *ZsYellow* expression in all 6 populations (Fig. 1e) with *nkx2.5* transcripts most abundant in the LPM and cardiomyocyte clusters (Fig. 1f). This pattern is consistent with previous observations demonstrating persistence of *ZsYellow* transcripts and fluorescence after downregulation of *nkx2.5*[4].

Because we suspected that the LPM cluster contained undifferentiated pharyngeal and lateral progenitors, we performed high resolution sub-clustering at each time point, which uncovered three transcriptionally distinct populations we termed LPM-1, −2, and −3 (Fig. 1g; and Supplementary Fig. 2b). Among other unique molecular signatures, LPM-1 expressed *aplnra*, *hey1*, and the retinoic acid (RA)-eliminating enzyme *cyp26c1*, LPM-2 expressed *bmp4, bmp6*, and *cdh11*, and LPM-3 expressed the transcriptional repressor *prdm1a* and the RA-converting enzymes *rdh10a* and *aldh1a2* (Fig. 1h; and Supplementary Fig. 2c). These molecular profiles suggest that RA signaling, which is known to promote PAA fates at the expense of SHF-derived ventricular cardiomyocytes and the OFT[18], is intrinsically low in LPM-1 and high in LPM-3. Comparison of marker gene expression over time revealed highly similar expression profiles (Supplementary Fig. 3a), consistent with population stability between these embryonic stages.

Gene ontology (GO) categories enriched in LPM-1 included ameboidal-type cell migration (Supplementary Fig. 3b; Supplementary Data 2), which might foreshadow active cell movement. Specifically, we detected *aplnra/b*, *cxcl4-cxcr12*, and *robo4*, all of which have been implicated in cardiac cell migration[19–21]. Although not observed in LPM-1, terms related to striated muscle, including *tnnt2a*, *tpm4a*, and *myh7*, were enriched in LPM-2 and LPM-3 (Supplementary Fig. 3b; and Supplementary Data 2) along with unique categories such as epithelial tube morphogenesis and retinol metabolic process, respectively. Specifically, we detected *hand2*, *cdh1*, and *llgl1* in the epithelial category and *rdh10a* and *aldh1a2* in the retinol category. From these molecular and GO term profiles, we conclude that each LPM sub-cluster is undifferentiated and stable between the 14- and 20-ss with respect to their final fates.

Nkx2.5+ cell populations occupy distinct regions of the ALPM

To learn if each cell cluster occupies a distinct region of the ZsYellow+ ALPM, we performed MERFISH spatial transcriptomics[22] to measure transcript number and distribution on serial sections of 14-ss *Tg(nkx2.5:ZsYellow)* embryos (Fig. 2a; and Supplementary Fig. 4). Specifically, 10 µm sections were cut along the length of the anterior to posterior (A-P) axis of the ALPM and processed by immunostaining for ZsYellow to mark the *nkx2.5* lineage, β-catenin to mark cell borders, and counter-stained with DAPI to highlight nuclei (Fig. 2a–c; and Supplementary Fig. 4). Sections were simultaneously hybridized with MERFISH-encoding probes recognizing 133 transcripts representing unique markers from each cell population identified by our scRNA-seq (Fig. 1b, c) along with transcripts more broadly expressed in the ALPM (Supplementary Data 3). After processing, 35 sections with ZsYellow+ ALPM cells from 4 independent embryos maintained sufficient integrity for downstream analyses.

Cell segmentation and selection analysis were performed on the processed sections with representative images shown at five levels along the A-P axis (Fig. 2c, d; and Supplementary Fig. 4). To identify distinct cells, we used Cellpose[23] to integrate the ZsYellow, β-catenin, and DAPI signals. Subsequently, ZsYellow+ cells in the ALPM region were selected for downstream clustering and annotation. In total, 1032 ZsYellow+ cells, each expressing *nkx2.5* (Supplementary Fig. 5a), were included that separated into five clusters on a UMAP plot (Fig. 2e; and Supplementary Data 4). Marker gene analysis revealed that the five MERFISH populations corresponded to the cardiomyocyte and LPM populations identified by scRNA-seq with LPM-1 further sub-clustering into LPM-1a and LPM-1b (Fig. 2e, f; and Supplementary Fig. 5b). Feature plots of *nkx2.5* and its paralog *nkx2.7*[24] revealed overlapping expression with complementary patterns in transcript levels: *nkx2.5* was enriched in the cardiomyocyte, LPM-1a, and LPM-3 populations, whereas *nkx2.7* was highest in LPM-1b and LPM-2 (Supplementary Fig. 5c, d).

Robust expression of *rbfox1l*, *myl7*, and *tnnt2a* were observed in the cardiomyocyte cluster as shown in dot and feature plots with lower levels seen in LPM-2 (Fig. 2f, and Supplementary Fig. 6a, b), suggesting that these two populations are transcriptionally more similar. *Gata5* and *hand2*, which have also been associated with the cardiac lineage[25], were additionally enriched in cardiomyocytes and LPM-2, while *tbx1*, which has been associated with the pharyngeal lineage[25,26], was seen in a reciprocal pattern (Fig. 2f; and Supplementary Fig. 6a, b). LPM-1a, which had the fewest cells surveyed, was enriched for *cyp26c1*, *aplnra*, *cxcl12a*, and *fsta* (Supplementary Fig. 6a, b) and LPM-1b most uniquely expressed *pltp* (Fig. 2f; and Supplementary Fig. 6a, b). Manual inspection revealed the presence of two sub-groups in LPM1b that were defined by differential expression of *prdm1a*, *hoxb1b*, *hoxb3a*, and *meis3* (Supplementary Fig. 6b). Although we were unable to sub-cluster LPM-1b computationally, this likely reflects insufficient cell sampling, sequencing depths, and transcripts probed. Finally, in addition to *prdm1a*, LPM-3 was also enriched for the RA converting enzymes, *rdh10* and *aldh1a2* (Fig. 2f; and Supplementary Fig. 6a, b). From these molecular and GO term profiles, we conclude that each LPM sub-cluster is undifferentiated at the 14- and 20-ss with respect to their final fates.

Spatial mapping of the MERFISH populations revealed that each occupied a distinct and reproducible region of the ALPM (Fig. 2g). The relative position of each population was also evident in a 3D-reconstruction of the serial sections (Supplementary Fig. 4; and Supplementary Movie 1). Specifically, we found that LPM-1a was smaller in size, restricted to the anterior-most sections, and lateral of the LPM-1b

https://doi.org/10.1038/s41467-025-66646-3

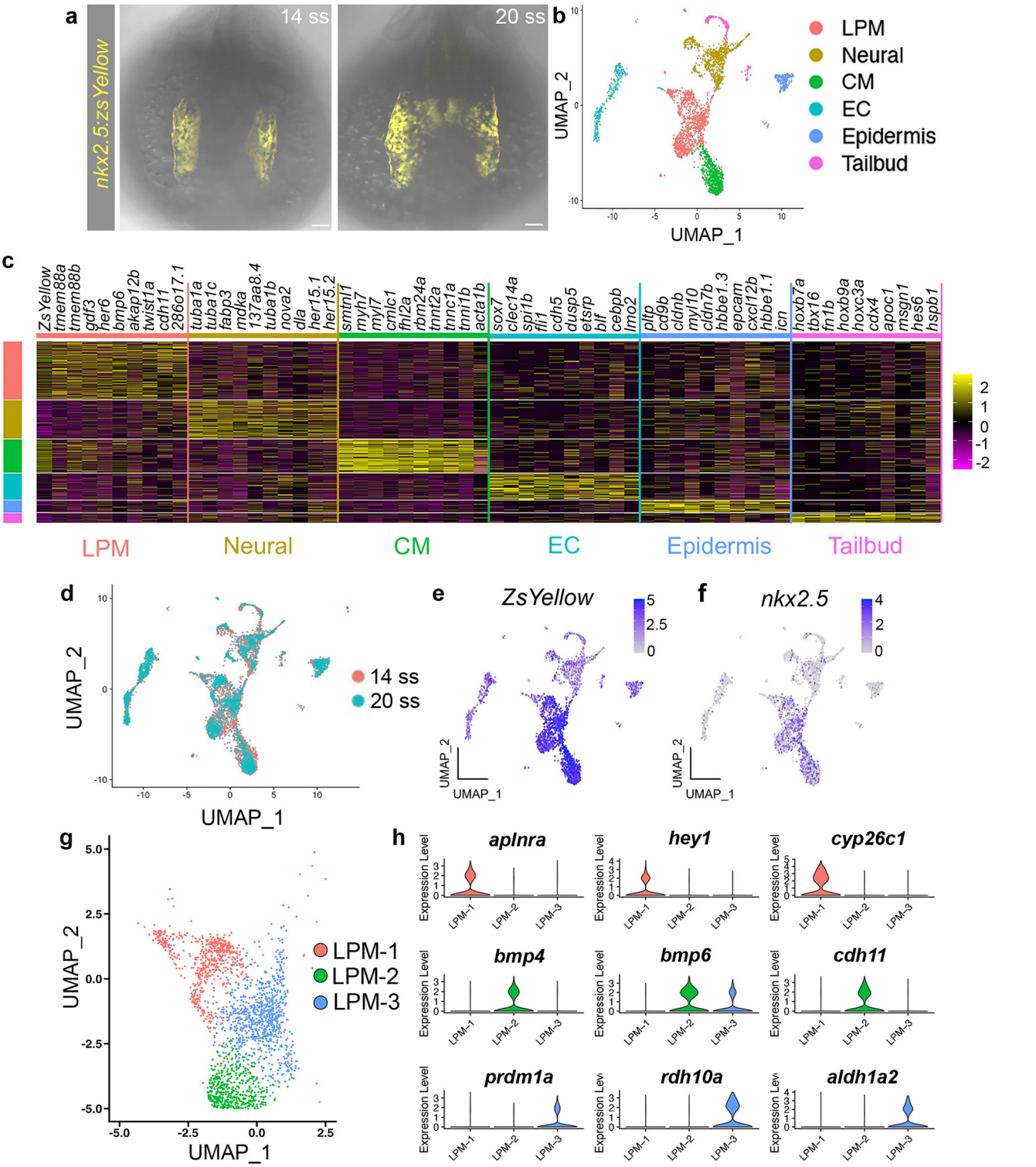
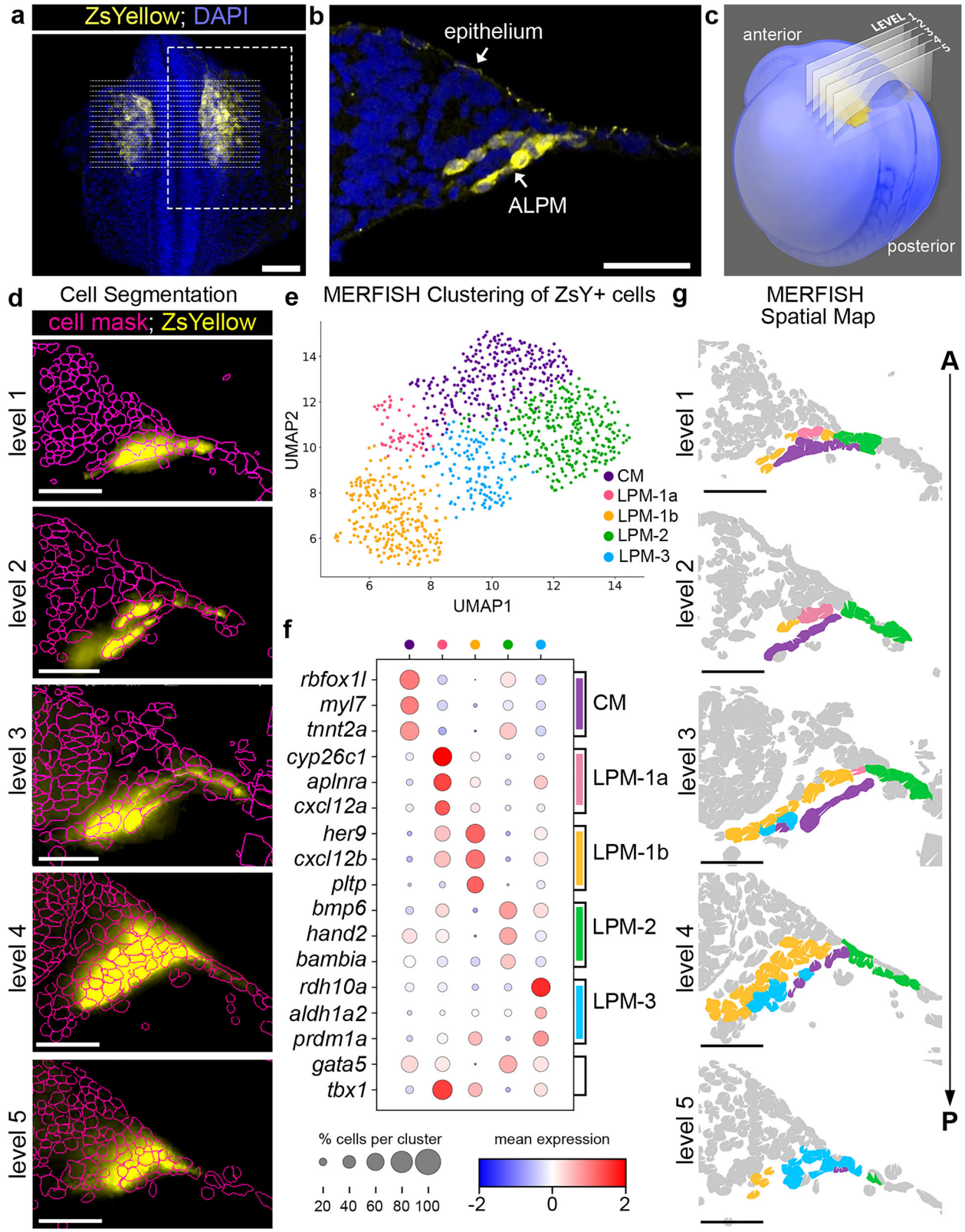


Fig. 1 | *Nkx2.5*+ progenitors are molecularly heterogenous in the ALPM. **a** Confocal Z-stacks of 14-ss and 20-ss *Tg(nkx2.5:ZsYellow)* embryos immunostained with ZsYellow antibodies. (14-ss, n = 12; 20-ss, *n* = 9). Dorsal views. **b** Integrated UMAP plot of scRNA-seq data from ZsYellow+ cells at the 14-ss and 20-ss. Each cell cluster is colored according to their identity. **c** Heat map of key marker genes used for cluster annotation. Gene expression levels are depicted by pink-to-yellow scaled Z-scores. **d** Overlaid UMAP plots of scRNA-seq data color coded according to developmental stage. **e**, **f** UMAP feature plots showing differential *ZsYellow* (E)and *nkx2.5* (**f**) expression between cell clusters. Expression values are represented by white-to-purple as shown. **g** Integrated UMAP plot depicting color-coded LPM sub-clustering into three distinct cell populations. **h** Volcano plots of LPM subpopulation marker gene expression that are colored coded by subpopulation identity as shown in (**g**). Abbr. *ss* somite stage, *LPM* lateral plate mesoderm, *CM* cardiomyocyte, *EC* endothelial cell. Scale bars: 25 microns.

https://doi.org/10.1038/s41467-025-66646-3



domain, which extended the length of the A-P axis. LPM-2 progenitors mapped to larger cells at the lateral edge, while LPM-3 progenitors were restricted to the posterior half of the ALPM and concentrated at the tip.

As observed in feature plots (Supplementary Fig. 5c, d), complementary patterns in *nkx2.5* and *nkx2.7* expression levels were also evident on MERFISH sections (Fig. 3a, b). Similarly, sub-partitioning of LPM-1b was also seen, as *prdm1a*, *hoxb1b*, *hoxb3a*, and *meis3* transcripts were enriched in the medial and posterior regions (Supplementary Fig. 7a–c). Interestingly, we found that *tbx1*-expressing LPM-1a and LPM-1b progenitors occupy a separate cell tier that is dorsal to underlying *myl7*-expressing cardiomyocytes (Fig. 3c, d). Furthermore, their relative positions in the ALPM predict their eventual migratory patterns. Specifically, *myl7*+ cardiomyocytes, which will migrate

Fig. 2 | *Nkx2.5*+ progenitor populations are spatially organized into discrete regions in the ALPM. a Confocal Z-stack of a 14-ss *Tg(nkx2.5:ZsYellow)* embryo immunostained for ZsYellow (yellow) and counterstained with DAPI (blue). Dashed lines represent transverse sections cut along the embryonic anterior-posterior (A-P) axis of the ALPM. ($n = 12$) Dorsal view, anterior up. Boxed region encompasses the right half of the embryo that is shown in remaining figure panels. **b** Confocal image of the right half of a representative transverse section through an embryo as shown in (**a**). Arrows point to the ZsYellow+ ALPM and epithelium. ($n = 7$) (**c**) Schematic of the embryos shown in panel (**a**) depicting five section levels along the A-*P* axis that are referenced in (**d**) and (**g**). Artwork by Kai-Ou Tang. **d** MERSCOPE image of half of the ALPM in representative transverse sections at each of the five levels depicted in (**c**). Cell segmentation is represented by software-generated pink masks based on beta catenin and DAPI signals. ZsYellow is shown. **e** UMAP plot of color-coded MERFISH cell populations. **f** Dot plot of key cell fate markers used to annotate MERFISH clusters. Dot size corresponds to the percentage of cells expressing the marker gene per cluster and the Z-score of the mean expression level is shown in the blue-to-red scale. **g** Software-generated MERFISH sections that correspond to the embryonic sections shown in (**d**). Each ZsYellow+ cell is color coded based on its identity as shown in (**e**) and (**f**). Each gray cell was negative for ZsYellow signals. Abbr: ALPM, anterior lateral plate mesoderm. Scale bars: 50 microns.

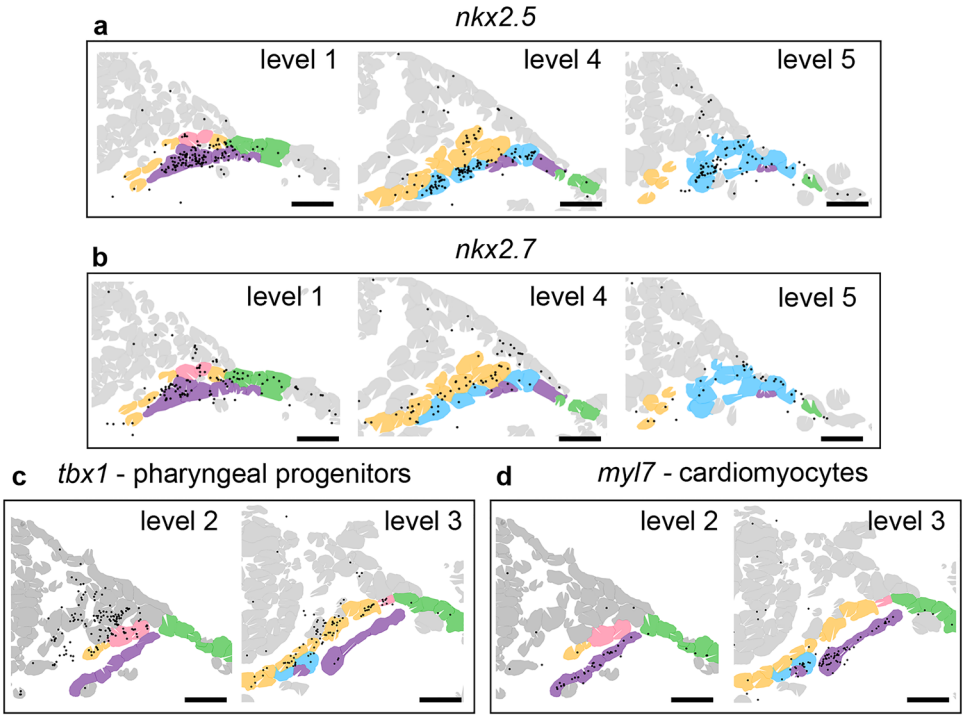


Fig. 3 | Reciprocal *nkx2.5* and *nkx2.7* expression in ALPM cell populations with LPM progenitors positioned dorsal to cardiomyocytes. a–d Software-generated MERFISH sections showing color-coded ZsYellow+ cell populations and individual transcripts as black dots. Levels 1-5 correspond to the anterior-posterior position in the ALPM as depicted in Fig. 2c and color legend is shown in Fig. 2e. (a,b) *Nkx2.5* and *nkx2.7* show overlapping expression patterns in the ALPM, but with reciprocal levels. **c**, **d** *Tbx1*-expressing LPM-1a and LPM-1b pharyngeal (SHF) progenitors reside in a separate cellular tier that is dorsal to *myl7*-expressing CMs. Scale bar: 25 microns.

medially and dive deep into the embryo[5], are located ventral to *tbx1*+ pharyngeal progenitors, which remain more superficial before becoming sequestered into the PAs.

To validate our MERFISH spatial mapping, we performed multiplexed RNAscope fluorescent in situ hybridization followed by antibody staining for ZsYellow in 14-ss *Tg(nkx2.5:ZsYellow)* embryos (Supplementary Fig. 8). In four-color confocal projections, we confirmed the expression patterns of LPM marker genes in the broader ZsYellow+ domain. Specifically, the LPM-1a marker *cyp26c1* was observed in the anterior-most region, the LPM-2 marker *cdh11* was seen in the lateral region, and the LPM-3 marker *aldh1a2* was found in the posterior tip (Supplementary Fig. 8a, b). Overlap between LPM marker gene expression and ZsYellow was confirmed in single optical sections (Supplementary Fig. 8c). ZsYellow+ cells that failed to express any of these LPM makers presumably represent LPM-1b progenitors and FHF-derived cardiomyocytes. This regionalized pattern of gene expression was consistent across biological replicates and corroborates our spatial transcriptomics. Overall, the spatial distribution of MERFISH clusters demonstrates that transcriptionally distinct *nkx2.5*+ subpopulations compartmentalize within the ALPM into discrete cellular communities prior to their migration and differentiation.

Identification of LPM derivatives by prospective fate mapping

To identify the derivatives of each LPM subpopulation, we used prospective fate mapping in 14-ss (16 hpf) *Tg(nkx2.5:Kaede)* embryos by targeted photoconversion of Kaede (Kaede^{PC}) expressed from *nkx2.5* cis-regulatory sequences[4,9]. As depicted schematically (Fig. 4a, d, g, j), we used a confocal laser to achieve regionalized photoconversion where Kaede instantly switches from green to red fluorescence (pseudo-colored yellow to magenta) following exposure to ultraviolet light. Photoconverted animals were imaged immediately, raised to 28 hpf, and imaged again.

Following unilateral photoconversion in the anterior-most ALPM, which contains cardiomyocytes and LPM-1a (Fig. 4a, b), we observed Kaede^{PC} signals in the linear heart tube and PA2 (Fig. 4c). While consistent with prior fate-mapping results[5,27], unconverted Kaede was also detected in PA2 (Fig. 4c), suggesting contributions from additional domains. To test this directly, we photoconverted the middle ALPM, which includes cardiomyocytes, LPM-1b, and LPM-3 (Fig. 4d, e) and observed Kaede^{PC} in the linear heart tube, adjacent ventricular SHF precursors, PA2, and PA3 (Fig. 4f). Photoconversion of the posterior ALPM, occupied exclusively by LPM-3 (Fig. 4g, h), resulted in Kaede^{PC} in PAs 3-4 (Fig. 4I). Together, these experiments demonstrate the

https://doi.org/10.1038/s41467-025-66646-3

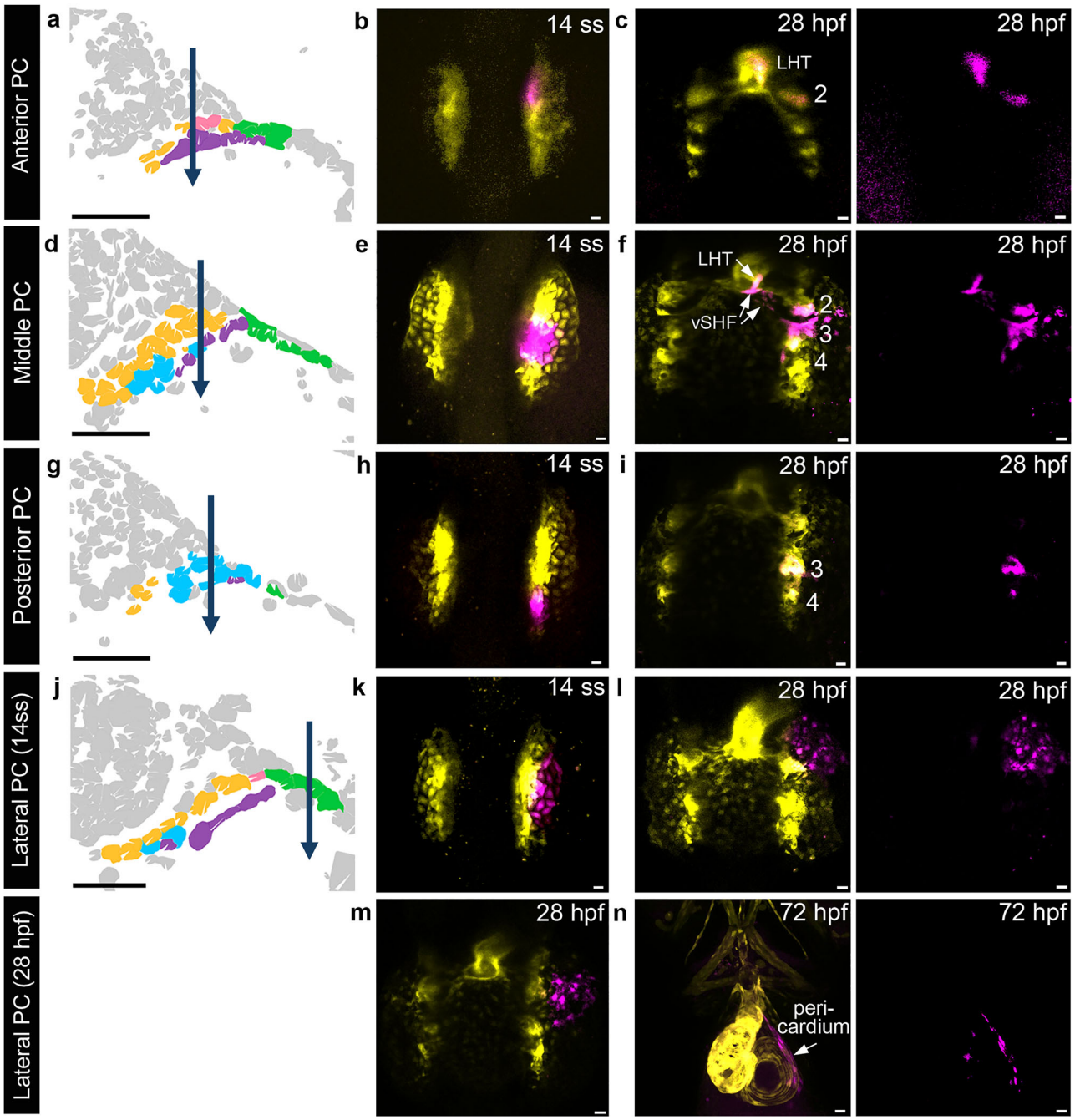


Fig. 4 | LPM subpopulations give rise to distinct cardiovascular derivatives. **a**, **d**, **g**, **j** Software generated MERFISH sections depicting *nkx2.5* + ALPM populations in the path of the UV laser. Population identity is color coded as shown in Fig. 2e: CM, purple; LPM-1a, pink; LPM-1b, yellow; LPM-2, green; LPM-3, blue. Scale bars: 50 microns. **b**, **e**, **h**, **k**, **m** Merged confocal Z-stacks of Kaede (yellow) and Kaede^{PC} (magenta) immediately following photoconversion in a *Tg(nkx2.5:Kaede)* embryo. Dorsal views, anterior up. Scale bars: 25 microns. **c**, **f**, **i**, **l**, **n** Merged and single-channel confocal Z-stacks at 28 or 72 hpf of the same embryo shown in panel to the left. Dorsal views, anterior up. Sample sizes: **b**, **c** $n = 6$; **e**, **f** $n = 4$; **h**, **i** $n = 7$; **k**, **l** $n = 5$; **m**, **n** $n = 7$.Scale bars: 25 microns. Abbr: *LHT* linear heart tube; *vSHF* ventricular second heart field; 2-4, pharyngeal arch number.

following: LPM-1a contributes to PA2, which gives rise to OFT and HM fates[5]; LPM-1b gives rise to SHF ventricular precursors that become the distal ventricle[8–12] and PA2 (OFT and HM fates)[5]; and LPM-3 contributes to PAs 3-4, which become the endothelium of their respective PAA[4]. Although this Kaede tracing could not directly link LPM-1a or LPM-1b to OFT or HM fates, the enrichment of *cyp26a1* and *fsta* in LPM-1a suggests that these are the HM precursors, consistent with prior associations of these genes with HM fates in mice[28].

To track LPM-2 derivatives, we photoconverted Kaede in the lateral-most cells (Fig. 4j, k). At 28 hpf, Kaede^{PC} was detected in a small number of cells in the atrium (Supplementary Fig. 9), consistent with previous observations showing contribution to late-differentiating atrial and pacemaker cardiomyocytes at the venous pole[7]. We also observed labeling in larger cells positioned lateral to the PA clusters (Fig. 4l). Because Kaede^{PC} fluorescence fades over time, we performed a second photoconversion in these larger cells of new 28 hpf embryos and analyzed their derivatives at 72 hpf (Fig. 4m, n). At this stage, Kaede^{PC} was broadly distributed throughout the pericardium (Fig. 4n), consistent with prior reports of ZsYellow fluorescence in the pericardium of *Tg(nkx2.5:ZsYellow)* transgenic embryos[5,27] and previous

lineage tracing analyses[6,29]. Although we were unable to computationally resolve late-differentiating atrial/pacemaker from pericardial cells in our datasets, our lineage tracing to both structures is consistent with their established derivation from the lateral-most ALPM[6,7,29]. Supporting this conclusion, MERFISH expression analyses showed that LPM-2 progenitors, which have largely downregulated *nkx2.5*, are enriched for *hand2* (Supplementary Fig. 9), a gene previously linked to mesothelial development, including the pericardium[6]. Overall, these data demonstrate that the spatial organization of *nkx2.5*+ populations in the ALPM corresponds with their future developmental trajectories[4–6,9].

Mutant analyses highlight functional requirements for LPM-enriched marker genes

Next, we performed mutant analyses to determine whether marker genes enriched in specific progenitor populations (Fig. 5a) are functionally required for the development of their end structures and to help link LPM subpopulations to their final fates. Each mutant was compared to 72 hpf controls (Fig. 5b–f) and evaluated for known derivatives of *nkx2.5*-expressing ALPM progenitors[4,5,9]. Specifically, we assessed ventricular chamber development (Fig. 5b, d) and PA2-derived structures, including the five HMs and OFT (Fig. 5b, c). PA1-derived HMs were also examined (Fig. 5b, c), despite their progenitors upregulating *nkx2.5* after ALPM stages[5]. Ventricular cardiomyocyte numbers, reflecting contributions from both the FHF and SHF, were quantified (Fig. 5d; Supplementary Fig. 10a). Finally, we analyzed the body vasculature, including the PAAs (Fig. 5e, f), and the pericardium (Supplementary Fig. 10e).

We first selected Follistatin a (Fsta), a potent BMP antagonist[30], because its transcripts were enriched in LPM-1a with limited expression in LPM-1b (Fig. 5g). We generated a stable *fsta* mutant line carrying a 14-base pair deletion predicted to truncate the protein (Supplementary Fig. 11a–c). Development appeared grossly normal aside from defects in retinal pigmentation (Supplementary Fig. 11d), likely resulting from excessive BMP signaling[31,32]. In contrast, PA2-derived structures arising from LPM1a and LPM1b were differentially affected. In *fsta* mutants, PA1- and PA2-derived HMs were largely absent, whereas the OFT was well-formed (Fig. 5h, i). LPM-1b derivatives were also differentially impacted: although the OFT was present, ventricular cardiomyocyte numbers were significantly reduced (Fig. 5h–j; and Supplementary Fig. 10b), likely reflecting a SHF deficiency. The body vasculature, PAAs, and pericardium developed normally (Fig. 5k, l; and Supplementary Fig. 10 f). These findings indicate that LPM-1a is the progenitor source of PA2-derived HMs, which require Fsta for their development on day 3. Similarly, *fsta* mutants from a previous study also showed defects in PA2-derived structures, including the hyoid joint and symplectic cartilage[33], suggesting this arch is particularly sensitive to Fsta deficiency. Our data also suggest that a subset of LPM-1b cells, which also express *fsta* (Supplementary Fig. 6b and 7 d), correspond to SHF ventricular progenitors.

Next, we selected *ephrinb2a* (*efnb2a*), which encodes a transmembrane ligand with known roles in angiogenesis[34]. *Efnb2a* transcripts were enriched in LPM-1a and LPM-1b with modest levels in LPM-3 (Fig. 5m). We generated a stable *efnb2a* mutant line harboring a 13-base pair deletion predicted to truncate the protein to 8 amino acids (Supplementary Fig. 12a–c). As anticipated from its expression pattern, *efnb2a* mutants (Supplementary Fig. 12 d) exhibited defects in PA1- and PA2-derived HM, ventricular cardiomyocyte numbers, and OFT development (Fig. 5n–p; and Supplementary Fig. 10c). In contrast, the body vasculature, PAAs, and pericardium formed normally (Fig. 5q, r; and Supplementary Fig. 10 g). These phenotypes support the conclusion that LPM-1a gives rise to PA2 HM progenitors, while LPM-1b contains SHF ventricular and OFT precursors, all of which are disrupted by Efnb2a deficiency on day 3.

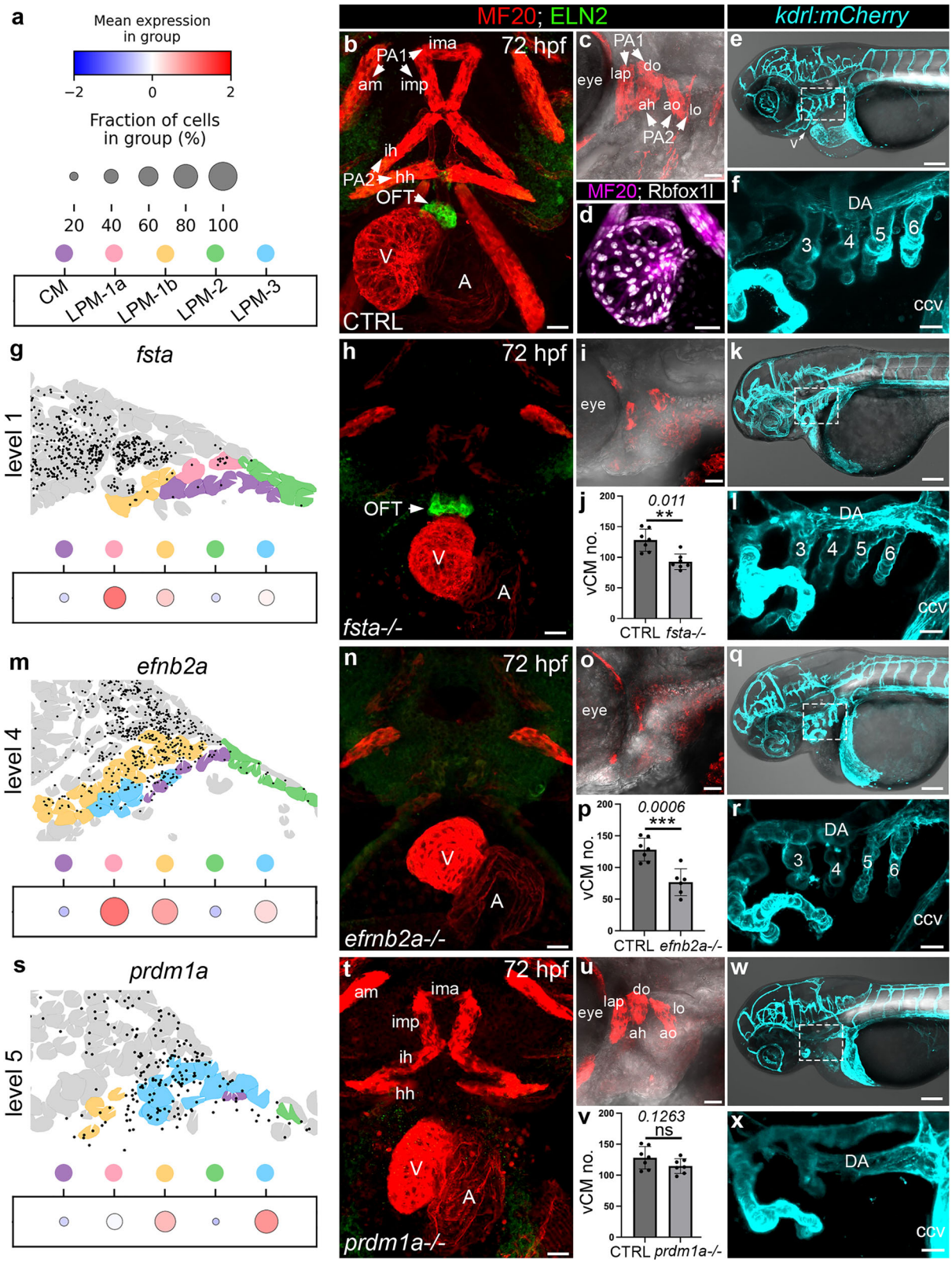
Finally, we characterized the cardiovascular phenotypes in a previously published mutant in *prdm1a*[35], because expression was enriched in posterior LPM-1b cells and LPM-3 (Fig. 5s; and Supplementary Fig. 6b; Supplementary Fig. 7c). Based on these expression patterns, we hypothesized that the HMs and ventricle might form normally, while defects in the OFT and PAAs might be observed. As anticipated, the HMs were well-developed in *prdm1a* mutants (Fig. 5t, u) and the ventricle was of normal size and cardiomyocyte number (Fig. 5v; and Supplementary Fig. 10 d). In addition, the OFT failed to form (Fig. 5t) and the PAAs were lacking (Fig. 5w, x) despite a normal body vasculature and pericardium (Fig. 5w; and Supplementary Fig. 10 h). This outcome suggests that more posterior LPM-1b progenitors, which have relatively higher *prdm1a* expression than those in the anterior, represent OFT precursors that are sensitive to loss of Prdm1a. Based on our collective findings, we conclude that LPM-1a are PA2 HM precursors, LPM-1b are SHF ventricular and OFT progenitors, LPM-2 are SHF atrial/pacemaker and pericardial progenitors, and LPM-3 are PAA endothelial progenitors (Fig. 6).

Discussion

Here, we present an initial characterization of the heterogeneous *nkx2.5*+ populations in the ALPM and describe their relative organization. Further, our data indicate that their spatial arrangement predicts their future migratory paths and end fates. Comparable pre-patterning has been documented in the gastrulating mouse embryo, where the sequence of myocardial progenitor cell specification in the nascent mesoderm correlates temporally with their migration through the primitive streak and ultimate arrangement in the cardiac crescent[36]. Notably, the relative positioning of myocardial progenitors in the mouse FHF, SHF, and juxtacardiac field (JCF), which contributes to the left ventricle[25,28,36,37], have been grossly mapped through a combination of morphometric analyses, single-cell RNA sequencing, and gene expression profiling. While an analogous JCF has yet to be defined in zebrafish, it is characterized by the lack of *nkx2.5* expression[28,37], suggesting this population would be missed in our analyses. While these previous studies have provided valuable insights into cardiovascular development, a 3D spatial map of the ALPM at single-cell resolution that contains progenitors of additional derivatives, including the HMs, OFT, PAAs, and pericardium[4,5,38–41], has been lacking. Our datasets fill this gap and open new avenues for interrogating how cardiac progenitor communities interact with each other and their microenvironments to pre-pattern the cardiovascular system.

Progenitor regionalization in the zebrafish ALPM is not only supported by our data but by a long history of lineage tracing studies employing increasingly sophisticated methods. Early photoactivation experiments with caged fluorescein relative to the notochord showed that the anterior half of the presumptive *nkx2.5*+ domain, now known to contain cardiomyocytes, gave rise to the heart, whereas posterior cells, now known to correspond to LPM-3, contributed mesenchyme beneath the otic vesicle[3], likely the PAAs. Using the same strategy, ventricular and atrial cardiomyocytes were found to be spatially segregated at the blastula stage, with ventricular precursors positioned closer to the organizer and margin than the atrial progenitors[42]. Fate mapping analyses performed at later developmental stages revealed that ventricular cardiomyocytes arise from medial ALPM cells co-expressing *nkx2.5* and *hand2*[1], while atrial cardiomyocytes were positioned more laterally[1,10] and expressed *hand2* alone[1]. Non-cardiac contributions to the PAs were also observed[1], presumably reflecting simultaneous labeling of LPM progenitor cells and underlying cardiomyocytes. A more recent *hand2* genetic lineage tracing study confirmed contributions to both cardiac chambers[43]. Our scRNA-seq and MERFISH data are consistent with these studies but show co-expression of *nkx2.5* with *hand2* in the cardiomyocyte population

https://doi.org/10.1038/s41467-025-66646-3



(Supplementary Fig. 9b), likely reflecting greater sensitivity compared to traditional methods. Because chamber-specific cardiomyocyte subpopulations were not detected in our scRNA-seq, chamber-specific markers were not included in our MERFISH probe panel. While regional assignments of ventricular versus atrial fates await further resolution, we anticipate their medial to lateral orientation will be corroborated.

Additional lineage tracing studies also support regionalization of non-myocardial progenitors. Fluorescein labeling of presumptive anterior and medial ALPM cells, where we identified LPM-1b, yielded contributions to the SHF-derived distal ventricle, OFT, and PAs[10]. Consistent with this outcome, a medial ALPM position for SHF progenitors was inferred from in situ hybridization analyses of marker genes identified by scRNA-seq of *gata5:GFP*+ cells[25]. Likewise, the

Fig. 5 | Mutant analyses highlight functional requirements for genes enriched in unique LPM populations. a Dot plot codes are shown in g,m,s where scale depicts mean expression level, dot size depicts the percentage of cells with expression, and color represents each subpopulation. **b**, **c** Confocal Z-stacks of representative control embryos immunostained for MF20 (red) and Elastin (green). Ventral view, anterior up (**b**); $n = 10$) and lateral view, anterior left (**c**) $n = 6$). Scale bars: 25 microns. **d** Confocal Z-stack of a representative control embryo immunostained for MF20 and Rbfox1l for ventricular CM quantifications in **j**, **p**, and **v** ($n = 6$). Scale bar: 25 microns. **e**, **f** Confocal Z-stack of lower and higher magnification views of a *Tg(kdrl:mCherry)* embryo immunostained for mCherry to highlight body vasculature (**e**) and PAAs (**f**). Lateral views, anterior left, $n = 12$. **g**, **m**, **s** Software-generated MERFISH sections showing color-coded ZsYellow+ LPM populations and individual transcripts (black dots). Dot plots depict expression levels in each subpopulation. **h**, **i**, **n**, **o**, **t**, **u** Confocal Z-stack of mutant embryos immunostained for MF20 (red) and Elastin (green). Ventral views, anterior up, $n = 5$ (**h**), $n = 8$ (**i**), $n = 7$ (**n**), $n = 9$ (**o**), $n = 10$ (**t**), $n = 10$ (**u**). **j**, **p**, **v** Dot plots of ventricular CM numbers. Each dot represents one biological replicate (1 embryo). **k**, **l**, **q**, **r**, **w**, **x** Confocal Z-stacks of lower and higher magnification views of mutants carrying the *Tg(kdrl:mCherry)* reporter following mCherry immunostaining to highlight the body vasculature (**k**, **q**, **w**) and PAAs (**l**, **r**, **x**). Lateral views; anterior left; $n = 6$ (**k**, **l**); $n = 7$ (**q**, **r**); $n = 15$ (**w**, **x**). Scale bars: 100 microns (**e**, **k**, **q**, **w**); 25 microns (**b**–**d**, **f**, **h**, **i**, **l**, **n**, **o**, **r**, **t**, **u**, **x**). Abbr: Ventral pharyngeal arch 1 (PA1) muscle: intermandibular anterior. Middle PA1 muscles: *imp* intermandibular posterior; *am* adductor mandibulae. Dorsal PA1 muscles: *lap* levator arcus palatine; *do* dilator opercula. Ventral pharyngeal arch 2 (PA2) muscle: *ih* interhyal. Middle PA2 muscle: *hh* hyohyal. Dorsal PA2 muscles: *ah* adductor hyomandiublae; *ao* adductor opercula; *lo* levator opercula. *OFT* outflow tract, *V* ventricle; *A* atrium; *DA* dorsal aorta; *CCV* common cardinal vein. *vCM no* ventricular cardiomyocyte number; numbers show PAAs. Statistics: Graphs show mean +/- SEM with two-sided student's t-tests used to quantify *p* values. No correction for multiple comparisons was applied. Source data are provided for ventricular cardiomyocyte counts as a Source Data file.

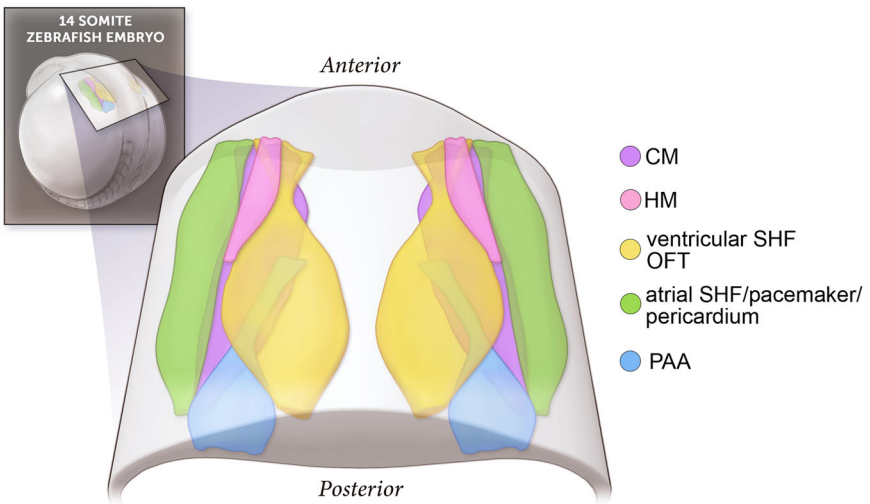


Fig. 6 | Schematic of *nkx2.5*+ progenitor cell organization in the ALPM. 3D illustration is based on serial sections that can be viewed in Supplemental Movie 1. While relative locations of each progenitor sub-type can be visualized, domain shapes are approximations. Abbr: *CM* cardiomyocyte; *HM* head muscle; *SHF* second heart field; *OFT* outflow tract; *PAA* pharyngeal arch artery. Artwork by Kai-ou Tang.

anterior and posterior tips of the *nkx2.5* + ALPM were reported to co-express *tcf21*, and their lineage tracing showed contributions to the HMs, OFT, and PAAs[27]. While these findings largely agree with ours, we observed a broader distribution of LPM-1b progenitors that represent SHF ventricular and OFT precursors along the A-P axis than previously described. Although we could not molecularly distinguish these progenitor sub-types in LPM-1b, deeper cell sampling should resolve their identities.

Beyond positional information, our transcriptomic and spatial datasets revealed potential signaling cues that likely contribute to cell fate establishment. For example, late-differentiating atrial cardiomyocytes, pacemaker cells, and pericardial progenitors within LPM-2 likely have differential requirements for Wnt signaling. A previous study found that pacemaker specification in the ALPM depends on canonical Wnt5b signaling[7], whereas a separate study reported that pericardial progenitors express *sfrp5*[6], an established Wnt antagonist[44]. We also detected enriched *sfrp5* expression in LPM-2 along with another Wnt agonist, *follistatin-like 1a* (*fstl1a*). Therefore, it is possible that differential Wnt activity segregates pacemaker from late-differentiating atrial cardiomyocyte and pericardial fates. We also identified potential interplay between BMP signaling and HM specification. Specifically, LPM-1a HM progenitors express the BMP antagonist *fsta*[30], while neighboring LPM-2 progenitors express BMP ligands. This spatial arrangement raises the possibility that Fsta protects HM fates by buffering BMP signals from adjacent LPM-2 cells. Together, these data highlight the potential importance of intercellular signaling both within and between progenitor communities in regulating cell identity.

In summary, our study provides an initial framework for the transcriptional and spatial organization of ALPM progenitors. However, lineage relationships among ALPM subtypes remain unresolved, including when progenitors become fate restricted. Although a study reported that single cells labeled at the 18-ss (cardiac cone stage) contribute to multiple cardiovascular derivatives[29], this conclusion requires confirmation with additional approaches such as barcode lineage tracing. Looking ahead, more comprehensive spatial datasets coupled with advanced computational pipelines (e.g., NicheNet[45], CellChat[46], and MEBOCOST[47]) will be invaluable for uncovering intercellular communication pathways that regulate progenitor allocation to LPM subtypes and establish boundaries between progenitor cell compartments. Ultimately, defining the heterogeneity and spatial logic of ALPM progenitor populations will deepen our understanding of vertebrate heart development and inform how disruptions in these processes contribute to congenital heart disease.

Methods

Zebrafish husbandry and strains

Zebrafish (Danio rerio) were bred and maintained following animal protocols approved by the Institutional Animal Care and Use Committees at Boston Children's Hospital. All protocols and procedures

followed the guidelines and recommendations outlined by the Guide for the Care and Use of Laboratory Animals. All experimental data was obtained from zebrafish embryos prior to sex determination. The following zebrafish strains were used in this study: *TgBAC(nkx2.5:ZsYellow)*^{fb7} [8], *Tg(nkx2.5:Kaede)*^{fb9} [9], *Tg(kdrl:mCherry)*^{s896} [48], *TgBAC(hand2:EGFP)*^{pd24} [49], *prdm1a*^{m805} [50], *fsta*^{chb11} (this study), and *efnb2*^{chb12} (this study). Zebrafish lines generated in the Burns lab are available upon request.

Isolation of *nkx2.5:ZsYellow*+ cells

14- and 20-ss wildtype (TU) and *Tg(nkx2.5:ZsYellow)* embryos were dissociated to single cells as previously described[51]. Briefly, 750–1000 embryos were collected and their chorions removed by incubation in 1 mg/ml Pronase for 3–4 min followed by washing in 0.5X Danieau Buffer. Yolks were removed in deyolking buffer followed by embryo dissociation in FACSmax Cell Dissociation solution. After passage through a 40 µm strainer, single cells were resuspended in fresh FACSmax Cell Dissociation solution and passed through a fluorescent activated cell sorter (BD FACSMelodyTM Cell Sorter) to separate the ZsYellow positive fraction. Sorted cells were collected in 0.1% BSA D-PBS.

Single-cell RNA sequencing

Single-cell RNA sequencing (scRNA-seq) of ZsYellow+ cells was performed using the inDrops platform following the user guide. Briefly, the cells were encapsulated in droplets as previously described[16,52] with the following primer sequence modifications that prevent the need for custom primers:

RT primers on hydrogel beads:

5′CGATTGATCAACGTAATACGACTCACTA-TAGGGTGTCGGGTGCAG[bc1,8nt]GTCTCGTGGGCTCGGA-GATGTGTATAAGAGACAG[bc2,8nt] NNNNNNTTTTTTTTTTTTTTTTTTV- 3′

PCR primer sequences (steps 157 and 160 in the library prep protocol from[16]):

5′-AATGATACGGCGACCACCGAGATCTA-CACXXXXXXXXTCGTCGGCAGCGTC-3′, where XXXXXX is an index sequence for multiplexing libraries.

5′- CAAGCAGAAGACGGCATACGAGATGGGTGTCGGGTGCAG-3′

The library prep protocol was based on SMART-Seq protocols[53] with modifications. The workflow of the library preparation can be summarized as follows: Reverse transcription with Maxima H Minus RTase at 50 °C for 60 min, followed by enzyme inactivation at 85 °C for 3 min -> post droplet breaking cleanup with MyOne SILANE beads -> Template switching using Maxima H Minus RTase (incubate for 30′ at 25 °C, 90′ at 42 °C and hold at 4 °C; with lid set to 50 °C.) -> 1X SPRI purification -> cDNA amplification with PCRI Fwd Primer V2 and PCRI Rev Primer -> 0.6X SPRI purification. The gene expression libraries were made by subjecting the amplified cDNA to Fragmentation, end repair and A-tailing, followed by ligation using NEBNext UltraII FS DNA Library Prep kit for Illumina -> 0.8X SPRI cleanup -> sample indexing by PCR -> double-sided SPRI cleanup.

scRNA-seq (inDrops) bioinformatics analysis

Raw sequencing data were converted to FASTQ files using bcl2fastq. 25 bases were trimmed from the 3′ end of read 1 using cutadapt[54]. Next, the zUMIs pipeline[55] was used to demultiplex the sequencing data, remove reads with low-quality cell barcodes or UMIs, map the filtered reads to the zebrafish genome (version danRer11) plus *ZsYellow*, and generate count matrices, which were processed by the Seurat package (version 4.3.0)[56]. Cells with less than 500 genes detected, more than 3000 genes detected, or more than 20% mitochondrial reads were filtered from the analysis. In addition, transcripts encoding ribosomal subunits were also removed before log scaling, data normalization, principal component analysis (PCA), cell clustering, and UMAP visualization.

Datasets from each developmental timepoint were integrated using canonical correlation analysis (CCA)[57]. Integrated data were then used for data scaling, PCA, cell clustering, and UMAP visualization. Differentially expressed genes between cell clusters were identified by the FindAllMarkers function using the default Wilcoxon test method. The heatmap showing the top 10 most highly expressed genes in each cluster was generated with the DoHeatmap function. Cell types were annotated based on the top differentially expressed genes. After initial annotation, all cells in the LPM cluster were selected and reprocessed through data scaling, PCA, clustering, and UMAP embedding. Differentially expressed genes among the three resulting subclusters were identified using the FindAllMarkers function (Wilcoxon test), and subcluster identities were assigned based on top marker genes. CMs were not included in the subclustering because they expressed only myocardial markers and did not show further heterogeneity. The expression of selected marker genes for each LPM subcluster were visualized using the VlnPlot function. Gene Ontology (GO) enrichment analysis was performed for marker genes in each LPM subcluster using the clusterProfiler package (version 4.6.2)[58].

MERFISH Spatial Transcriptomics

MERFISH[22] was performed using the commercialized MERSCOPE platform (Vizgen, MA, US). For the 133 probe panel, we selected the top 7–15 most specifically expressed genes identified for each cell cluster of our scRNA-seq samples ($n = 94$ genes), 35 genes from a previously published scRNA-seq of *gata5*:GFP^{+} cells from 8-ss zebrafish embryos[25], 4 genes previously reported in the literature, and *ZsYellow*. Probes were designed with the Vizgen Gene Panel Portal (https://portal.vizgen.com) and manufactured by Vizgen.

14-ss *Tg(nkx2.5:ZsYellow)* embryos were fixed overnight with 4% paraformaldehyde (PFA) at 4 °C and rinsed the following day in PBSTw (0.1% Tween-20 in PBS). Embryos were dehydrated in a graduated PBSTw:methanol series and stored in 100% methanol at -20 °C. Embryos were cleared in Butyl alcohol (1-Butanol) for 10 min and embedded in paraffin wax. 10 µm transverse embryonic sections were cut, floated in a nuclease-free water bath, mounted onto MERSCOPE Slides, and incubated for 4 hours at RT prior to storing at -20 °C. MERFISH was performed according to the manufacturer's instructions for formalin-fixed paraffin-embedded (FFPE) tissues. Briefly, autofluorescence was quenched from sections with MERSCOPE Photobleacher for three hours. Quenched sections were stained with anti-β-catenin and anti-ZsYellow antibodies using the Vizgen Staining Kit and counterstained with DAPI. After 48 hours of probe incubation at 37 °C, the sections were covered with glass plate coating gel for tissue clearing following the manufacturer's guidelines for non-resistant samples. Imaging was performed on the MERSCOPE Platform using a 140-gene imaging kit.

MERFISH bioinformatics analysis

Two slides, each containing 10 sections, were processed using the MERSCOPE™ Image Processing Software, which generated a list of all detected transcripts (CSV format), raw images (TIFF format), and image metadata (CSV format) for four staining channels, including DAPI, Poly T, anti-β-catenin, and anti-ZsYellow.

For cell segmentation, DAPI was used to identify cell nuclei, ß-catenin to delineate cell boundaries and shapes, and ZsYellow to identify the *nkx2.5* lineage in the ALPM. Based on the raw images and metadata from these three staining channels, cell boundaries were defined using the Cellpose algorithm (version 1.0.2)[59] within the Vizgen Post-processing Tool vpt (version 1.3.0). Next, these defined boundaries were used to group transcripts into cells using vpt. Subsequently, cell metadata, including location, size, and shape, were calculated, and

cell signals, such as the brightness of each mosaic TIFF image within each cell, were summarized. Finally, we used vpt to adjust the cell boundaries in the Vizgen MERSCOPE™ Visualizer. All downstream analyses were performed directly on the segmentation coordinates obtained from cellpose_micron_space.parquet. The sharp, inward-pointing concavities created by MERSCOPE_Vizualizer are visualization artifacts and do not reflect the true segmentation; the raw coordinates show smooth and continuous boundaries.

After cell segmentation, we obtained a list of files including the cells-by-genes file (CSV format), cell metadata file (CSV format), and sum signal file (CSV format). ZsYellow$^+$ cells were selected through UMAP clustering based on these files by the following steps:

(1) The sum signal file was used to analyze ZsYellow$^+$ signals, which were detected in 11 of 20 regions. The remaining 9 regions showing no ZsYellow signals were discarded.

(2) ZsYellow+ cells from these remaining 11 regions were pooled.

(3) The Python package Scanpy (version 1.9.3)[60] was used to conduct default preprocessing, including quality control, normalization, scaling, and dimensionality reduction. Cells with fewer than three detected transcripts total and genes with only one transcript observed were filtered out. Transcript counts were normalized per cell using Scanpy, followed by logarithmic transformation, principal component analysis, and computation of a neighborhood graph.

(4) The UMAP method[61] within Scanpy was used to embed the neighborhood graph and cluster cells into subgroups.

(5) Cell cluster annotation was visualized on UMAP and spatial coordinates using the Python package Squidpy (version 1.3.1)[62]. Spatial coordinates were mapped to cell clusters for each region, and clusters were selected based on spatial distribution.

(**6**) Steps 4 and 5 were repeated twice to refine ZsYellow+ cell selection, resulting in 1,032 ZsYellow$^+$ cells for downstream analysis.

Next, cell types were annotated using marker genes and the Python package Decoupler (version 1.6.0)[63]. Single-cell RNA-seq data provided cell type-specific markers. To assign cell type labels to single-cell clusters, we identified communities of cells with similar transcriptional profiles and checked for the expression of cell type-specific markers. Functional enrichment scores, measuring the over-representation of gene sets at the top or bottom of a ranked list, were inferred using the Over Representation Analysis (ORA) method within Decoupler[64]. The built-in function rank_sources_groups in Decoupler was used to identify the top predicted cell types per cluster. This process enabled the annotation of single-cell clusters with specific cell types.

3D reconstruction

We reconstructed the three-dimensional organization of the zebrafish embryo dataset (13 serial sections) using PASTE2 (v1.0.1)[65]. Sections were ordered anterior to posterior, and pairwise alignment between consecutive sections was performed with the partial_pairwise_align function, using a similarity threshold of 0.2 for preliminary alignment and 0.3 for refinement. The resulting pairwise alignments were integrated with the PASTE2 projection module, which iteratively stacks sections into a common coordinate space while preserving spatial correspondence. To generate a 3D point cloud, spatial coordinates and cluster annotations were exported in VTK format using pyevtk (v1.6.0). Two-dimensional spatial coordinates for each section were augmented with section-specific z-values reflecting physical section spacing, and the concatenated dataset was written to a.vtk file via the pointsToVTK function. Visualization and final 3D renderings were produced in ParaView (v5.11), where points were displayed by cluster identity and rendered with ray tracing.

Zebrafish whole mount immunofluorescence

Embryos less than 24 hpf were fixed in PBS containing 4% paraformaldehyde (PFA) overnight at 4 °C. Embryos greater than 24 hpf were transferred to E3 containing 0.003% Phenylthiourea (1-phenyl-2-thiourea; PTU) to inhibit pigment formation. At the desired developmental stage, embryos were anesthetized with 0.16% Tricaine and fixed overnight at 4 °C in PBS containing 4% paraformaldehyde (PFA). For all stages, embryos were dehydrated and rehydrated the next day in a PBSTw:methanol series and antigen retrieval performed in 150 mM Tris-HCl buffer (pH 9) for 15 min at 70 °C. After washing with PBSTw, embryos were placed in blocking solution (5% FBS, 1% BSA in PBSTw) for 2 hours at room temperature and incubated at 4 °C overnight with primary antibodies diluted in blocking solution. The following primary antibodies were used: anti-ZsYellow (rabbit polyclonal, Takara Bio; 1:50), anti-myosin heavy chain (MF20 mouse monoclonal, Developmental Hybridoma Bank, University of Iowa; 1:50), anti-Elastin-2 (*34*) (gift from Fred Keeley; 1:500), and anti-Rbfox1l[66] (gift from S. Amacher; 1:200). Embryos were washed in PBSTw and incubated with Alexa-Fluor conjugated secondary antibodies (1:500; Thermo Fisher Scientific) overnight at 4 °C. The following secondary antibodies were used: Goat anti-Mouse IgG 546, Goat anti-Rabbit IgG 546, Goat anti-Mouse IgG2a 488, and Goat anti-Rabbit IgG 488. Following incubation, embryos were washed in PBSTw and nuclei labeled with 1 μg/ml DAPI in PBSTw for 5 minutes before imaging.

RNAscope whole mount in situ hybridization

14-ss embryos were fixed for 1 hour in 4% PFA, washed with PBSTw (PBS with 0.1% Tween), dehydrated in methanol, and kept at -20 °C overnight. Whole mount fluorescent in situ hybridization was performed using the RNAscope Multiplex Fluorescent Reagent Kit (Advanced Cell Diagnostics) as previously described[67]. RNAscope riboprobes to *cyp26c1*, *cdh11* and *aldh1a2* were hybridized overnight at 40 °C at a 50:1:1 ratio, respectively. Following the final labeling reaction, embryos were incubated with ZsYellow primary antibody overnight at 4 °C, rinsed the following day in PBSTw for several hours, and incubated with the secondary antibody overnight at 4 °C. Embryos were washed in PBSTw prior to imaging.

Kaede photoconversion

Tg(nkx2.5:Kaede)^{fb9} embryos were photoconverted by exposing Kaede protein to fluorescent light passed through a DAPI filter on an Olympus FV3000R resonant scanning confocal microscope with a 20X objective. 14-somite stage (16 hpf) and 28 hpf embryos were mounted in 1% low melting point agarose (Sigma-Aldrich) in a 35 mm MatTek glass bottom dish (MatTek Corporation). Prior to photoconversion, embryos were imaged using a 488 laser to detect green fluorescence. Embryonic regions were exposed to UV light using the 405 mm laser for 1 min. Embryos were immediately imaged using the GFP and RFP filters, removed from agarose, arrayed individually in six-well plates, and incubated in the dark until 28 hpf or 72 hpf when they were imaged again by confocal microscopy. During imaging, embryos were anesthetized in 0.16% Tricaine to stop cardiac contractions.

Microscopy

For compound microscopy, embryos were anesthetized in 0.16% Tricaine, mounted in 1% low melting agarose dissolved in E3 medium on depression slides, and imaged using fluorescence and/or brightfield optics on a Nikon Eclipse 80i compound microscope with a QImaging Retiga 2000R CCD camera (QImaging).

For live imaging on a confocal microscopy, embryos carrying fluorescent reporters were anesthetized in 0.16% Tricaine, mounted on 35 mm MatTek glass bottom dishes in 1% low melting point agarose containing 0.16% Tricaine, and covered in E3 containing 0.4% Tricaine. For confocal imaging of fixed embryos, animals were embedded in 1% low melting point agarose and mounted on 35 mm MatTek glass bottom dishes. Z-stack images were taken using 405 nm, 488 nm, 561 nm, 641 nm lasers on a Olympus FV3000R resonant scanning confocal microscope and processed using Fiji software[68].

Ventricular cardiomyocyte counts

Embryos were fixed, immunostained for Rbfox1l to mark striated muscle nuclei[69,70] and MF20 to mark striated muscle cells and imaged by confocal microscopy. For all animals, Rbfox1L+ nuclei within the ventricular chamber proper, defined as the region between the myocardial collar of the outflow tract and the atrioventricular canal, were counted.

Generation and genotyping of zebrafish mutants lines

Mutant alleles in *fsta* and *efnb2a* were generated using previously published sgRNAs for *fsta* 5′-GGTGGATGATCTTCAATGGCGGG-3′ and *efnb2a* 5′-TCTTTGTGGAGATATTACTTTGG-3′ (PAM site underlined)[33,71]. Synthetic sgRNA stocks were prepared as described[72]. A solution consisting of 1 µl of 25 mM sgRNA, 1 µl 25 mM Cas9 protein, and 3 µl nuclease free-H_2O was incubated at 37 °C for 5 min and then placed at room temperature before injecting 1 nL into embryos at the one-cell stage. Germ-line transmission of CRISPR/Cas9-induced deletions were detected by PCR. For *fsta*, 14 base pairs (TGGCGGGGCTCCAA) in exon 2 were deleted, which led to loss of an *Aci1* restriction site. For *efnb2a*,13 base pairs (ATTACTTTGGAGT) in exon 1 were deleted, which lead to loss of a *BsaXi* restriction site. For genotyping, genomic DNA was amplified by PCR, the products digested with *Aci1* or *BsaX1*, and the banding pattern analyzed on a 2% agarose gel. For *fsta*, the PCR primers 5′- TGGGTCATTGGCTTCAGACT-3′ and 5′-TCCAGGTGACGTTAGAGCAG-3′ were used to produce a 562 bp wildtype amplicon and a 548 bp mutant amplicon. For *efnb2a*, the PCR primers: 5′-CCCGCGAAACGGAATTTGACCG -3′ and 5′-CTAAATGAACGAGTTGTCGCC -3′ were used to produce a 309 bp wildtype amplicon and a 296 bp mutant amplicon.

Artwork

The artwork in Fig. 2c was drawn and labeled entirely in Adobe Photoshop. Section placement was estimated visually. The artwork in Fig. 6 was created using Adobe Photoshop, Adobe Illustrator, and Cinema 4D. First, the image was layered and visually aligned in Adobe Photoshop. The alignment was done by lowering the opacity of each section layer and visually matching the overlapping anatomical regions. Using the Pen tool, each colorized population was outlined broadly and labeled, resulting in 4-6 splines per section. An approximate spline outline of the zebrafish embryo in the transverse plane was also drawn. The Photoshop file was then opened in Adobe Illustrator as an.ai file and saved. The Illustrator file was opened in Cinema 4D and splines imported into the C4D Viewport. Each group of splines, separated by section, were moved 4 cm apart along the *y* axis. This distance was the best visual match in comparison to the 3-D rendering of MERFISH sections. Using the “Extrude” modifier, the embryo outline was extruded to a length of 16 cm. The “Loft” modifier in Cinema 4D was used to “connect” each progenitor population spline level together such that they made several 3-D objects spanning sections 1 to 5. Each progenitor population 3-D object was assigned a Material corresponding to the color on the original reference sections. A dorsal view of the 3D object was rendered. The zebrafish embryo was illustrated in Adobe Photoshop and the 3-D render was added to it as a composite callout image.

Statistics

Data are expressed as the mean +/- 1 SD of at least 3 independent experiments. Statistical significance was evaluated by unpaired two-tailed Student’s t-test in GraphPad Prism 10. $P < 0.05$ was considered significant. All sample sizes and *P* values are reported in the figure or figure legends. No statistical method was used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized. The Investigators were blinded to genotype during experiments and outcome assessments.

Reporting summary

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

Data availability

All data are available in the main text or the supplementary materials. The single-cell RNA sequencing data generated in this study have been deposited in the GEO database under accession code GSE294086. The MERFISH data generated in this study have been deposited in the GEO database under accession number GSE294469. Source data are provided with this paper.

Code availability

The full code used for cell segmentation in this study has been deposited on GitHub and is openly accessible at (https://github.com/Yunxia-wang/MerfishCardioMap) and (https://doi.org/10.5281/zenodo.17419819[73]).

References

1. Schoenebeck, J. J., Keegan, B. R. & Yelon, D. Vessel and blood specification override cardiac potential in anterior mesoderm. *Developmental cell* **13**, 254–267 (2007).
2. Chen, J. N. & Fishman, M. C. Zebrafish tinman homolog demarcates the heart field and initiates myocardial differentiation. *Dev. (Camb., Engl.)* **122**, 3809–3816 (1996).
3. Serbedzija, G. N., Chen, J.-N. & Fishman, M. C. Regulation in the heart field of zebrafish. *Development* **125**, 1095–1101 (1998).
4. Paffett-Lugassy, N. et al. Heart field origin of great vessel precursors relies on nkx2.5-mediated vasculogenesis. *Nat. cell Biol.* **15**, 1362–1369 (2013).
5. Paffett-Lugassy, N. et al. Unique developmental trajectories and genetic regulation of ventricular and outflow tract progenitors in the zebrafish second heart field. *Dev. (Camb., Engl.)* https://doi.org/10.1242/dev.153411 (2017).
6. Prummel, K. D. et al. Hand2 delineates mesothelium progenitors and is reactivated in mesothelioma. *Nat. Commun.* **13**, 1677 (2022).
7. Ren, J. et al. Canonical Wnt5b signaling directs outlying Nkx2.5+ mesoderm into pacemaker cardiomyocytes. *Dev. Cell* 50, 729–743 (2019).
8. Zhou, Y. et al. Latent TGF-β binding protein 3 identifies a second heart field in zebrafish. *Nature* **474**, 645–648 (2011).
9. Guner-Ataman, B. et al. Zebrafish second heart field development relies on progenitor specification in anterior lateral plate mesoderm and nkx2.5 function. *Dev. (Camb., Engl.)* **140**, 1353–1363 (2013).
10. Hami, D., Grimes, A. C., Tsai, H.-J. & Kirby, M. L. Zebrafish cardiac development requires a conserved secondary heart field. *Dev. (Camb., Engl.)* **138**, 2389–2398 (2011).
11. Pater et al. Distinct phases of cardiomyocyte differentiation regulate growth of the zebrafish heart. *Dev. (Camb., Engl.)* **136**, 1633–1641 (2009).
12. Lazic, S. & Scott, I. Mef2cb regulates late myocardial cell addition from a second heart field-like population of progenitors in zebrafish. *Dev. Biol.* **356**, 172 (2011).
13. Martin, K. E. & Waxman, J. S. Atrial and sinoatrial node development in the zebrafish heart. *J. Cardiovasc. Dev. Dis.* **8**, 15 (2021).
14. Rodriguez, E. R. & Tan, C. D. Structure and anatomy of the human pericardium. *Prog. Cardiovasc. Dis.* **59**, 327–340 (2017).
15. Abrial, M. et al. TGF-β signaling is necessary and sufficient for pharyngeal arch artery angioblast formation. *Cell Rep.* **20**, 973–983 (2017).
16. Zilionis, R. et al. Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44–73 (2017).
17. Hashimshony, T. et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol.* **17**, 77 (2016).

18. Duong, T. B., Holowiecki, A. & Waxman, J. S. Retinoic acid signaling restricts the size of the first heart field within the anterior lateral plate mesoderm. *Dev. Biol.* **473**, 119–129 (2021).
19. Zhao, J. & Mommersteeg, M. T. M. Slit–Robo signalling in heart development. *Cardiovasc. Res.* **114**, 794–804 (2018).
20. Paskaradevan, S. & Scott, I. C. The Aplnr GPCR regulates myocardial progenitor development via a novel cell-non-autonomous, Gαi/o protein-independent pathway. *Biol. Open* **1**, 275–285 (2012).
21. Döring, Y., Pawig, L., Weber, C. & Noels, H. The CXCL12/CXCR4 chemokine ligand/receptor axis in cardiovascular disease. *Front. Physiol.* **5**, 212 (2014).
22. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
23. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* **100**, 106 (2021).
24. Ford, C. et al. Nkx2.7 is a conserved regulator of craniofacial development. *Nat. Commun.* **16**, 3802 (2025).
25. Song, M. et al. GATA4/5/6 family transcription factors are conserved determinants of cardiac versus pharyngeal mesoderm fate. *Sci. Adv.* **8**, eabg0834 (2022).
26. Guner-Ataman, B. et al. Failed progenitor specification underlies the cardiopharyngeal phenotypes in a zebrafish model of 22q11.2 deletion syndrome. *Cell Rep.* **24**, 1342–1354.e5 (2018).
27. Nagelberg, D. et al. Origin, specification, and plasticity of the great vessels of the heart. *Curr. Biol.: CB* **25**, 2099–2110 (2015).
28. Tyser, R. C. V. et al. Characterization of a common progenitor pool of the epicardium and myocardium. *Science* **371**, eabb2986 (2021).
29. Mao, L. M. F., Anderson, E. A. T. B. & Ho, R. K. Anterior lateral plate mesoderm gives rise to multiple tissues and requires tbx5a function in left-right asymmetry, migration dynamics, and cell specification of late-addition cardiac cells. *Dev. Biol.* **472**, 52–66 (2021).
30. Mulloy, B. & Rider, C. C. Chapter two the bone morphogenetic proteins and their antagonists. *Vitam. Horm.* **99**, 63–90 (2015).
31. Zhu, D., Wu, J., Spee, C., Ryan, S. J. & Hinton, D. R. BMP4 mediates oxidative stress-induced retinal pigment epithelial cell senescence and is overexpressed in age-related macular degeneration*. *J. Biol. Chem.* **284**, 9529–9539 (2009).
32. Zhu, D., Deng, X., Xu, J. & Hinton, D. R. What determines the switch between atrophic and neovascular forms of age related macular degeneration? - the role of BMP4 induced senescence. *Aging (Albany NY)* **1**, 740–745 (2009).
33. Askary, A. et al. Genome-wide analysis of facial skeletal regionalization in zebrafish. *Development* **144**, 2994–3005 (2017).
34. WEINSTEIN, B. M. & LAWSON, N. D. Arteries, Veins, Notch, and VEGF. *Cold Spring Harb. Symp. Quant. Biol.* **67**, 155–162 (2002).
35. Hernandez-Lagunas, L. et al. Zebrafish narrowminded disrupts the transcription factor prdm1 and is required for neural crest and sensory neuron specification. *Dev. Biol.* **278**, 347–357 (2005).
36. Dominguez, M. H., Krup, A. L., Muncie, J. M. & Bruneau, B. G. Graded mesoderm assembly governs cell fate and morphogenesis of the early mammalian heart. *Cell* **186**, 479–496.e23 (2023).
37. Zhang, Q. et al. Unveiling Complexity and Multipotentiality of Early Heart Fields. *Circ. Res.* **129**, 474–487 (2021).
38. Swedlund, B. & Lescroart, F. Cardiopharyngeal Progenitor Specification: Multiple Roads to the Heart and Head Muscles. *Csh Perspect. Biol.* **12**, a036731 (2019).
39. Lescroart, F. et al. Clonal analysis reveals common lineage relationships between head muscles and second heart field derivatives in the mouse embryo. *Dev. (Camb., Engl.)* **137**, 3269–3279 (2010).
40. Buckingham, M., Meilhac, S. & Zaffran, S. Building the mammalian heart from two sources of myocardial cells. *Nat. Rev. Genet.* **6**, 826–835 (2005).
41. Kelly, R. G., Brown, N. A. & Buckingham, M. E. The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm. *Dev. Cell* **1**, 435–440 (2001).
42. Keegan, B. R., Meyer, D. & Yelon, D. Organization of cardiac chamber progenitors in the zebrafish blastula. *Dev. (Camb., Engl.)* **131**, 3081–3091 (2004).
43. Ming, Z. et al. Lineage labeling with zebrafish hand2 Cre and CreERT2 recombinase CRISPR knock-ins. *Dev. Dyn.* https://doi.org/10.1002/dvdy.70022 (2025).
44. Cruciat, C.-M. & Niehrs, C. Secreted and transmembrane Wnt inhibitors and activators. *Cold Spring Harb. Perspect. Biol.* **5**, a015081 (2013).
45. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).
46. Jin, S. et al. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
47. Cui, K. et al. Epsin nanotherapy regulates cholesterol transport to fortify atheroma regression. *Circ. Res.* **132**, e22–e42 (2023).
48. Chi, N. C. et al. Foxn4 directly regulates tbx2b expression and atrioventricular canal formation. *Genes Dev.* **22**, 734–739 (2008).
49. Kikuchi, K. et al. Retinoic acid production by endocardium and epicardium is an injury response essential for zebrafish heart regeneration. *Developmental cell* **20**, 397–404 (2011).
50. Artinger, K. B., Chitnis, A. B., Mercola, M. & Driever, W. Zebrafish narrowminded suggests a genetic link between formation of neural crest and primary sensory neurons. *Development* **126**, 3969–3979 (1999).
51. 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).
52. Klein, A. M. et al. Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell* **161**, 1187–1201 (2015).
53. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
54. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetJ* **17**, 10–12 (2011).
55. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *Gigascience* https://doi.org/10.1101/153940 (2018).
56. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
57. 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).
58. Yu, G. Thirteen years of clusterProfiler. *Innov* **5**, 100722 (2024).
59. Pachitariu, M. & Stringer, C. Cellpose 2.0: how to train your own model. *Nat. Methods* **19**, 1634–1641 (2022).
60. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
61. Dorrity, M. W., Saunders, L. M., Queitsch, C., Fields, S. & Trapnell, C. Dimensionality reduction by UMAP to visualize physical and genetic interactions. *Nat. Commun.* **11**, 1537 (2020).
62. Palla, G. et al. Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022).
63. Badia-i-Mompel, P. et al. decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform. Adv.* **2**, vbac016 (2022).
64. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1–13 (2009).
65. Liu, X., Zeira, R. & Raphael, B. J. Partial alignment of multislice spatially resolved transcriptomics data. *Genome Res* **33**, 1124–1132 (2023).

66. Berberoglu, M. A. et al. Satellite-like cells contribute to pax7-dependent skeletal muscle repair in adult zebrafish. *Dev. Biol.* **424**, 162–180 (2017).
67. Gross-Thebing, T., Paksa, A. & Raz, E. Simultaneous high-resolution detection of multiple transcripts combined with localization of proteins in whole-mount embryos. *BMC Biol.* **12**, 55 (2014).
68. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. methods* **9**, 676–682 (2012).
69. Huang, M. et al. Intrinsic myocardial defects underlie an Rbfox deficient zebrafish model of hypoplastic left heart syndrome. *Nat. Commun.* **13**, 5877(2022).
70. Gallagher, T. L. et al. Rbfox-regulated alternative splicing is critical for zebrafish cardiac and skeletal muscle functions. *Developmental Biol.* **359**, 251–261 (2011).
71. Chen, J. et al. Acute brain vascular regeneration occurs via lymphatic transdifferentiation. *Dev. Cell* **56**, 3115–3127.e6 (2021).
72. Hoshijima, K. et al. Highly efficient CRISPR-Cas9-based methods for generating deletion mutations and F0 embryos that lack gene function in Zebrafish. *Dev. Cell* **51**, 645–657.e4 (2019).
73. Coskun et al. 3D spatial organization of heterogeneous *nkx2.5*+ progenitors in the zebrafish heart field pre-patterns cardiovascular development. *Github* https://doi.org/10.5281/zenodo.17419819 (2025).

Acknowledgements

We would like to thank the Boston Children's Hospital aquatics team for their care of the zebrafish and the Single Cell Core at Harvard Medical School, Boston, MA for performing the single-cell RNA sequencing sample preparation. Part of this work was performed in conjunction with the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD), which is funded by grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA). Additional information about CIFASD can be found at (www.cifasd.org). We also acknowledge Kai-ou Tang for the artwork illustrations shown in Figs. 2 and 6. HC was supported by a National Institutes of Health (NIH) training grant awarded to Boston Children's Hospital (BCH; T32HL007572; Principal Investigator (PI), William T. Pu); The K.C. lab is supported in part by NIH grants R01GM138407 (PI: K.C.), R01GM125632 (PI: K.C.), R01HL148338 (PI: K.C.), R01CA278832 (PI: K.C.), R01HL133254 (PI: K.C.). This work was primarily supported by R35HL135813 (PI: C.E.B.) and R01HL171206 (C.E.B.) with contributions from U01AA030185 (PIs: C.E.B. and C.G.B.). The Chen and Burns laboratories were also supported by philanthropic funds from the Boston Children's Hospital Department of Cardiology.

Author contributions

H.C. designed, performed, and interpreted the scRNA-seq, MERFISH, and wet bench experiments and interpreted bioinformatic analyses; Y.W. analyzed the spatial transcriptomics analyses and integrated the findings with the scRNA-seq analyses; X.G. analyzed the scRNA-seq data; K.C. oversaw all bioinformatics analyses. K.C., C.G.B., and C.E.B. jointly supervised the study, acquired funding, designed experiments, and interpreted data. H.C., C.G.B., and C.E.B. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-66646-3.

Correspondence and requests for materials should be addressed to Kaifu Chen, C. Geoffrey Burns or Caroline E. Burns.

Peer review information *Nature Communications* thanks Zexian Zeng and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.