Multi-modal refinement of the human heart atlas during the first gestational trimester

Christopher de Bono, Yichi Xu, Samina Kausar, Marine Herbane, Camille Humbert, Sevda Rafatov, Chantal Missirian, Mathias Moreno, Weiyang Shi, Yorick Gitton, et al.

Multi-modal refinement of the human heart atlas during the first gestational trimester

Authors

Christopher De Bono[1], Yichi Xu[2,*], Samina Kausar[1,*,£], Marine Herbane[1], Camille Humbert[1], Sevda Rafatov[1], Chantal Missirian[1,3], Mathias Moreno[1], Weiyang Shi[4], Yorick Gitton[5], Alberto Lombardini[6], Ivo Vanzetta[6], Séverine Mazaud-Guittot[7], Alain Chédotal[5], Anaïs Baudot[1], Stéphane Zaffran[1,✪] and Heather C. Etchevers[1,✪]

Affiliations

[1] Aix Marseille Univ, INSERM, MMG (Marseille Medical Genetics), Marseille, France

[2] Department of Systems Biology for Medicine and Frontier Innovation Center, School of Basic Medical Sciences, Fudan University, Shanghai 200032, China

[3] Medical Genetics Department, Assistance Publique Hôpitaux de Marseille, La Timone Children's Hospital, Marseille, France

[4] Department of Laboratory Medicine, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

[5] INSERM, CNRS, Institut de la Vision, Sorbonne Université, Paris, France

[6] Aix Marseille Univ, CNRS UMR 7289, INT (Institut de Neurosciences de la Timone), Marseille, France

[7] Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail), UMR_S1085, Université Rennes, Rennes, France

* These authors contributed equally to this work.

✪ Co-last and corresponding authors.

[£] Current address, Wellcome – MRC Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge, CB2 0AW, UK.

Abstract

Understanding human cardiac development is essential to improving the diagnosis and treatment of congenital heart defects. Here, we present a multi-modal atlas of the developing human fetal heart during the critical first trimester. Using single-nucleus RNA sequencing, we sampled nearly 50,000 cardiac nuclei from three human fetuses at 8.6, 9.0, and 10.7 post-conceptional weeks (pcw). This dataset enabled distinction of 21 cell types, including novel contractile, conductive, and stromal cells. Lymphatic endothelial, epicardial and autonomic neural and glial cells were among the new, smaller populations for which we established high-resolution transcriptional profiles. We further integrated the snRNAseq data with published single-cell RNAseq data from hearts between 5 to 7 pcw. Combined trajectory analysis allowed us to identify a new human cardiomyofibroblast progenitor, preceding the segregation and diversification of cardiomyocyte and many stromal lineages. To refine cell-type annotations, we turned to spatial transcriptomics. Analysis of six Visium sections from two additional hearts was aided by deconvolution of spots with our snRNAseq data. We further confirmed key markers using *in situ* hybridization or immunofluorescence on sectioned or whole hearts, followed by standard or light-sheet confocal microscopy. Altogether, these complementary approaches permitted us to add anatomical-positional features such as chamber specificities, innervation and conduction system components, or subdomains of the atrioventricular septation complex. Each translates cell identity into specialized cardiac functions. A total of forty first-trimester hearts were studied to discover and validate cellular characteristics across a wide range of scales, from the individual nucleus to the entire organ. Our results offer a comprehensive resource for understanding the cellular dynamics of human cardiac development and lay the groundwork for further studies into the molecular mechanisms underlying congenital heart malformations. This atlas adds unprecedented spatial and temporal resolution to the characterization of human-specific aspects of early human heart formation.

Introduction

The heart is the first permanent organ to assume a vital function in all vertebrate embryos as well as in many invertebrates (Maldonado et al., 2019). However, stage- and species-specific anatomy and function vary widely across the animal kingdom. Understanding the normal developmental trajectories of the cells and tissues of the human heart is the requisite benchmark for improving diagnoses, therapies and prognoses for congenital heart defects, the most frequent type of malformations present at birth.

Due to the challenges of studying human tissue directly, most of what we know about the molecular, cellular, and morphogenetic processes that drive heart development comes from research on animal models. While many severe congenital heart defects incompatible with continued pregnancy are typically established by the end of the first six weeks of gestation, numerous other serious or fatal defects arise in regions of the heart affected by human-specific parameters of shear stress, somatic growth, and anatomical changes that continue to modulate cardiac function up to and including birth (Oparil et al., 1984). These unique aspects are often not fully replicated in the vertebrate models commonly used for molecular and genetic research.

The critical window of development known as remodeling, during late human embryonic and early fetal life at the end of the first gestational trimester, has been inaccessible to comprehensive cellular investigation. As a result of sample scarcity, the atlas of human prenatal cardiac cell types, functions, anatomical positions and interactions continues to be enriched with ever more samples to better represent cellular and individual diversity (Asp et al., 2019; Cao et al., 2020; Farah et al., 2024; Hou et al., 2024; Lázár et al., 2024; Leshem et al., 2024; Xu et al., 2023). Each new paper, often adding both confirmatory and additional, complementary information, demonstrates the need for continuous or at least periodic reintegration of older with newer findings.

However, such integration is not a trivial task. Fetal heart cells produce hundreds to thousands of stage-specific transcripts and paralogous genes to those that define cell types in other parts of the body or at later life stages (D'Antonio et al., 2022). This has made it challenging or inappropriate to assign all cell identities at all stages on the basis of prior knowledge. Unique cellular functions also underlie small but crucial structures such as the valves, the interfaces of the autonomic nervous system with the cardiac conduction system, and specialization of the surfaces in direct contact with blood, requiring the transcriptomes of tens of thousands of cells to best characterize minority profiles. It is also well established that the fetal heart in other mammals presents distinct physiological properties that are mediated in part by organ-specific connective tissues from multiple developmental lineages, about which little is known in

humans (Ali et al., 2014; Deng et al., 2023; Dettman et al., 1998; Dewing et al., 2022; Gittenberger-de Groot et al., 1998; Norris et al., 2009; Vrancken Peeters et al., 1999).

These distinctions highlight the importance of refining the definitions of what cell types make up the fetal human heart and discovering mechanisms involved in its maturation during development. To address this problem, we thoroughly characterized multiple human hearts between 8 to 11 post-conceptional weeks (pcw), corresponding to the transition between the embryonic and fetal period. We generated *de novo* molecular and anatomically informed datasets about their composition using integrative spatial and single-nucleus transcriptomics, multi-dimensional transcript and protein localization *in situ*, including the available human cardiac cell atlases from earlier stages. These have allowed us to find evidence for previously unidentified intermediates in the lineage diversification of major and minor cardiac subpopulations, including novel markers and developmental features of a transitory subdomain of the atrioventricular septation complex. The greater definition of cell types and states in this atlas of the developing human heart promises more accurate understanding of all relevant influences on cellular differentiation, function and appropriate spatial organization.

Results

Increasing the resolution of biologically meaningful human cardiac cell types

To refine and enrich a multimodal cardiac atlas for the scientific community, we performed a detailed analysis of 40 human embryonic and fetal hearts staged between 6.5 and 12.3 post-conceptional weeks (pcw). This took the form of histological and complementary transcriptomics approaches, followed by spatial validation of transcript and protein expression in sliced and/or intact organs.

Assignment of cell types at single-nucleus resolution

Because cardiac cell types may vary in size across two orders of magnitude, we recovered the nuclei from whole hearts with proximal great vessels from one female (8.6 pcw) and two male (9.0 pcw and 10.7 pcw) fetuses (Supplemental Materials and Methods; Fig. 1A). 10X Chromium libraries were then prepared for single-nucleus RNA-sequencing (snRNAseq) and sequenced using an Illumina sequencer. After quality control and filtering with Cell Ranger (10X Genomics) and Seurat software (Hao et al., 2021), snRNAseq yielded transcriptomes for 18,129, 8,750, and 22,348 individual nuclei respectively, for a total of 49,227 fetal cardiac nuclei analyzed. Unsupervised clustering of the integrated snRNAseq datasets was then performed on the basis of the most highly variable genes, yielding 21 clusters (Table S1).

We assigned cell identities (Fig. 1B) using a combination of commonly used markers in animal models and a post-hoc literature search for co-expressed, differentially expressed genes (DEGs) when comparing a cluster against all others. Each cluster was present in all three hearts, although the arterial smooth muscle cell (SMC) cluster 12 had proportionately fewer cells at 10.7 pcw relative to earlier stages (Fig. 1C arrow, D and Supplementary Materials and Methods). This specimen (Fig. 1A) had a relatively truncated ascending aorta and pulmonary trunk, distal segments of the outflow tract, as compared to the others.

Some well-documented DEGs enabled cluster groupings into functionally identifiable classes (Fig. 1D-F). Markers for these included *MYH6* and *MYH7* for cardiomyocytes (I), *POSTN* and *PECAM1* for cardiovascular endothelial cells (II), and *EBF2 and PRRX1* for interstitial fibroblasts and smooth muscle in the stroma (III), respectively. Epicardial (IV), blood and immune cell (V) and neural crest-derived (VI) cells were also designated as classes (Fig. 1E).

Iterative refinement of annotations

Class I comprised 29,233 cardiomyocyte nuclei, which constituted 59.4% of the nuclei sampled across the three hearts (Table S1, tab 2). *MYH6* and *MYH7*, classical indicators of atrial and ventricular cardiomyocytes respectively (Lu et al., 2022), are markers for Class I. Atrial and ventricular chambers are histologically (Fig. S1) and functionally different from one another. The relative levels of expression allowed discrimination, as both markers were broadly transcribed in the entire class (Fig. 1E-G). The nine clusters assigned to Class I represented distinctly atrial (CmA1-A4) and ventricular (CmV1, CmV2, CmVCnD) subclasses based on cell type–specific DEGs, and intermediate cell types (CmG2M, CmMit). Significant DEGs from each class, as well as within- or between-class DEGs, are available in Table S1.

The “CmG2M” cluster between atrial and ventricular subclasses comprised 8.3% of the cardiomyocyte class. It expressed characteristics of the G2M transition, including among its most significant DEGs the centriolar assembly proteins *CENPF* and *CENPP*; *ANLN*, which encodes an actin-binding cytokinesis protein; *KIF18B* and *KIF15*, among nine kinesin superfamily genes; and *BRIP1*, encoding a BRCA1-interacting helicase. These are not transcripts specific to cardiomyocytes but more represented in CmG2M among other cardiomyocytes. The “CmMit” cluster referred to the 3% of cardiomyocytes potentially in active mitosis at the stages examined. Its significant DEGs included proliferation marker *MKI67*, centriolar proteins *CENPE* and *CENPF*; the cell cycle regulatory kinase, *AURKA*; and the regulator of mitotic initiation, *CCNB1,* among others. The most populous cluster CmV1 (9398 cells, or 32.2% of the class), like other ventricular cardiomyocytes expressed *MYH7* and *TTN*, but its cells also specifically co-transcribed higher levels than any other cardiomyocytes of the potassium channel *KCNQ5* (Kv7.5) and the PI3K signaling suppressor, *INPP4B* (Fig. 1G; Fig. S2A). This transcriptional combination may underscore the specific electrical properties and functional requirements of ventricular cardiomyocytes.

Another pair of interesting clusters spanned atrial and ventricular subclasses; these were respectively termed CmA2 and CmV2 (4.4 and 3.4% of cells in Class I, Fig. 1D and Table S1). CmA2 transcribed *PAM*, *TBX5* and *MYH6* as DEGs compared with all other cell types, placing this population in the atrium. CmA2 also expressed the *CACNA1D* gene more than other atrial cell types (Table S1, tab 7). *CACNA1D* encodes the L-type voltage-activated calcium channel CaV1.3, a marker of the sinoatrial node (van Eif et al., 2019). CmV2, like CmVCnD, expressed significantly more *KCNIP4* (voltage-gated potassium channel-interacting protein 4) but, like CmA2, the AV bundle marker *CNTN5* (contactin-5, a homophilic adhesion molecule; Fig. 1G) (Kanemaru et al., 2023; Smirnov et al., 2018). CmVCnD was distinguished from CmV2 and all other ventricular cardiomyocytes (Table S1, tab 6) by expressing more *NAV3* (neuron navigator 3), a conserved axon guidance cue necessary for zebrafish heart morphogenesis

(Lv et al., 2022). Other CmVCnD markers included *NLGN1*, encoding a transmembrane ligand for neurexins found in post-synaptic densities, including neuromuscular junctions (Ramesh et al., 2021) and *SORCS3*, a neurotrophin uptake receptor necessary for post-synaptic vesicular trafficking (Oetjen et al., 2014) (Fig. 1G). From the snRNAseq data alone, it was possible to ascertain that CmA2, CmV2 and CmVCnd represented distinct cell states and/or locations in the developing human pacemaker (sinoatrial and atrioventricular nodes) and conduction system.

Assignment of rare cell clusters was aided by first identifying their DEGs as compared to all other heart nuclei, then as compared to the other clusters within their class, and further using Enrichr-KG (Evangelista et al., 2023) to examine significantly enriched cell-type annotations from the Descartes human fetal atlas (Cao et al., 2020), the Human Gene Atlas (Su et al., 2004), the Tabula Sapiens (The Tabula Sapiens Consortium, 2022), the KEGG 2021 Human, and the Gene Ontology Biological Process databases (Fig. S2B). In this manner, the depth and breadth of our snRNAseq dataset enabled us to independently find a previously unidentified, but anticipated, population of 210 cells termed EnVL (venous-lymphatic endothelium, Fig. 1B, G). These were grouped in the EndoCV (endocardiovascular) Class II and distinguished by the co-expression of such well-known markers as *FLT4* (*VEGFR3*) and *TBX1* (Fig. 1G; Table S1) (Wang et al., 2023).

Other small clusters could be assigned firm identities as well. Although blood had intentionally been rinsed away during our collection procedure, 655 cells in Class V reflected fetal hemoglobin-expressing (*HBG1*/*HBG2*) erythrocytes, *RUNX1*-expressing immature and *PTPRC* (CD45)-transcribing differentiated hematopoietic progeny and other potentially resident immune cells, such as *SPP1*-positive macrophages (Fig. 1D-E). Likewise, the *NRXN1* (neurexin 1) and *CDH6* (cadherin-6)-transcribing neural crest-derived cell Class (VI) subdivided into distinct clusters. These represented neuroendocrine (NE; *NRG1* [neuregulin 1], *GATA3* and *ALK* (Boeva et al., 2017), 227 cells; Fig. 1E-G, Fig. S2B) and Schwann cell precursors (SCP; *EDNRB* [endothelin receptor B] and *CDH19* [cadherin-19, (George et al., 2018)], 309 cells; Fig. 1G, Fig. S2C).

Integration of single-cell first-trimester cardiac atlases identifies an early cardiomyofibroblast progenitor

To see how cell transcriptomes changed over time, we integrated our data with two independently published human heart atlases. These had been established from dissociated single-cell (sc)RNAseq data of seven whole human embryos at stages corresponding to 4-6 pcw (Carnegie stages 13 to 16) (Xu et al., 2023) or from a single heart at 6.5-7 pcw (Asp et

al., 2019), representing 8,266 and 3,717 additional cardiac cell transcriptomes, respectively (Fig. 2A). Similar integration of scRNAseq and snRNAseq datasets has enabled identification of minority progenitor cell types in mouse skeletal muscle (McKellar et al., 2021).

Cardiomyocytes and stromal cells in the mouse derive from common mesodermal progenitors that begin the process of lineage restriction as early as gastrulation (Lescroart et al., 2014). To search for a human equivalent process, we selected and integrated only those cells assigned to stromal or cardiomyocyte classes (I and III, Fig. 1D) between 5 pcw (Carnegie Stage [CS]13) and 10.7 pcw. We then performed dimensional reduction with Potential of Heat-diffusion for Affinity-based Trajectory Embedding (PHATE), a diffusion-based manifold learning method (Moon et al., 2019) (Fig. 2A-C; Fig. S3). The new clusters corresponded on a one-to-one or one-to-many basis with the cell type assignments we had made in our snRNAseq dataset alone (Fig. S3). However, not all clusters were equally populated by stage, and one in particular appeared more represented at CS15-16 (6 pcw) than at later stages (Fig. 2A-C, box). Slingshot trajectory analysis (Street et al., 2018) superimposed on this new projection indicated that human fetal cardiomyocytes and fibroblasts could indeed be traced back to a common cardiomyofibroblast (CmF) cell type mostly present before CS16 (cluster 16, Fig. 2D).

A confusion matrix between the clustering of the full scRNAseq/snRNAseq dataset and our original clusters (Fig. 1; Fig. 2E) showed that the CmF cluster had little in common with the later cardiomyocyte clusters of Class I or stromal clusters of Class III. Although a small minority at later stages, cells assigned to the CmF cluster were a persistent fraction of the total population analysed at 10.7 pcw (Fig. 2C, F). Greater contributions from later stages, particularly among the cardiomyocytes, were found at more distal trajectory nodes (Fig. 2D-F). Some assignments were almost exclusively observed in cardiac cells from the third month, including clusters 8 (CmA3) and 2 (CmA4) as well as a stromal progenitor in cluster 15 (Fb3) (Fig. 2F). In contrast, cells from the 6.5-7 pcw dataset (Asp et al., 2019) were relatively impoverished in cardiomyocytes and their precursors but enriched in assigned stromal types (Fig. 2F), as might be expected from isolating cells over a wide size range rather than more homogeneous nuclei.

The CmMit (cluster 6) and Fb3 (cluster 15) types of mitotic cardiomyocytes and fibroblast progenitors were the nearest class-specific progenitors to the cardiomyofibroblast (Fig. 2D). In order to see which of these was closer to the prospective parental CmF cell type, we classified the transcription of all significant DEGs into “high” or “low” bins (Table S2), and scored how these categories were inherited over the trajectory from cluster 16 to 6 or to 15 during differentiation. Most transcribed genes were maintained from CmF (cluster 16) to

CmMit (cluster 6) but newly transcribed in Fb3 (cluster 15) (Fig. 2G), implying a closer relationship of CmF to proliferating cardiomyocyte progenitors.

By integrating former and novel data acquired over the second and third months of gestation, it was possible in this manner to identify new cardiac progenitor states and model potential cell lineage relationships in the developing human heart.

Spatial transcriptomics of multiple 8.4 and 9.7 pcw heart sections

As transcriptomics of single nuclei is high-resolution but sparse, and removes informative positional information from samples, we also undertook 10x Genomics Visium spatial transcriptomics analysis on two sections from a male heart collected at 8.4 pcw and four sections from a female one at 9.7 pcw (Fig. 1A, Figs. S4-9). Over 12,000 regularly spaced tissue spots from the six sections were analyzed (detailed in Supplementary Materials and Methods).

Sections were first analyzed individually with Seurat, yielding similar numbers of unsupervised spot clusters (eight for sections at 8.4 pcw, six to eight for the 9.7 pcw sections (Figs. S4-9). Cluster information was then projected onto each section at the appropriate coordinates. With the aid of other hearts (Fig. S1), histological landmarks and known markers permitted us to assign cluster identities. Clusters corresponding to compact or trabecular *MYH7*+ cardiomyocytes of the ventricle expressed *HEY2 and IRX3*, respectively (Table S3). *MYH6*+ cardiomyocytes expressed *BMP10* and *TBX5* in the right atrium and *PITX2* in the left atrium. A cluster corresponding to arterial smooth muscle and enriched in *ELN* and *MYH11* was identified at both stages, as well as discrete clusters containing valve (*SCX*, *PENK*) or subvalvular (*POSTN, PRRX1*) stroma. Full lists of marker genes for individual section spot clusters are provided in Table S3.

We further analyzed these data by performing integration of the six spatial transcriptomic datasets using Seurat. Fourteen distinct clusters were identified by this means, allowing comparison of more detailed cell type distributions across time and space than when each section was analyzed individually (Fig. 3A, B, and Fig. S10). Left and right, compact and trabecular cardiomyocyte clusters expressed expected markers (Fig. 3B,C). Full lists of marker genes for integrated spatial clusters are also provided in Table S4.

This approach also enabled discrimination of relevant non-contractile cell types. Cluster 6 was enriched in transcripts for mesenchymal markers like *PRRX1*, *DCN, FN1* and multiple collagens (*COL1A1, COL1A2, COL3A1*), as well as markers enabling unambiguous assignment to encased autonomic neurons (*NEFL*, *TUBB3*, *PRPH*, *PHOX2B*). Thick-walled, embedded blood vessels clustered separately (as cluster 11). The ascending aorta and

ventricles were surrounded by a transcriptionally distinct *tunica adventitia* (cluster 7, *COL1A1*+; Table S3). The valves themselves (cluster 10) expressed *SCX* (Fig. 3C) and *POSTN*, but clusters 5 and 12 were also strongly enriched in *POSTN*. The ascending aortic elastic smooth muscle cluster 12 (*ELN*+) was clearly visible and distinct at both stages. Many individual spots dominated by an epicardial identity (cluster 9, with co-expression of the transcription factor *TBX18* and *ITLN1*+, an epicardial adipokine (Sacks and Fain, 2007)) could be distinguished. The cluster with the fewest spots was labeled “cardiac conduction system” on the strength of its enriched expression of *SHOX2* and *TBX3* (Fig. 3 A-C). It corresponds more specifically to cells of the conduction nodes, demonstrating the improved precision of cell type annotation made possible by adding positional information (Table S4; Fig. 3B, C).

To validate some of these findings, we carried out immunofluorescent labeling of human embryonic heart sections against MYL7, MYH6 and MYH7 proteins (Fig. 4). Myosin is a large hexamer composed of four light and two heavy chains, necessary for cardiomyocyte and muscle contractility. Myosin regulatory light chain 7 (encoded by *MYL7*) is a human atrial-specific gene, expressed in the fetal heart from 8 pcw until 15 days postnatal, and to a lesser extent in the ventricles throughout life (Hailstones et al., 1992). While *MYL7* was expressed throughout the four heart chambers at 8.4 pcw (Fig. 4A-B), there were more transcripts in the atria than the ventricles by 9.7 pcw (Fig. 4C), which was also borne out in terms of protein immunoreactivity at 9.0 pcw (Fig. 4D). This confirms *in situ* hybridization results from the same stage (Hailstones et al., 1992). MYH6 and MYH7 proteins were found at 9.0 and 12.0 pcw respectively, but not exclusively, in the atria and the ventricles, as had been noted in the snRNAseq transcriptome dataset (Fig. 4E-L). We also validated that *BMP10* was enriched in trabeculae of human ventricular and right atrial chambers relative to myocardium (Teichmann and Kessel, 2004). Suggested by the snRNAseq data (Fig. 4M), *BMP10* expression was confirmed by RNAscope (FISH) analysis at 9.0 pcw (Fig. 4N-P).

Thus, integration of spatial transcriptomic data increased the resolution of human early fetal heart atlas, enabling finer identification of cell types in sections.

Deconvolution of spot transcriptomes

Each spot in a sampling approach to sequencing-based spatial transcriptomics data such as the one we chose may contain distinct cell types in varying proportions. To measure this heterogeneity, we deconvoluted each spot using the Robust Cell Type Decomposition (RCTD) method (Cable et al., 2022), with our integrated snRNAseq data as a reference. The estimations of relative proportions of cell types for each spot of spatial transcriptomic data was

then projected onto the spots of the six heart sections in our dataset, turning each into a pie chart of relative proportions of each scRNAseq cluster (Fig. 5, Fig. S11-16).

Figure 5 shows examples of 8.4 pcw heart section A91 and 9.7 pcw heart section A17. For each spot, the relative proportion of each snRNAseq cell type is depicted as a pie chart (Fig. 5 A,C). Spot deconvolution results were further averaged in each of the 14 clusters across the 6 integrated spatial transcriptomic data sets (Fig. 5 A,C; Table S5). Large pie charts in Fig. 5A and 5C present the average relative proportions of snRNAseq cell types in clusters 0 (compact myocardium RV), 2 (trabeculae), 3 (right atrial cardiomyocytes), 4 (left atrial cardiomyocytes), 7 (tunica adventitia), 10 (valves), 12 (aorta, vascular smooth muscle) and 13 (conduction system) at each stage. Table S5 provides detailed deconvolution information (relative proportions as percentages) for each spatial transcriptomic cluster in each section.

Our deconvolution analysis quantified the degree of cell-type heterogeneity per spot. For example, at 8.4 pcw (Fig. 5A), the largest cluster 0 (right ventricular compact myocardium) comprised on average 40% of cellular transcriptomes most like CmVCnD, 15% of CmV1, 19% of CmG2M and 12% of EnTh cell types identified in the heart snRNAseq integrated data. Cluster 10, corresponding to valvular cells in the spatial transcriptomics dataset, is estimated to contain 73% transcriptomes resembling FbVIC cells, but also 6% of EndoCd cells and 7% of EnTh cells.

Fig. 5B and 5D show estimated relative proportions of the following clusters from integrated snRNAseq data in each spot of heart section A91 at 8.4 pcw and section A17 at 9.7 pcw after deconvolution: CmVCnD, CmV1, CmG2M, CmA1, CmA2, CmA3, CmA4, CmMit, EnTh, EndoCd, Fb1, Fb2, FbVIC, SMC and EpiCd. Similar proportions of cell types per spot per cluster were obtained across all sections at 8.4 pcw and 9.7pcw (Figs. S11-16; Table S5).

The four heart chambers develop discrete identities by the third month of gestation

These spatial and deconvolution analyses in turn further refined the identities of clusters assigned after integrated snRNAseq (Fig. 1). For example, at 8.4 pcw (section A91), spots of the right atrium (cluster 3 in integrated spatial transcriptomic data) were composed on average by 30% cellular transcriptomes most like CmA4 and 12% cellular transcriptomes most like CmA1. In contrast, those of the left atrium (cluster 4 in integrated spatial transcriptomic data) were composed of an average of 40% cellular transcriptomes most like CmA3 but also 5% most like CmV1, 6% most like CmG2M and 3.5% most like CmA1 (Table S5). CmA4 is thus

associated with right atrial identity while CmA3 is associated with left identity, while CmA1, their purported progenitor (Fig. 2D), is found in both atria (Fig. 5B, D).

The CmA2 cluster mapped specifically to sinoatrial and atrioventricular nodes, which for the former was observed across multiple sections at both stages (Fig. 5B, arrows; Figs. S11-16). Likewise, the CmVCnD cluster was more represented in the right ventricle, while the CmV1 cluster was more present in the left ventricle; these cluster identities were observed across both compact and trabecular myocardium (Fig. 5B, D). Valvular interstitial fibroblast (FbVIC) cellular transcriptomes were present in both atrioventricular and aortic valves (Fig. 5B, D; Figs. S11-16).

For other spatially restricted populations, CmV2 cardiomyocytes were concentrated in a subvalvular location, particularly around the ventricular septum, while the EndoCd cluster mapped clearly to trabecular endocardium (Fig. 5B, D). The SMC cluster corresponded to vascular, rather than intracardiac, smooth muscle cells of the aorta and extracardiac arteries, while the Fb1 cluster mapped to the *tunica adventitia*, capping the ventricles under the atria (Fig. 5B, D). Figs. S11-16 show the estimated proportion of each cluster assigned from our integrated snRNAseq data mapped to each spot of the Visium data, section by section.

In this manner, by combining snRNAseq and spatial transcriptomic data and deconvoluting the latter, we not only investigated cell-type heterogeneity of spatial transcriptomic spots wholescale, but we were also able to better define the identity of snRNAseq clusters. For example, we confirmed the distinct identities of endocardium and vascular endothelium by adding spatial distribution information. We appended complementary spatial terms (*eg*. left and right atria, subvalvular and atrioventricular node) to the snRNAseq cluster names in Fig. 5.

A multi-omic approach identifies small cell clusters and rare transcripts

Integrating the two modalities allowed us to identify genes expressed in localized subpopulations of cardiac cells, including some not previously known to be involved in human cardiac development. In light of their calculated spatial expression, we revisited the list of markers of the integrated snRNAseq data clusters (Table S1) and cross-referenced them with the list of cluster markers from the integrated spatial transcriptomic data (Table S4). Fibroblast and mural cell types in the mouse are functionally diverse and site-specific (Muhl et al., 2020). To identify similar specializations, we compared the combined list of genes significantly enriched in the clusters of snRNAseq stromal Class III with those significantly enriched in six

spatial clusters corresponding to valves, subvalvular and connective tissue, tunica adventitia, extracardiac vessels and aortic SMC (Table S4). One intriguing gene was *PENK*, which encodes proenkephalin, a component of the neuroendocrine opioid system expressed in murine neural crest-derived mesenchyme and valve primordia (Chen et al., 2022; Liu et al., 2019; Soldatov et al., 2019). *PENK* was a significant DEG in FbVIC by snRNAseq (Table S1). It was strikingly expressed both in and between the valves in spatial transcriptomic sections (Fig. 6A, Fig. S17). Other DEGs enriched in the valve spatial cluster included *BOC*, *INHBA*, *LTBP2*, *MEOX1*, *PTN*, *PTPRZ1*, *PRRX1*, *ROBO1,* and *RSPO3* (Fig. 6A, Fig. S17).

RNAscope-based *in situ* hybridization confirmed that *PENK* was indeed strongly expressed in and between atrioventricular valves in additional specimens (Fig. 6D, F, F', M, N). Although *ACTA1* (Actin Alpha 1, Skeletal Muscle), was the top differentially expressed gene in the subvalvular spatial transcriptomic cluster 5 (Table S4), it was slightly but significantly enriched in the CmMit snRNAseq cluster relative to all others (Table S1). RNAscope confirmed restricted spatial distribution of *ACTA1* transcripts at the ventricular attachment points of the valvular leaflets, corresponding anatomically to the chordae tendoneae (the "heart strings"; Fig. 6E, F, F'). Furthermore, our spatial transcriptomic data suggested that *ACTA1* and *PENK* were expressed in adjacent but mutually exclusive domains (Fig. 6G), a finding we validated with distinct fluorochromes to co-hybridize the probes in the same tissue (Fig. 6G').

A comparison of genes expressed in the subvalvular spatial cluster 5 with other stromal class spatial clusters (6, 7, 10, 11, 12) showed that only two genes, both associated with fetal cardiomyocyte mitosis, *ANLN* (anillin, actin binding protein) and *CDK1* (cyclin-dependent kinase 1) (Alharbi et al., 2024), were exclusively enriched (Table S4). This showed that the subvalvular state was largely determined by position, average state of the cell cycle and degrees of gene expression, rather than unique markers. While CmV2 made up on average 17% of the subvalvular cluster at 8.4 pcw and 14% at 9.7 pcw, CmV1 increased from 7 to nearly 12%, CmVCnd composed another 15% of the population at both ages, and CmG2M were the most populous cells in the subvalvular domain, at over 19% at both ages (Table S5). The septal area bridging and ventricular to the valves thus appears to harbor cycling stromal and myocardial cells at varying degrees of maturation.

The mesenchymal homeobox-1, *MEOX1*, which plays a key role in regulating fibrosis in injured or stressed postnatal hearts (Alexanian et al., 2021), was transcribed in a restricted manner in the atrioventricular valve (Fig. 6D); from the snRNAseq data it was significantly enriched in both FbVIC and Fb1 fibroblasts (Table S1). We also confirmed the concomitant expression of *PRRX1*, a mesenchymal marker widespread in the spatial and snRNAseq data but also among the top DEG characteristic of CmVCnd. It was widely expressed in aortic and tricuspid atrioventricular valves as well as in the subvalvular domain and ventricular trabeculae (Fig.

6E). Altogether, these observations suggest that many DEGs characteristic of spatial clusters modulate the mechanical and signaling environment of the valve and its interface with contractile cells during morphogenesis.

Finally, we observed a subpopulation of spots within cluster 6 (connective tissue) of the integrated spatial transcriptomic data, where many genes associated with peripheral nervous system development appeared to be co-expressed. These included *ISL1*, *TH*, *PRPH*, *PHOX2A*, *PHOX2B*, *PRRX1*, *DLK1*, *PENK*, *TBX2, GAP43*, *L1CAM*, *STMN2* and *TUBB3* (Fig. S18). RNAscope and immunofluorescence analysis confirmed expression of PRPH protein and *ISL1* and *PRRX1* transcripts intermingled in complementary domains (Fig. 7A-C). This population thus most likely corresponds to a cardiac parasympathetic ganglion.

Toward a four-dimensional atlas

Cell types along reticulated networks of small blood vessels, connective tissues and nerves are represented in all organs but can be difficult to detect because of low absolute cell numbers or difficult to distinguish their spatiotemporally specific properties. The former difficulty was encountered for the 536 (1.1% of total) cells assigned to the two clusters of the neural crest-derived Class VI. The latter was the case for the 7,224 transcriptomes of endocardial, vascular or lymphatic endothelial cells (Table S1). In particular, although we had identified clusters corresponding to neuronal (*PRPH*+) and endothelial (*PECAM1*+) cells (Fig. 1 and Fig. 8A, D), their anatomic localizations were incompletely represented in the Visium spatial transcriptomics data (Fig. 8B, E). Similarly, the architecture of blood vessels and peripheral nerves within and surrounding the heart is also incompletely reflected in sections (Fig. 8C, F).

Three-dimensional immunofluorescence and tissue clearing have successfully allowed observation of the ramifications of peripheral nerves and lacrimal glands to better study human-specific morphogenesis over time (Belle et al., 2017; Blain et al., 2023). We also used whole-organ immunofluorescence to examine the spatial arrangements of blood vessels and peripheral nerve in ten human fetal hearts (Fig. 8G-L and Supplemental Movies 1-4), as well as markers of other cell types (*eg*. labeled by FABP4, MYH6, POSTN; Fig. S19). Light-sheet confocal microscopy resolved expression of these proteins at a cell-to-tissue scale, to bridge annotations of cell types and their spatial positions in the organ.

We could thus demonstrate that the great arteries of the outflow tract at 9-10 pcw are far more densely innervated than the cardiac apex, and that innervation progresses in a proximal-to-distal manner over time. Nerves tracked with blood vessels on the outside of the heart before penetrating the muscle more robustly in the ventricles than the atria (Fig. 8G, J). The endocardial lining of all valvular cusps, like the arterial trunk and coronary veins, expressed

PECAM1 in a sparse manner as compared to the strong immunofluorescence in ramifications of the coronary arteries (Fig. S19), while the valves themselves were devoid of nerves at all stages examined (Fig. 8I, L). DEGs for the FbVIC, NE and EndoCd clusters (Table S1), respectively, encode candidate repulsion mediator molecules such as *ROBO1*, *SEMA6A*, *SEMA3D* and *PLXNA4* in the valvular interstitial fibroblasts; *SLIT3*, *SEMA3C*, *SEMA6D*, *UNC5B, UNC5*C, *UNC5D*, *DCC* and *DSCAM* in cardiac neurons; and *NTN1*, *PLXNA4* and *PLXND1* in the endocardium.

Discussion

In this work, we independently analyzed whole fetal human hearts using single-cell and spatial transcriptomics, and enriched cell type annotations with spatial and anatomical information from additional samples, to obtain a comprehensive overview into a vulnerable window of human cardiac development. Recent work, by others (Farah et al., 2024; Knight-Schrijver et al., 2022) and here, highlights the importance of iterative refinements and fully accessible data to refine the “ground truth” for developing organ atlases, not limited to the heart. While new contributions, whether replicative or novel, could seem incremental, they play a critical role in addressing gaps in current knowledge by increasing statistical power. Building a comprehensive, community-driven developmental human cell atlas requires an open, iterative process where scientists can continuously reinterpret and enhance their findings in light of new findings and technical approaches (Haniffa et al., 2021).

Identification of previously undescribed cell types and states

We distinguished nine populations of cardiomyocytes during analysis of 49,227 transcriptomes from individual human fetal cardiac nuclei at 8.6, 9.0, and 10.7 pcw. One of these was preparing for, and another seemingly in, active proliferation (CmG2M and CmMit). In the minority even at these stages (14% and 3% of total cardiomyocytes, respectively), such cell types are at least ten times scarcer in postnatal human hearts (Bergmann et al., 2009). In the mouse fetus, most cardiomyocytes undergo a final round of nuclear mitosis and cell division around the time of birth that yields the post-mitotic, multinucleated or polyploid cells capable of constructing the sarcomeric scaffold needed for a lifetime of contraction (Soonpaa et al., 1996). Likewise, most cells in the postnatal human heart are post-mitotic. This atlas provides a robust reference of normal human development to which improvements to the maturation of induced pluripotent stem cells for cell therapy in such conditions as hypoplastic left heart syndrome (Krane et al., 2021) or infarction (Jiang et al., 2024) can be compared.

Furthermore, after integration with earlier available datasets, we discovered evidence for a trajectory diversifying multiple novel fibroblast types. The earliest of these seems to correspond to a common progenitor to both fibroblast and cardiomyocyte lineages. Genetic lineage tracing in stem cell-derived organoid or gastruloid models may resolve the potential and contributions of such cardiomyofibroblast progenitors (Argiro et al., 2023). Integration with further datasets (Cranley et al., 2024; Lázár et al., 2024) could resolve the nature, contributions and competence of these stromal cell types in and beyond the human fetal heart over time.

Stromal diversity reflects anatomical specializations

Fibroblasts are extracellular matrix and growth factor-secreting cells of remarkable functional, organizational and lineage diversity. Epicardial cells are the predominant contributors to cardiac fibroblasts, giving rise to the majority of resident fibroblasts in mice (Ali et al., 2014). However, endocardial and endothelial cells also contribute to fibroblast development, particularly in the ventricular septum and chamber walls (Deng et al., 2023; Liu et al., 2018). Furthermore, neural crest cells have been shown, through lineage tracing and transcriptomic or proteomic profiling, to differentiate into migrating pericytes, fibroblasts, smooth muscle cells and even melanocytes in the proximal great vessels, valves, right atrium and septal bridge (Etchevers et al., 2001; Hulin et al., 2019; Jiang et al., 2000; Le Lièvre and Le Douarin, 1975). We found no evidence of melanocytes in these first-trimester human hearts, but the differential expression of *CHD6* (cadherin-6) as a marker of neural crest-derived Class VI cells is compatible with their mesenchymal phenotype at the stages assessed (Clay and Halloran, 2014).

According to the specific valve and even the leaflet, the fibroblasts known as valvular interstitial cells (VICs) have been shown in vertebrate models to have distinct developmental origins and convergent functions. In atrioventricular or inlet valves, VICs arise mostly from endocardial and epicardial cells, whereas in the arterial or outflow valves, they are derived from neural crest cells, alongside contributions from endocardial and second heart field cells (El Robrini et al., 2016; Etchevers et al., 2001; Jiang et al., 2000; Le Lièvre and Le Douarin, 1975; Liu et al., 2018; Liu et al., 2023; Waldo et al., 1998). Our findings contribute further evidence that many human cardiac fibroblasts differentiate from non-epicardial progenitors, as demonstrated in mice (Deng et al., 2023). Trajectory analysis over integrated 5-10 pcw data shows that the epicardial lineage had, as expected, segregated at an earlier time in human mesoderm from the common cardiomyofibroblast lineage branch. By 8.6 pcw, CmF had diverged into dividing progenitors inclined to contribute to either stromal or cardiomyocyte lineages but not both. CmF-derived stroma is predicted to contribute via the Fb3 type of fibroblast to a range of

mesenchymal cell types, with earlier segregation of VICs than other fibroblasts or smooth muscle.

Highly differential *TBX5* expression relative to other stromal subtypes was a marker of Fb3 cells, which were simultaneously enriched in *FHL2*, a titin ligand implicated in cardiac mechanotransduction (Strom et al., 2024) (Table S1). *Fhl2* has also been shown to be a transcriptional target of Tbx5 in mice, where downregulation of either is associated with fatal arrhythmias (Rathjens et al., 2021). *TBX5* is expressed in all human fetal atrial cardiomyocyte subtypes (CmA1, CmA2, CmA3, CmA4; Fig. 3C). This is to be expected from its evolutionarily conserved role in promoting atrial-specific cardiomyocyte transcription (Sweat et al., 2023) and its missense or truncating mutations in Holt-Oram syndrome, where arrhythmias and atrial septal defects are common (Møller Nielsen et al., 2024). Separate enhancers or genomic contexts for TBX5 may distinguish atrial cardiomyocytes from *TBX5*-expressing Fb3 fibroblasts. TBX5 is one of the three early developmental transcription factors able to reprogram human fibroblasts directly into induced cardiomyocyte-like cells (Miyamoto et al., 2018), suggesting that the persistent Fb3 population may retain some CmF potential.

Complementary information from spatial transcriptomics establishes a more detailed atlas

Spot clusters, dominated by cells with similar transcriptional profiles, established new links between cardiac cell localization, transcriptomic profile and cellular function when mapped back to anatomically and functionally distinct heart regions. For example, we identified and validated specific expression of *BMP10* in ventricular trabeculae, as shown in chicken and mouse embryos (Chen et al., 2004; Teichmann and Kessel, 2004), but also in the right atrium, consistent with its asymmetric enrichment of pectinate muscles in mice and humans (Ho et al., 2002; Sedmera et al., 2000). The role of *BMP10* in their elaboration is currently unknown, but thickened junctions of pectinate muscles with the lateral right atrial wall, studied in diseased adult hearts *ex vivo*, have been proposed to drive their persistent atrial fibrillation (Zhao et al., 2023). Our multi-modal atlas shows for the first time that *BMP10* co-varies spatially in the right atrium and in specific cardiomyocytes (CmA1/CmA4 populations) with the *PRKAG2* gene. PRKAG2 loss-of-function is responsible for congenital, progressive cardiomyopathies with high rates of life-threatening atrial fibrillation and/or ventricular hypertrophy, glycogenosis or tachycardia (Blair et al., 2001; Gollob et al., 2001; White-Brown et al., 2024). A human microdeletion syndrome affecting BMP dosage phenocopies *PRKAG2* hypomorph alleles (Lalani et al., 2009). Our observation provides further evidence that BMP receptor-mediated signal transduction is needed to establish the *annulus fibrosis* at the

atrioventricular junction (Gaussin et al., 2005) and thereby, may prevent inappropriate electrical coupling outside of the pacemaker circuit.

We also found specific expression of *PITX2* in the left atrium in both types of transcriptome data. Early fate lateralization reinforced by hemodynamics are known to enhance the molecular differences between cells of the left and right heart and in particular, *Pitx2* expression in distinct chambers in embryos of many vertebrate models (Liu et al., 2002), some of which persist in human adult tissues (Hill et al., 2019).

We countered a potential drawback in spatial detection of rarer cell types by integrating all spatial transcriptomic data. Increased resolution allowed identification of spots assigned to the sinoatrial node or an autonomic ganglion. Above all, the deconvolution of spatial spot transcriptomes with the identities established by our single-nuclear analyses allowed us to achieve two key outcomes. First, we could predict the heterogeneous cellular composition of each spot, and second, we improved the identification of snRNAseq clusters using spatial information. This was true for neuroendocrine, endothelial *versus* endocardial, and trabecular cardiomyocyte identities.

Our integrative approach allowed us to better characterize and identify new marker genes of subpopulations involved in atrioventricular septation of the heart. Septation is affected in Holt-Oram (*TBX5*) syndrome as above, but also in DiGeorge (22q11.2 deletion, implicating *TBX1*) syndrome. Mouse embryos with reduced retinoic acid signaling show a direct effect on the expression of Tbx5 in a subpopulation of Tbx1+ cardiac progenitors in the second heart field, necessary to populate and septate the venous pole (De Bono et al., 2018). Further investigation of the cell populations involved in atrioventricular septation is a prerequisite to understand the etiology of congenital defects affecting this region of the human heart.

There has been historical ambiguity about whether the term “endocardial” is only a positional or cell-autonomous distinction from endothelial cells of arteries, veins, lymphatics and capillaries (Wang et al., 1998); we have demonstrated here that human fetal endocardium has both spatial and molecular specificities. Furthermore, the multi-modal transcriptomic approach enabled identification of spatially restricted expression when a gene was transcribed in multiple lineages, such as *PENK*. The tissue of the *tunica adventitia*, apposed to the proximal ventricles and outflow tract, was best distinguished in the sections subjected to spatial transcriptomics, and was enriched in macrophage-specific markers. We could more efficiently predict left *versus* right chambers, compact *versus* trabecular myocardium and conduction nodes *versus* Purkinje cells in subpopulations of the snRNAseq datasets, and find both specific and sensitive markers to validate with these complementary techniques.

In this study, we combined both snRNAseq and spatial transcriptomics approaches to better define and characterize cell types for the human fetal cardiac atlas. The former had the advantage of providing quantitative co-transcriptional information at the resolution of individual nucleus while the latter provided physical landmarks such as chambers and valves, or coordinates relative to the body plan (dorsal/ventral, anterior/posterior) and the organ itself (proximal/distal, inside/outside). Organ atlases can thus be refined even with a few well-chosen samples, adding informative dimensions by combining them with existing data and new sets available for reanalysis.

Sparse but widespread biological coordinate systems

The aforementioned methods can still bypass important components made of rare and/or sparse cells present within a sample, including nerves, lymphatics, resident immune cell types, vasculature and fascia. Synergy between next-generation sequencing or image-driven transcriptomics, multidimensional protein and/or mRNA labelling approaches, and qualitative as well as quantitative measurements adds functional significance to cellular identity labels. The 3D immunofluorescence image datasets we acquired for a few time points and markers are a proof-of-concept that multiple networked biological structures could serve as anatomical coordinates in the developing fetus as well as in the adult (Boppana et al., 2023). They may also be re-used for quantitative segmentation measurements over time and to train deep-learning models (Kaltenecker et al., 2024).

Until this year, catalogues of types of cells identified in the human heart had characterized relatively few individual cells leading up to the critical time window of the end of the third month of gestation and had thereby neglected important cell types and states. As performed for reasons of scale, sampling representative portions of larger organs (Litviňuková et al., 2020) or regressing out highly variable cell-cycle or ribosomal transcripts as once performed in single-cell analysis pipelines (Butler et al., 2018), can exclude expected and functionally important minority populations from developing tissues. Autonomic neurons and glia, specialized attachment zones between tissues such as the chordae tendineae, endothelia and valvular interstitial fibroblasts and likely cardiomyocyte and mesenchymal progenitor states are among these. Our atlas iteration of the developing human heart has placed each type in its spatiotemporal and functional context, to serve as a reference and roadmap for future work.

Methods

Human embryo collection, staging and quality control

First trimester human embryos and fetuses (6-13 pcw) were obtained from induced terminations of pregnancy performed legally in France and sought for reasons other than known fetal abnormality. Tissues were collected after obtaining written consent from donors following procedural initiation, a protocol approved by the national Agency for Biomedical Research (authorization #PFS14-011 to SZ; Agence de la Biomédecine) and declared by the Human Developmental Cell Atlas (HuDeCA; https://hudeca.com) consortium to the French Ministry of Higher Education and Research under reference DC-2022-5011.

Sample collection

Terminations of pregnancy were induced using a standard combined mifepristone and misoprostol protocol, followed by aspiration. Gestational age was initially determined by ultrasound, refined by measurement of foot length (Evtouchenko et al., 1996; Hern, 1984; O'Rahilly and Müller, 1987; O'Rahilly and Müller, 2010), and confirmed through comparison and contribution of high-resolution photographs of our samples to a staged morphometric and histological cardiac atlas hosted at the Human Developmental Biology Resource (Gerrelli et al., 2015) (https://hdbratlas.org/organ-systems/cardiovascular-system/heart/dissections/gallery.html or https://bit.ly/HumanFetalHeart).

Intact hearts were recovered from aspiration detritus under a Zeiss binocular microscope, rinsed in ice-cold phosphate-buffered saline (PBS) and maintained on ice in Hank's buffered saline solution until further processing. Retained tissues were documented in a secure HuDeCA-wide database https://hudeca.genouest.org) using the OpenSpecimen biobanking LIMS (https://github.com/krishagni/openspecimen/; Krishagni Solutions, India).

Genotyping

Quantitative PCR was performed in duplicate on genomic DNA derived from lung or umbilical cord fragments using primers within *SRY* (ACAGTAAAGGCAACGTCCAG and ATCTGCGGGAAGCAAACTGC) to amplify a Y chromosome-specific fragment of 293 bp (Lardenois et al., 2023), and normalized to a 183 bp fragment of the *GAPDH* promoter on chromosome 12, amplified by CCACAGTCCAGTCCTGGGAACC and GAGCTACGTGCGCCCGTAAAA.

Genomic DNA from all samples were further annotated by molecularly determined sex and exclusion of large chromosomal-level copy number variations using a custom digital-droplet PCR analysis (StemGenomics iCS-Digital karyotypes (Assou et al., 2020)). Samples with ambiguous results were reanalyzed using subtelomeric Multiplex Ligation-dependent Probe Amplification (Aneuploidy MLPA probemix [P095], MRC Holland) or for sample HEF_DN_F0000467, with a diagnostic CGH array to exclude a smaller potential copy-number anomaly on chr22 with a high degree of confidence.

Single-nucleus RNA-sequencing

Three intact hearts from one female and two male fetuses (validated in the transcriptional data) were dissociated for single-nucleus incorporation into the GEMs of the 10X Chromium platform. Dissociation was begun by an initial crush in liquid nitrogen, continued with 10-20 strokes in a Dounce homogenizer over ice, and nuclei further separated from debris over a discontinuous density gradient at 30-35% iodixanol as per a community protocol (Martelotto, 2021). Libraries were prepared by the Genomics and Bioinformatics facility (GBiM) of the U1251 / Marseille Medical Genetics lab using the Chromium Next GEM Single Cell 3' Kit v3.1 (PN-1000269) with Dual Index Kit TT Set A (PN-1000215, 10X Genomics), according to manufacturer's instructions.

Sequencing and raw data processing

All libraries were sequenced on a Nextseq 500 (Illumina) by the GBiM facility, except for samples HEF_DN_F0000467 and HEF_DN_F0000530 on a Nextseq 500 by Integragen (Evry, France). The sequenced snRNAseq libraries were processed and aligned to the human reference genome (GRCh38-2020-A) using Cell Ranger software v.3.1.0 (10x Genomics), with unique molecular identifiers (UMI) counted for each barcode.

Analysis, visualization and integration of snRNAseq datasets

Three datasets of snRNAseq from euploid hearts sequenced on different runs (HEF_DN_F0000530, 8.6 pcw, female; HEF_DN_F0000374, 9.0 pcw, male; HEF_DN_F0000467, 10.7 pcw, male) were analyzed using default parameters in the Seurat 4.1.1 pipeline and 40 PCA dimensions (Hao et al., 2021; Stuart et al., 2019). The pipeline included QC and data filtration, identification of highly variable genes, dimensional reduction, graph-based clustering, and differential expression analysis for the identification of cluster markers. Each dataset was preprocessed and analyzed separately, and later integrated with

anchor-based batch correction. Further sample details and photographs are in Supplemental Materials and Methods.

Integration of heart scRNA-seq and snRNA-seq datasets

We integrated our dataset (snRNA-seq) with scRNA-seq datasets of human heart from early stages using the top 1000 HVGs and reciprocal PCA (RPCA) in Seurat; specifically, the datasets included CS12-16 (Xu et al., 2023) and 6.5-7 post-conceptional weeks (Asp et al., 2019). The similarity between cell types across datasets was defined as 1/(1 + distance), where the distance between cell types was determined using a Mehalanobis-like distance metric in Slingshot (Street et al., 2018) based on the top 30 principal components (PCs) of the integrated dataset. Cell types were matched between adjacent developmental stages. The best-matching cell type was initially linked if similarity > 0.06. A secondary match was linked if the similarity z-score across all cell types was > 1 and similarity > 0.06.

Integration of cardiomyocyte and stromal cells from three datasets

The UMI count matrices of cardiomyocyte and stromal cells from each dataset were combined in Seurat. Integration was performed using the top 1000 HVGs and RPCA, followed by reclustering with top 30 PCs and resolution 0.4. Sixteen clusters were obtained in reclustering for cardiomyocyte and stromal cells across three datasets. PHATE analysis (Moon et al., 2019) was performed on the cardiomyocyte and stromal cells using PCA input and 20 KNNs (k-nearest neighbors). Clusters identified during reclustering were mapped onto the PHATE embedding. The trajectories of those clusters were computed using Slingshot [PmID 29914354] with PCA input.

To classify gene expression along the bifurcation of CmF, we defined four distinct patterns: (1) CmMit > Fb3 and CmF > Fb3; (2) CmMit > CmF and CmMit > Fb3; (3) CmMit < CmF and CmMit < Fb3; (4) CmMit < Fb3 and CmF < Fb3. Here, the ‘<’ and ‘>’ were defined by at least 2-fold change and mean expression > 0.5.

Spatial transcriptomics

Fresh fetal hearts were staged, photographed under a Leica dissecting microscope and flash-frozen in Leica Frozen Section Compound (FSC)22 in polypropylene molds floated over isopentane cooled by liquid nitrogen. Blocks were wrapped in aluminum and stored at -80°C

pending quality control (staging, sexing and euploidy). One male heart at 8.4 pcw (HEF_DN_F0000476) and a female heart at 9.7 pcw (HEF_DN_F0000298) were chosen for spatial transcriptomic library generation. Blocks were trimmed to a 6.5 x 6.5 mm square, sectioned at 10 µm and slices melted onto one of: a 10X Genomics Tissue Optimization Slide to determine optimal time for protease digestion (18 minutes); Superfrost Plus (Menzel Gläser) slides for subsequent validation; or a Visium v1 Spatial Gene Expression slide (10X Genomics, one individual heart per slide). All were stored with desiccant at -20°C until further processing. As per the Visium v1 Gene Expression protocol, sections were stained in hematoxylin/eosin (H&E), coverslipped in a glycerol-based solution, imaged at 10X on an AxioScan 7 (Zeiss), then barcoded mRNA content was captured to generate DNA libraries. Sequencing of libraries was done by the GBiM platform as above.

After standard quality control, reads were aligned to the human reference genome (GRCh38-2020-A) with SpaceRanger v2.0.1 to generate a UMI count matrix per spot. Spots overlaid by tissue were designated from the H&E scans using the manual alignment tool from Loupe Browser v6.4.1 (10X Genomics) to retain only tissue-associated barcodes. Two poorly adherent tissue slices of the 8.4 pcw heart were excluded from further analyses. Two (A90-A91) and four slices (A17-A20) of the younger and older hearts were retained for further analysis and respectively had >3000 median UMI counts per spot under tissue.

Analysis, visualization and integration of spatial transcriptomic datasets

Spatial transcriptomic data were first explored by individual section using the provided Loupe software v8.0 (10X Genomics) and in further depth with Seurat v5.0.3 (Hao et al., 2024), using the first 30 dimensions. The number of clusters (communities based on a shared nearest-neighbor modularity optimization-based algorithm) was obtained by applying resolutions of 0.45 for section A17, 0.65 for section A18, 0.35 for section A19, 0.20 for section A20, 0.30 for section A90 and 0.35 for section A91. The R package ‘scTransform’ was used for normalization and variance stabilization (Hafemeister and Satija, 2019). After individual slices were examined and compared, data were integrated using Seurat v5.0.3, spots re-clustered by applying resolutions of 0.47, and clusters manually annotated based on prior knowledge of DEGs and anatomy.

Spatial transcriptomic data deconvolution

RCTD (robust cell-type decomposition (Cable et al., 2022)) from the ‘SpaceX’ (spatially dependent gene co-expression network) R package was used to deconvolute spatial

transcriptomic data. Using our integrated snRNAseq data as a reference, we mapped the 21 identified cell types to their spatial distributions on each section. The RCTD “full mode” was used to assign any number of cell types per spot of spatial transcriptomic data. Cell-type weights for each slice, which represent the estimated proportion of each cell type for each spot, were assigned and normalized to sum to 1. This was represented in the form of pie charts for each spot and projected on to the spatial transcriptomic section images to show the estimated proportions of each cell type per spot. CmMit was manually excluded from the snRNAseq reference for deconvolution as it prevented visualization of the true cell type heterogeneity in each spot with strong expression of generic genes as *MYH6* and *MYH7, TOP2A*, *MT-CO1* and *COL1A2*, masking other cell types in the other groups.

Functional analyses of term annotations

Enrichr-KG (Evangelista et al., 2023) analyses of gene lists were annotated by terms from 2021-2022 releases of the Descartes human fetal atlas (Cao et al., 2020), the Human Gene Atlas (Mabbott et al., 2013), the Tabula Sapiens (The Tabula Sapiens Consortium, 2022) Gene Ontology Biological Processes, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) Human Gene and Pathway databases. Boolean operations on lists of genes were made with GeneVenn (Pirooznia et al., 2007) and/or “Draw Venn Diagrams” (Vandepeer, 2011).

Histology

Hematoxylin and eosin staining

Heart formalin-fixed, paraffin-embedded (FFPE) sections were deparaffinized twice in xylene and progressively rehydrated in ethanol baths of decreasing concentrations diluted in H_2O. Rehydrated sections were incubated in Mayer’s hematoxylin (Sigma-Aldrich, MHS32) for 2 minutes, rinsed in H_2O, differentiated briefly in Acid Alcohol (Sigma-Aldrich, A3179), rinsed again for 2 minutes then incubated 5 minutes in aqueous eosin Y (Sigma-Aldrich, HT110232). Sections were dehydrated then mounted with EuKitt (Sigma-Aldrich, 03989). Heart sections were imaged using a Zeiss AxioScan Z1.

Immunofluorescence

Immunofluorescence on cryosection

Hearts were collected in 1X PBS (phosphate-buffered saline) at 4°C and fixed in 4% aqueous buffered paraformaldehyde in PBS (PFA, Electron Microscopy Sciences, 15714) overnight. Hearts were then rinsed in 1X PBS and placed in 15% followed by 30% sucrose in 1X PBS until equilibration at 4°C. Hearts were embedded in OCT (Optimal Cutting Temperature compound, Leica Biosystems, 3801480), oriented, flash-frozen over dry ice or liquid nitrogen and stored at -80°C. Hearts were sectioned to 10 µm thickness using a Leica cryostat. Sections were washed in 1X PBS and fixed 10 minutes in 2% PFA. Sections were permeabilized in 0.1% Triton X-100 (Sigma-Aldrich) in PBS for 20 minutes then incubated in an excess of blocking solution (1% bovine serum albumin, Sigma-Aldrich, A7906; 1% inactivated horse serum, GE Healthcare B15-023; 0.1% Tween-20, Sigma-Aldrich, P1379; in 1X PBS) for 1h. Sections were incubated with primary antibodies diluted in blocking solution overnight at 4°C. After rinsing in PBS, heart sections were incubated with secondary antibodies (Invitrogen, ThermoFisher Scientific) diluted to 1/500 in blocking solution containing 300 nM DAPI (4',6-diamidino-2-phénylindole) at room temperature for 2 hours. After washes in 1X PBS, slides were mounted in Fluoromount-G (ThermoFisher Scientific, 00-4958-02). Sections were imaged using a Zeiss LSM800 confocal microscope or a Zeiss AxioImager Z2 stereomicroscope with an ApoTome device.

Rabbit anti-ALDH1A1 from GeneTex (GTX123973) was diluted to 1/100. Rabbit anti-CRABP2 (Sigma, HPA004135) was diluted to 1/100. Mouse anti-MYL7 (also known as MLC-2A, Synaptic Systems, 311011) was used at 1/200. Mouse anti-MYH6 (Sigma, AMAb90950) and mouse anti-MYH7 (Abcam ab207926) were diluted to 1/500. Highly cross-absorbed secondary antibodies against relevant species, coupled to AlexaFluor-488, -555 or -647, were all purchased from Thermo Fisher Scientific.

Immunofluorescence on whole heart

Hearts were collected in cold HBSS, rinsed in PBS and fixed 1 hour in PFA at 4°C with gentle agitation. After washes in 1X PBS, hearts were dehydrated at room temperature in progressive methanol baths (50%, 80% and 100% methanol, 1.5 hour each under agitation). Hearts were then incubated overnight at 4°C in a solution of 80% methanol, 6% H_2O_2. After progressive rehydration, hearts were incubated in PBSGT (0.2% porcine skin gelatin (VWR 24350.262), 0.5% Triton X-100, 0.01% sodium azide (Sigma-Aldrich, S2002) in 1X PBS) during 3 days under constant agitation at room temperature. Hearts were then incubated with primary antibodies against CD31/PECAM1 (Cell Signaling Technology, #3528, clone 89C2 at 1/200) and PRPH1 (Sigma-Aldrich AB1530, at 1/1000) at 37°C for 17 to 24 days proportional to relative size, under constant mild agitation. Alternatively, they were incubated in antibodies

against ALDH1A1 (Abcam ab52492 clone EP1933Y), S100B (GeneTex GTX14849 clone 4B3), EGR2 (Millipore ABE1374), FABP4 (Bioss BSM-51247M clone 5C1), POSTN (Origene TA804575S clone OTI2B2), or MYH6 (Sigma-Aldrich AMAb90950 clone CL2162). After 24h of washing in multiple changes of PBT (0.5% Triton-X100 in 1X PBS), hearts were incubated with secondary antibodies diluted to 1/1000 (ThermoFisher Scientific, Invitrogen) in PBSGT during 2 days at 37°C. After multiple washes in PBT, hearts were cleared (Belle et al., 2017). After progressive dehydration in methanol, anhydrous hearts were incubated overnight in glass vials with a solution containing two parts dichloromethane (Sigma-Aldrich, 270997, CAS 75-09-2), one part methanol, followed by 100% dichloromethane overnight. Hearts were next incubated in 100% (di)benzyl ether (Sigma-Aldrich, 108014, CAS 103-50-4) for at least 7 hours. Transparent hearts were conserved in benzyl ether but transferred to ethyl cinnamate (Sigma-Aldrich, 112372, CAS 103-36-6) for imaging with a Blaze light-sheet UltraMicroscope (Miltenyu Biotec). Imaris software (Oxford Instruments) was used to reconstruct 3D images.

In situ hybridization on paraffin sections with RNAscope

Hearts were collected in cold 1X PBS and fixed overnight in 4% PFA. Hearts were then rinsed in 1X PBS and dehydrated in progressive ethanol baths until complete dehydration in 100% ethanol, equilibrated to xylene (twice, 1 hour each), embedded in paraffin and stored at 4°C. 10 µm thick sections were made using a microtome and mounted on SuperFrost Ultra Plus GOLD slides (Microm Microtech, Epredia, K5800AMNZ72). RNAscope *in situ* hybridization was performed using the RNAscope Multiplex Fluorescent v2 kit from Advanced Cell Diagnostics (ACD), following the manufacturer's protocol. Before the probe hybridization step, sections were incubated in the "protease PLUS" solution for 15 minutes at 40°C. *CRABP1* (855321-C2), *ISL1* (478591-C3), *PRRX1* (511001-C2), *PENK* (548301-C3), *MEOX1* (564321), *ACTA1* (454311) and *BMP10* (459641) RNAscope probes were then hybridized. Nuclei were counterstained with the provided DAPI solution. Opal 520, Opal 570 and/or Opal 690 reagent packs (Akoya Biosciences) were used at 1/200 dilution to localize probes. Heart sections were imaged with tiling on a Zeiss LSM800 confocal microscope and stitched with Zen 3.7 software.

Websites

Loupe Browser

https://www.10xgenomics.com/products/loupe-browser/downloads/eula

Seurat vignettes and use-cases

https://satijalab.org/seurat/articles/pbmc3k_tutorial

https://satijalab.org/seurat/articles/sctransform_vignette

https://satijalab.org/seurat/articles/spatial_vignette

Enrichr-KG

https://maayanlab.cloud/enrichr-kg

Venn diagrams and subsets of co-expression

https://www.bioinformatics.org/gvenn/

https://bioinformatics.psb.ugent.be/webtools/Venn/

Human Developmental Cell Atlas (France)

https://hudeca.com/

https://hudeca.genouest.org/

References

Alexanian, M., Przytycki, P. F., Micheletti, R., Padmanabhan, A., Ye, L., Travers, J. G., Gonzalez-Teran, B., Silva, A. C., Duan, Q., Ranade, S. S., et al. (2021). A transcriptional switch governs fibroblast activation in heart disease. *Nature* 595, 438–443.

Alharbi, H. O., Sugden, P. H. and Clerk, A. (2024). Mitogen-activated protein kinase signalling in rat hearts during postnatal development: MAPKs, MAP3Ks, MAP4Ks and DUSPs. *Cell. Signal.* 124, 111397.

Ali, S. R., Ranjbarvaziri, S., Talkhabi, M., Zhao, P., Subat, A., Hojjat, A., Kamran, P., Müller, A. M. S., Volz, K. S., Tang, Z., et al. (2014). Developmental heterogeneity of cardiac fibroblasts does not predict pathological proliferation and activation. *Circ. Res.* 115, 625–635.

Anderson, R. H., Webb, S., Brown, N. A., Lamers, W. and Moorman, A. (2003). Development of the Heart: (1) Formation of the Cardiac Chambers and Arterial Trunks. *Heart* 89, 806–814.

Argiro, L., Chevalier, C., Choquet, C., Nandkishore, N., Ghata, A., Baudot, A., Zaffran, S. and Lescroart, F. (2023). Cardiopharyngeal mesoderm specification into cardiac and skeletal muscle lineages in gastruloids. 2023.05.15.540476.

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wärdell, E., Custodio, J., Reimegård, J., Salmén, F., et al. (2019). A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell* 179, 1647-1660.e19.

Assou, S., Girault, N., Plinet, M., Bouckenheimer, J., Sansac, C., Combe, M., Mianné, J., Bourguignon, C., Fieldes, M., Ahmed, E., et al. (2020). Recurrent Genetic

Abnormalities in Human Pluripotent Stem Cells: Definition and Routine Detection in Culture Supernatant by Targeted Droplet Digital PCR. *Stem Cell Rep.* 14, 1–8.

Belle, M., Godefroy, D., Couly, G., Malone, S. A., Collier, F., Giacobini, P. and Chédotal, A. (2017). Tridimensional Visualization and Analysis of Early Human Development. *Cell* 169, 161-173.e12.

Bergmann, O., Bhardwaj, R. D., Bernard, S., Zdunek, S., Barnabé-Heider, F., Walsh, S., Zupicich, J., Alkass, K., Buchholz, B. A., Druid, H., et al. (2009). Evidence for cardiomyocyte renewal in humans. *Science* 324, 98.

Blain, R., Couly, G., Shotar, E., Blévinal, J., Toupin, M., Favre, A., Abjaghou, A., Inoue, M., Hernández-Garzón, E., Clarençon, F., et al. (2023). A tridimensional atlas of the developing human head. *Cell* 186, 5910-5924.e17.

Blair, E., Redwood, C., Ashrafian, H., Oliveira, M., Broxholme, J., Kerr, B., Salmon, A., Ostman-Smith, I. and Watkins, H. (2001). Mutations in the gamma(2) subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum. Mol. Genet.* 10, 1215–1220.

Boeva, V., Louis-Brennetot, C., Peltier, A., Durand, S., Pierre-Eugène, C., Raynal, V., Etchevers, H. C., Thomas, S., Lermine, A., Daudigeos-Dubus, E., et al. (2017). Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat. Genet.* 49, 1408–1413.

Boppana, A., Lee, S., Malhotra, R., Halushka, M., Gustilo, K. S., Quardokus, E. M., Herr, B. W., Börner, K. and Weber, G. M. (2023). Anatomical structures, cell types, and biomarkers of the healthy human blood vasculature. *Sci. Data* 10, 452.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420.

Cable, D. M., Murray, E., Zou, L. S., Goeva, A., Macosko, E. Z., Chen, F. and Irizarry, R. A. (2022). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* 40, 517–526.

Cao, J., O'Day, D. R., Pliner, H. A., Kingsley, P. D., Deng, M., Daza, R. M., Zager, M. A., Aldinger, K. A., Blecher-Gonen, R., Zhang, F., et al. (2020). A human cell atlas of fetal gene expression. *Science* 370, eaba7721.

Chen, H., Shi, S., Acosta, L., Li, W., Lu, J., Bao, S., Chen, Z., Yang, Z., Schneider, M. D., Chien, K. R., et al. (2004). BMP10 is essential for maintaining cardiac growth during murine cardiogenesis. *Dev. Camb. Engl.* 131, 2219–2231.

Chen, T., Song, S., Jiang, H., Lian, H. and Hu, S. (2022). Single Cell Sequencing Reveals Mechanisms of Persistent Truncus Arteriosus Formation after PDGFRα and PDGFRβ Double Knockout in Cardiac Neural Crest Cells. *Genes* 13, 1708.

Clay, M. R. and Halloran, M. C. (2014). Cadherin 6 promotes neural crest cell detachment via F-actin regulation and influences active Rho distribution during epithelial-to-mesenchymal transition. *Dev. Camb. Engl.* 141, 2506–2515.

Cranley, J., Kanemaru, K., Bayraktar, S., Knight-Schrijver, V., Pett, J. P., Polanski, K., Dabrowska, M., Mulas, I., Richardson, L., Semprich, C., et al. (2024). Multiomic analysis reveals developmental dynamics of the human heart in health and disease. 2024.04.29.591736.

D'Antonio, M., Nguyen, J. P., Arthur, T. D., Matsui, H., Donovan, M. K. R., D'Antonio-Chronowska, A. and Frazer, K. A. (2022). In heart failure reactivation of RNA-binding proteins is associated with the expression of 1,523 fetal-specific isoforms. *PLoS Comput. Biol.* 18, e1009918.

De Bono, C., Thellier, C., Bertrand, N., Sturny, R., Jullian, E., Cortes, C., Stefanovic, S., Zaffran, S., Théveniau-Ruissy, M. and Kelly, R. G. (2018). T-box genes and retinoic acid signaling regulate the segregation of arterial and venous pole progenitor cells in the murine second heart field. *Hum. Mol. Genet.* 27, 3747–3760.

Deng, Y., He, Y., Xu, J., He, H. and Li, G. (2023). Heterogeneity and Functional Analysis of Cardiac Fibroblasts in Heart Development. *BioRxiv Prepr. Serv. Biol.* 2023.07.30.551164.

Dettman, R. W., Denetclaw, W., Ordahl, C. P. and Bristow, J. (1998). Common epicardial origin of coronary vascular smooth muscle, perivascular fibroblasts, and intermyocardial fibroblasts in the avian heart. *Dev. Biol.* 193, 169–81.

Dewing, J. M., Saunders, V., O'Kelly, I. and Wilson, D. I. (2022). Defining cardiac cell populations and relative cellular composition of the early fetal human heart. *PLOS ONE* 17, e0259477.

El Robrini, N., Etchevers, H. C., Ryckebüsch, L., Faure, E., Eudes, N., Niederreither, K., Zaffran, S. and Bertrand, N. (2016). Cardiac outflow morphogenesis depends on effects of retinoic acid signaling on multiple cell lineages. *Dev. Dyn.* 245, 388–401.

Etchevers, H. C., Vincent, C., Le Douarin, N. M. and Couly, G. F. (2001). The cephalic neural crest provides pericytes and smooth muscle cells to all blood vessels of the face and forebrain. *Development* 128, 1059–68.

Evangelista, J. E., Xie, Z., Marino, G. B., Nguyen, N., Clarke, D. J. B. and Ma'ayan, A. (2023). Enrichr-KG: bridging enrichment analysis across multiple libraries. *Nucleic Acids Res.* 51, W168–W179.

Evtouchenko, L., Studer, L., Spenger, C., Dreher, E. and Seiler, R. W. (1996). A mathematical model for the estimation of human embryonic and fetal age. *Cell Transplant.* 5, 453–464.

Farah, E. N., Hu, R. K., Kern, C., Zhang, Q., Lu, T.-Y., Ma, Q., Tran, S., Zhang, B., Carlin, D., Monell, A., et al. (2024). Spatially organized cellular communities form the developing human heart. *Nature* 627, 854–864.

Gaussin, V., Morley, G. E., Cox, L., Zwijsen, A., Vance, K. M., Emile, L., Tian, Y., Liu, J., Hong, C., Myers, D., et al. (2005). Alk3/Bmpr1a receptor is required for development of the atrioventricular canal into valves and annulus fibrosus. *Circ. Res.* 97, 219–226.

George, D., Ahrens, P. and Lambert, S. (2018). Satellite glial cells represent a population of developmentally arrested Schwann cells. *Glia* 66, 1496–1506.

Gerrelli, D., Lisgo, S., Copp, A. J. and Lindsay, S. (2015). Enabling research with human embryonic and fetal tissue resources. *Dev. Camb. Engl.* 142, 3073–3076.

Gittenberger-de Groot, A. C., Vrancken Peeters, M. P., Mentink, M. M., Gourdie, R. G. and Poelmann, R. E. (1998). Epicardium-derived cells contribute a novel population to the myocardial wall and the atrioventricular cushions. *Circ. Res.* 82, 1043–1052.

Gollob, M. H., Seger, J. J., Gollob, T. N., Tapscott, T., Gonzales, O., Bachinski, L. and Roberts, R. (2001). Novel PRKAG2 mutation responsible for the genetic syndrome of ventricular preexcitation and conduction system disease with childhood onset and absence of cardiac hypertrophy. *Circulation* 104, 3030–3033.

Hafemeister, C. and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 20, 296.

Hailstones, D., Barton, P., Chan-Thomas, P., Sasse, S., Sutherland, C., Hardeman, E. and Gunning, P. (1992). Differential regulation of the atrial isoforms of the myosin light chains during striated muscle development. *J. Biol. Chem.* 267, 23295–23300.

Haniffa, M., Taylor, D., Linnarsson, S., Aronow, B. J., Bader, G. D., Barker, R. A., Camara, P. G., Camp, J. G., Chédotal, A., Copp, A., et al. (2021). A roadmap for the Human Developmental Cell Atlas. *Nature* 597, 196–205.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573-3587.e29.

Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., et al. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* 42, 293–304.

Hennings, E., Blum, S., Aeschbacher, S., Coslovsky, M., Knecht, S., Eken, C., Lischer, M., Paladini, R. E., Krisai, P., Reichlin, T., et al. (2023). Bone Morphogenetic Protein 10-A Novel Biomarker to Predict Adverse Outcomes in Patients With Atrial Fibrillation. *J. Am. Heart Assoc.* 12, e028255.

Hern, W. M. (1984). Correlation of fetal age and measurements between 10 and 26 weeks of gestation. *Obstet. Gynecol.* 63, 26–32.

Hill, M. C., Kadow, Z. A., Li, L., Tran, T. T., Wythe, J. D. and Martin, J. F. (2019). A cellular atlas of Pitx2-dependent cardiac development. *Dev. Camb. Engl.* 146, dev180398.

Ho, S. Y., Anderson, R. H. and Sánchez-Quintana, D. (2002). Atrial structure and fibres: morphologic bases of atrial conduction. *Cardiovasc. Res.* 54, 325–336.

Hou, X., Si, X., Xu, J., Chen, X., Tang, Y., Dai, Y. and Wu, F. (2024). Single-cell RNA sequencing reveals the gene expression profile and cellular communication in human fetal heart development. *Dev. Biol.* 514, 87–98.

Hulin, A., Hortells, L., Gomez-Stallons, M. V., O'Donnell, A., Chetal, K., Adam, M., Lancellotti, P., Oury, C., Potter, S. S., Salomonis, N., et al. (2019). Maturation of heart valve cell populations during postnatal remodeling. *Development* 146, dev173047.

Jiang, X., Rowitch, D. H., Soriano, P., McMahon, A. P. and Sucov, H. M. (2000). Fate of the mammalian cardiac neural crest. *Development* 127, 1607–1616.

Jiang, X., Lian, X., Wei, K., Zhang, J., Yu, K., Li, H., Ma, H., Cai, Y. and Pang, L. (2024). Maturation of pluripotent stem cell-derived cardiomyocytes: limitations and challenges from metabolic aspects. *Stem Cell Res. Ther.* 15, 354.

Kaltenecker, D., Al-Maskari, R., Negwer, M., Hoeher, L., Kofler, F., Zhao, S., Todorov, M., Rong, Z., Paetzold, J. C., Wiestler, B., et al. (2024). Virtual reality-empowered deep-learning analysis of brain cells. *Nat. Methods* 21, 1306–1315.

Kanemaru, K., Cranley, J., Muraro, D., Miranda, A. M. A., Ho, S. Y., Wilbrey-Clark, A., Patrick Pett, J., Polanski, K., Richardson, L., Litvinukova, M., et al. (2023). Spatially resolved multiomics of human cardiac niches. *Nature* 619, 801–810.

Knight-Schrijver, V. R., Davaapil, H., Bayraktar, S., Ross, A. D. B., Kanemaru, K., Cranley, J., Dabrowska, M., Patel, M., Polanski, K., He, X., et al. (2022). A single-cell comparison of adult and fetal human epicardium defines the age-associated changes in epicardial activity. *Nat. Cardiovasc. Res.* 1, 1215–1229.

Krane, M., Dreßen, M., Santamaria, G., My, I., Schneider, C. M., Dorn, T., Laue, S., Mastantuono, E., Berutti, R., Rawat, H., et al. (2021). Sequential Defects in Cardiac Lineage Commitment and Maturation Cause Hypoplastic Left Heart Syndrome. *Circulation* 144, 1409–1428.

Lalani, S. R., Thakuria, J. V., Cox, G. F., Wang, X., Bi, W., Bray, M. S., Shaw, C., Cheung, S. W., Chinault, A. C., Boggs, B. A., et al. (2009). 20p12.3 microdeletion predisposes to Wolff-Parkinson-White syndrome with variable neurocognitive deficits. *J. Med. Genet.* 46, 168–175.

Lardenois, A., Suglia, A., Moore, C., Evrard, B., Noël, L., Rivaud, P., Besson, A., Toupin, M., Léonard, S., Lesné, L., et al. (2023). Single-cell transcriptome landscape of developing fetal gonads defines somatic cell lineage specification in humans. 2023.08.07.552336.

Lázár, E., Mauron, R., Andrusivová, Ž., Foyer, J., He, M., Larsson, L., Shakari, N., Salas, S. M., Avenel, C., Sariyar, S., et al. (2024). Spatial Dynamics of the Developing Human Heart. 2024.03.12.584577.

Le Lièvre, C. S. and Le Douarin, N. M. (1975). Mesenchymal derivatives of the neural crest: analysis of chimaeric quail and chick embryos. *J. Embryol. Exp. Morphol.* 34, 125–154.

Lescroart, F., Chabab, S., Lin, X., Rulands, S., Paulissen, C., Rodolosse, A., Auer, H., Achouri, Y., Dubois, C., Bondue, A., et al. (2014). Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development. *Nat. Cell Biol.* 16, 829–840.

Leshem, R., Baker, S. M., Mallen, J., Wang, L., Dark, J., Sharrocks, A. D., Hanley, K. P., Hanley, N. A., Rattray, M., Bamforth, S. D., et al. (2024). A cell atlas of the developing human outflow tract of the heart and its adult derivatives. 2023.04.05.535627.

Litviňuková, M., Talavera-López, C., Maatz, H., Reichart, D., Worth, C. L., Lindberg, E. L., Kanda, M., Polanski, K., Heinig, M., Lee, M., et al. (2020). Cells of the adult human heart. *Nature* 588, 466–472.

Liu, C., Liu, W., Palie, J., Lu, M. F., Brown, N. a and Martin, J. F. (2002). Pitx2c patterns anterior myocardium and aortic arch vessels and is required for local cell movement into atrioventricular cushions. *Development* 129, 5081–91.

Liu, K., Yu, W., Tang, M., Tang, J., Liu, X., Liu, Q., Li, Y., He, L., Zhang, L., Evans, S. M., et al. (2018). A dual genetic tracing system identifies diverse and dynamic origins of cardiac valve mesenchyme. *Development* 145, dev167775.

Liu, X., Chen, W., Li, W., Li, Y., Priest, J. R., Zhou, B., Wang, J. and Zhou, Z. (2019). Single-Cell RNA-Seq of the Developing Cardiac Outflow Tract Reveals Convergent Development of the Vascular Smooth Muscle Cells. *Cell Rep.* 28, 1346-1361.e4.

Liu, K., Jin, H., Zhang, S., Tang, M., Meng, X., Li, Y., Pu, W., Lui, K. O. and Zhou, B. (2023). Intercellular genetic tracing of cardiac endothelium in the developing heart. *Dev. Cell* 58, 1502-1512.e3.

Lu, P., Wu, B., Feng, X., Cheng, W., Kitsis, R. N. and Zhou, B. (2022). Cardiac Myosin Heavy Chain Reporter Mice to Study Heart Development and Disease. *Circ. Res.* 131, 364–366.

Lv, F., Ge, X., Qian, P., Lu, X., Liu, D. and Chen, C. (2022). Neuron navigator 3 (NAV3) is required for heart development in zebrafish. *Fish Physiol. Biochem.* 48, 173–183.

Mabbott, N. A., Baillie, J. K., Brown, H., Freeman, T. C. and Hume, D. A. (2013). An expression atlas of human primary cells: inference of gene function from coexpression networks. *BMC Genomics* 14, 632.

Maldonado, E., Rangel-Huerta, E., González-Gómez, R., Fajardo-Alvarado, G. and Morillo-Velarde, P. S. (2019). Octopus insularis as a new marine model for evolutionary developmental biology. *Biol. Open* 8, bio046086.

Martelotto, L. G. (2021). S.O.F protocol for nuclei isolation from fresh and frozen tissues using OptiPrep® discontinuous gradient V.4.

McKellar, D. W., Walter, L. D., Song, L. T., Mantri, M., Wang, M. F. Z., De Vlaminck, I. and Cosgrove, B. D. (2021). Large-scale integration of single-cell transcriptomic data captures transitional progenitor states in mouse skeletal muscle regeneration. *Commun. Biol.* 4, 1280.

Miyamoto, K., Akiyama, M., Tamura, F., Isomi, M., Yamakawa, H., Sadahiro, T., Muraoka, N., Kojima, H., Haginiwa, S., Kurotsu, S., et al. (2018). Direct In Vivo Reprogramming with Sendai Virus Vectors Improves Cardiac Function after Myocardial Infarction. *Cell Stem Cell* 22, 91-103.e5.

Møller Nielsen, A. K., Dehn, A. M., Hjortdal, V. and Larsen, L. A. (2024). TBX5 variants and cardiac phenotype: A systematic review of the literature and a novel variant. *Eur. J. Med. Genet.* 68, 104920.

Moon, K. R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D. B., Chen, W. S., Yim, K., Elzen, A. van den, Hirn, M. J., Coifman, R. R., et al. (2019). Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* 37, 1482–1492.

Muhl, L., Genové, G., Leptidis, S., Liu, J., He, L., Mocci, G., Sun, Y., Gustafsson, S., Buyandelger, B., Chivukula, I. V., et al. (2020). Single-cell analysis uncovers fibroblast heterogeneity and criteria for fibroblast and mural cell identification and discrimination. *Nat. Commun.* 11, 1–18.

Norris, R. A., Potts, J. D., Yost, M. J., Junor, L., Brooks, T., Tan, H., Hoffman, S., Hart, M. M., Kern, M. J., Damon, B., et al. (2009). Periostin promotes a fibroblastic lineage pathway in atrioventricular valve progenitor cells. *Dev. Dyn.* 238, 1052–63.

Oetjen, S., Mahlke, C., Hermans-Borgmeyer, I. and Hermey, G. (2014). Spatiotemporal expression analysis of the growth factor receptor SorCS3. *J. Comp. Neurol.* 522, 3386–3402.

O'Rahilly, R. and Müller, F. (1987). *Developmental Stages in Humans: including a revision of Streeter's "Horizons" and a survey of the Carnegie collection*. Washington, D.C.: Carnegie Institution of Washington.

O'Rahilly, R. and Müller, F. (2010). Developmental stages in human embryos: revised and new measurements. *Cells Tissues Organs* 192, 73–84.

Pirooznia, M., Nagarajan, V. and Deng, Y. (2007). GeneVenn - A web application for comparing gene lists using Venn diagrams. *Bioinformation* 1, 420–422.

Ramesh, N., Escher, M. J. F., Mampell, M. M., Böhme, M. A., Götz, T. W. B., Goel, P., Matkovic, T., Petzoldt, A. G., Dickman, D. and Sigrist, S. J. (2021). Antagonistic interactions between two Neuroligins coordinate pre- and postsynaptic assembly. *Curr. Biol. CB* 31, 1711-1725.e5.

Rathjens, F. S., Blenkle, A., Iyer, L. M., Renger, A., Syeda, F., Noack, C., Jungmann, A., Dewenter, M., Toischer, K., El-Armouche, A., et al. (2021). Preclinical evidence for the therapeutic value of TBX5 normalization in arrhythmia control. *Cardiovasc. Res.* 117, 1908–1922.

Sacks, H. S. and Fain, J. N. (2007). Human epicardial adipose tissue: a review. *Am. Heart J.* 153, 907–917.

Sedmera, D., Pexieder, T., Vuillemin, M., Thompson, R. P. and Anderson, R. H. (2000). Developmental patterning of the myocardium. *Anat. Rec.* 258, 319–337.

Smirnov, A. V., Kontsevaya, G. V., Feofanova, N. A., Anisimova, M. V., Serova, I. A., Gerlinskaya, L. A., Battulin, N. R., Moshkin, M. P. and Serov, O. L. (2018). Unexpected phenotypic effects of a transgene integration causing a knockout of the endogenous Contactin-5 gene in mice. *Transgenic Res.* 27, 1–13.

Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V., et al. (2019). Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* 364, eaas9536.

Soonpaa, M. H., Kim, K. K., Pajak, L., Franklin, M. and Field, L. J. (1996). Cardiomyocyte DNA synthesis and binucleation during murine development. *Am. J. Physiol.* 271, H2183-2189.

Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E. and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 19, 477.

Strom, J., Bull, M., Gohlke, J., Saripalli, C., Methawasin, M., Gotthardt, M. and Granzier, H. (2024). Titin's cardiac-specific N2B element is critical to mechanotransduction during volume overload of the heart. *J. Mol. Cell. Cardiol.* 191, 40–49.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P. and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902.e21.

Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., Zhang, J., Soden, R., Hayakawa, M., Kreiman, G., et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci.* 101, 6062–6067.

Sweat, M. E., Cao, Y., Zhang, X., Burnicka-Turek, O., Perez-Cervantes, C., Arulsamy, K., Lu, F., Keating, E. M., Akerberg, B. N., Ma, Q., et al. (2023). Tbx5 maintains atrial identity in post-natal cardiomyocytes by regulating an atrial-specific enhancer network. *Nat. Cardiovasc. Res.* 2, 881–898.

Teichmann, U. and Kessel, M. (2004). Highly restricted BMP10 expression in the trabeculating myocardium of the chick embryo. *Dev. Genes Evol.* 214, 96–98.

The Tabula Sapiens Consortium (2022). The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* 376, eabl4896.

van Eif, V. W. W., Stefanovic, S., van Duijvenboden, K., Bakker, M., Wakker, V., de Gier-de Vries, C., Zaffran, S., Verkerk, A. O., Boukens, B. J. and Christoffels, V. M. (2019). Transcriptome analysis of mouse and human sinoatrial node cells reveals a conserved genetic program. *Dev. Camb. Engl.* 146, dev173161.

Vandepeer, Y. (2011). Venn Diagram tool. *Draw Venn Diagr.*

Vrancken Peeters, M.-P. F. M., Gittenberger-de Groot, A. C., Mentink, M. M. T., Poelmann, R. E., Groot, A. C. G., Mentink, M. M. T., Poelmann, R. E., Vrancken Peeters, M.-P. F. M., Gittenberger-de Groot, A. C., Mentink, M. M. T., et al. (1999). Smooth muscle cells and fibroblasts of the coronary arteries derive from epithelial-mesenchymal transformation of the epicardium. *Anat. Embryol. (Berl.)* 199, 367–378.

Waldo, K., Miyagawa-Tomita, S., Kumiski, D. and Kirby, M. L. (1998). Cardiac neural crest cells provide new insight into septation of the cardiac outflow tract: aortic sac to ventricular septal closure. *Dev. Biol.* 196, 129–44.

Wang, H. U., Chen, Z. F. and Anderson, D. J. (1998). Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. *Cell* 93, 741–753.

Wang, W., Li, X., Ding, X., Xiong, S., Hu, Z., Lu, X., Zhang, K., Zhang, H., Hu, Q., Lai, K. S., et al. (2023). Lymphatic endothelial transcription factor Tbx1 promotes an immunosuppressive microenvironment to facilitate post-myocardial infarction repair. *Immunity* 56, 2342-2357.e10.

White-Brown, A. M., Richard, M., Morency, A.-M., Maedler-Kron, C. and De Bie, I. (2024). PRKAG2-Related Lethal Congenital Glycogen Storage Disease of the Heart as Rare Cause of Fetal Hydrops With Bradycardia and Cardiomyopathy: Clinical Report and Literature Review. *Am. J. Med. Genet. A.* e63865.

Xu, Y., Zhang, T., Zhou, Q., Hu, M., Qi, Y., Xue, Y., Nie, Y., Wang, L., Bao, Z. and Shi, W. (2023). A single-cell transcriptome atlas profiles early organogenesis in human embryos. *Nat. Cell Biol.* 25, 604–615.

Zhao, J., Kennelly, J., Nalar, A., Kulathilaka, A., Sharma, R., Bai, J., Li, N. and Fedorov, V. V. (2023). Chamber-specific wall thickness features in human atrial fibrillation. *Interface Focus* 13, 20230044.

Acknowledgments

Prof. Aubert Agostini and staff of the Centre de Gynécologie Sociale de l'AP-HM in Marseille obtained patient consents and provided most of the tissue donations used in this work, to which effort Maryne Toupin also contributed from Rennes. Laurence Noël built and helped curate the ontologies used in HuDeCA's OpenSpecimen database. The authors gratefully acknowledge personnel of the GBiM and imaging platforms of Marseille Medical Genetics, in particular Catherine Robert, Christel Castro and Dr. Valérie Delague, as well as Céline Chevalier, Nicolas Lenfant, Alexandre Atkinson, Deniz Harcanoğlu and Dr. Raphaël Blain for their assistance in analyzing the sequencing and imaging data presented.

Funding: Institut National de la Santé et de la Recherche Médicale (INSERM) Programme Transversale "HuDeCA" (MH, YG, AC, SMG, SZ, HCE)

Association Française contre les Myopathies grant "MoThARD" (SK, AB, SZ, HCE)

CDB is supported by a grant from the Fondation pour la Recherche Médicale.

This publication is part of the Human Cell Atlas. https://www.humancellatlas.org/publications/

Author contributions:

SZ and HCE conceptualized the study and designed the experiments. CDB, SZ and HCE refined methodology, interpreted the data and wrote the original manuscript draft, to which AB and SK also contributed. MH, MM, WS, CM and SMG processed samples and performed quality control measures. MH carried out the histological, immunofluorescence and high-throughput sequencing-based experiments. CDB, CH, SK, SR, YX, AB and HCE designed, refined and analyzed bioinformatics experiments. AL, IV, YG and AC contributed to the 3D immunofluorescence experimental design, execution and data analyses. CDB, ZB, IV, AC, AB, SZ and HCE supervised the analyses. All authors reviewed and provided input on the final manuscript.

Competing interests: Authors declare that they have no competing interests.

Correspondence and requests for materials should be addressed to heather.etchevers@inserm.fr or stephane.zaffran@inserm.fr

Data and materials availability:

All data, code, and materials used in the analyses are available either in the main text and/or the Supplementary Materials, or deposited in the following repositories: Sequencing data and snRNAseq matrices are available from the Gene Expression Omnibus (GEO), https://www.ncbi.nlm.nih.gov/geo/ under accession GSExxxxxx; code from https://github.com/BAUDOTlab/Human_fetal_heart_atlas; movies, large images and interactive sections (*.cloupe and *.rds files) for spatial transcriptomics are available from Figshare at https://figshare.com/projects/Multi-modal_refinement_of_the_human_heart_atlas_during_the_first_gestational_trimester/213151 . Some images are reproduced at the Human Developmental Biology Resource (UK) fetal cardiac anatomy website (https://bit.ly/HumanFetalHeart).

Human Cell Atlas data from (Xu et al., 2023) is available both at:

https://explore.data.humancellatlas.org/projects/e255b1c6-1143-4fa6-83a8-528f15b41038 and from GEO under accession GSE157329.

Figure legends

Figure 1. Profiling human fetal heart development at the resolution of individual nuclei

(**A**) Samples of nuclei were derived from three whole dissociated hearts to generate distinct single-cell RNA-seq datasets at 8.6 post-conceptional weeks (pcw, confirmed XX genotype), 9.0 pcw (XY) and 10.7 pcw (XY). Samples of spatial transcriptomic analysis were derived from two whole dissociated hearts to collect cryosections and generate Visium datasets at 8.4 pcw (XY, 2 sections) and 9.7 pcw (XX, 4 sections). (**B**) Integrated UMAP representation of 49,227 profiled nuclei coloured by cell type. (**C**) UMAP representation of profiled nuclei as in (B) but separated by sample. Arrows indicate cluster 12 (SMC) in each, but that all cell types are present at each sample. (**D**) UMAP plots per sample indicating that all cell classes are also represented in each sample. (**E**) Dot plot of top marker genes for each Class (y axis: cardiomyocytes (I), endocardiovascular cells (II), stroma (III), epicardium (IV), blood (V) and neural crest progeny (VI). The size of the dot represents the percent of nuclei with transcripts at non-zero levels, and color intensity represents average log-normalized expression of the gene where relative abundance of typical markers is indicative of cell class. (**F**) UMAP feature plots of representative gene expression in Classes I to VI for *MYH7* (cardiomyocytes), *PECAM1* (endothelial and endocardial cells), *EBF2* (stroma), *TBX18* (epicardium), *SPP1* (immune cells) and *NRXN1* (neural crest). (**G**) Heatmap showing selected genes expressed in the nine cardiomyocyte clusters of Class I compared with the minority populations of lymphatic endothelium, Schwann cell precursors and neuroendocrine cells. Each column displays gene expression of an individual cell and genes are listed in the rows.

Figure 2. Trajectory analysis of Classes I and III over time suggests a transient cardiomyofibroblast progenitor present at 4-6 pcw

(**A**) PHATE visualization of cardiomyocyte and stromal cells, colored by developmental stage, from cardiac single-cell RNA-seq data from seven human embryos at Carnegie stage (CS)12, CS13-14 and CS15-16 (Xu et al., 2023), one human heart at 6.5-7 pcw (Asp et al., 2019), and the snRNAseq data from three hearts at 8.6, 9.0 and 10.7 pcw (this study). (**B**) Subset of PHATE plot generated in A with scRNA-seq data at 4-6 pcw, colored by original clusters identified in Xu et al., 2023. (**C**) Subset of PHATE plot generated in A with snRNAseq data at 8.6, 9.0 and 10.7 pcw, colored by clusters identified from their integration (cf. Fig. 1B). (**D**) Superimposition on PHATE plot in A-C of Slingshot trajectory analysis, predicting cluster lineage bifurcations over time (black lines) from cluster 16, the “cardiomyofibroblast” (CmF). (**E**) Confusion matrix showing percentages of each cell type from our snRNAseq clusters represented in C (y axis) and each newly generated cluster represented in D, after integration of all datasets (x axis). Fewer than 20% of the cells in cluster 16 (CmF) are present after 8.6 pcw. (**F**) Histogram of the percentages of cells/nuclei from each stage in each cluster from in D, aligned with the x-axis of E. Fb2, SMC and CmVCnd cell types are only present after 6.5 pcw, while CmA3, CmG2M and Fb3 were novel cell types represented in our dataset after 8.6 pcw. (**G**) The top panel describes four inheritance patterns (1-4) of the trajectory from cluster 16 (CmF) to 6 (CmMit) and 15 (Fb3). Bottom panel: DEGs of A>B are defined as A > 0.5 and A/B > two-fold. 1: CmF and CmMit > Fb3 (inherited DEGs). 2: CmMit > CmF and Fb3 (non-inherited). CmF and Fb3 > CmMit (inherited DEGs). Fb3 > CmF and CmMit (non-inherited). Patterns 1 and 4 best fit the data, such that CmF is closer to a cardiomyocyte than a stromal identity.

Figure 3. Spatial transcriptomics of 8.4 pcw and 9.7 pcw human heart sections demonstrate regionalized distributions of cardiac cell types

(**A**) Seurat UMAP (Uniform Manifold Approximation and Projection) plot of integrated spatial transcriptomic data from two 8.4 postconceptional week (pcw) heart sections and four 9.7 pcw heart sections. Each dot represents a spot covered by a histological section on the Visium spatial capture slides. 14 clusters of the developing heart corresponding to cardiac cell types are listed. (**B**) Visualization of clustering on each heart section after integration of spatial transcriptomic data. Cluster annotations and colors match those of the UMAP in A. (**C**) Spatial plots showing gene expression on heart sections, with highest expression in red and lower expression in blue. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta.

Figure 4. Immunostaining or transcript corroboration of selected Class I marker genes

(First column: **A, E, I, M**) UMAPs showing gene expression of *MYL7*, *MYH6*, *MYH7* and *BMP10* respectively in cardiomyocyte Class I, indicating atrial and ventricular subclasses. (**B, C**) Spatial plots showing expression of *MYL7* in representative sections of 8.4 pcw and 9.7 pcw hearts, highest in red and lowest in blue. (**D**) MYL7 protein detected by immunofluorescence was more strongly expressed in the atria than the ventricles of a 9.0 pcw heart section. (**F, G**) Spatial plots showing expression of *MYH6* in representative sections of 8.4 pcw and 9.7 pcw hearts, highest in red and lowest in blue. (**H**) MYH6 protein was more strongly expressed in the atria than the ventricles at 9.0 pcw. (**J, K**) Spatial plots showing expression of *MYH7* in representative sections of 8.4 and 9.7 pcw hearts, highest in red and lowest in blue. (**L**) MYH7 protein was detected by immunofluorescence in the ventricular myocardium of a 12.0 pcw heart section. (**N, O**) Spatial plots showing expression of *BMP10* in representative sections of 8.4 and 9.7 pcw hearts, highest in red and lowest in blue. (**P**) *BMP10* transcripts were detected by RNAscope in ventricular trabeculae and the pectinate muscles of the right atrium in a 9.0 pcw heart section. Ao, aorta; IVS, interventricular septum; LA, left atrium; LV, left ventricle; RA, right atrium; RV, right ventricle. Scale bars: 250μm.

Figure 5: Deconvolution of heart spatial transcriptomic data with snRNAseq data as reference reveals transcriptional heterogeneity in spatial transcriptomic spots

(**A, C**) Deconvolution of spatial transcriptomic data from 8.4 pcw heart section A91 (**A**) and 9.7 pcw section A17 (**C**) with RCTD software. The proportion of each cell type (cluster) identified in integrated snRNAseq data (Figure 1) is estimated and represented in pie charts for each spatial spot. Large pie charts represent the average proportion of each cell type in indicated spatial clusters (C0, C2, C3, C4, C10 and C13) identified in Figure 3B. (**B, D**) Calculated cell-type proportions for CmVCnD, CmV1, CmG2M, CmA1, CmA4, CmA3, FbVIC, CmV2/SubValvular, EnTh, EndoCd, Fb1, Fb2, SMC, EpiCd and CmA2/Node in each spatial spot of 8.4 pcw heart section A91 (**B**) and 9.7 pcw section A17 (**D**). Highest proportion is represented in red and lowest in blue. Note the coherency between the cell types identified in snRNAseq data and the predicted spatial localization of these types.

Figure 6: Spatial gene expression in outflow and atrioventricular valves

(**A-C, G-I**) Spatial plots showing distribution of stromal class transcripts on representative sections of 8.4 pcw heart, with highest expression in red and least expression in blue. (**C, I**) co-expression plots of *PENK* (red) and *ACTA1* or *MEOX1* (green) highest expression of each in most saturated color and of both in intermediate shades, with yellow being equal co-expression. (**C', I'**) Enlargement of the atrioventricular cushion region of section showed in panels **C** and **I**. (**D, E**) Spatial expression of *PENK* and *ACTA1* detected by RNAscope *in situ* hybridization in a 9.0 pcw heart section. (**F, F'**) Double-staining showing expression of *PENK* (red) and *ACTA1* (green) mRNAs in the same section. *PENK* is expressed in the atrioventricular valve leaflets, while *ACTA1* labeled the chordae tendineae (**F'**) Enlargement of the area delimited by a dotted square in the adjacent section of the section showed in panel F. (**J**) Expression of *PRRX1* was detected in the aortic valve of 7.0 pcw heart. (**K, L**) Spatial expression of *MEOX1* and *PRRX1* was detected by RNAscope *in situ* hybridization in 9.0 pcw heart sections. (**M, M'**) Double-staining showing expression of *PENK* (red) and *MEOX1* (green) mRNAs in the same section. (**M'**) Enlargement of the area delimited by a dotted square in the adjacent section of the section showed in panel M. *PENK* is strongly expressed by the inferior atrioventricular endocardial cushion, partly overlapping (arrows in **I'**) with the more restricted domain of *MEOX1*. Ao, aorta; CT, chordae tendineae; iavc, inferior atrioventricular endocardial cushion (Anderson et al., 2003); LA, left atrium; LCA, left coronary artery; LV, left ventricle; MV, mitral valve; RA, right atrium; RCA, right coronary artery; RV, right ventricle; TV, tricuspid valve; sb, subvalvular domain (septal bridge); *, blood clots. Arrows in **J** show coronary arteries. Scale bars: 250 μm except D, 500 μm.

Figure 7: Parasympathetic ganglia and nerve markers

(**A, B**) Spatial plots showing spatial expression of genes enriched in the conduction and nervous system genes on a representative 8.4 pcw heart section, with highest expression in red and lowest expression in blue. (**C**) Co-expression plots of *ISL1* (red) and *PRRX1* (green) highest expression of each in most saturated color and of both in intermediate shades, with yellow being equal co-expression. **(A', B'**) Spatially restricted transcription of *ISL1* and *PRRX1* were detected in the same sections by RNAscope *in situ* hybridization. (**C**) Co-expression of *ISL1* (green) and *PRRX1* (red) reveals that these genes are expressed in the parasympathetic ganglia. Ao, aorta; RA, right atrium. Scale bars: 250μm.

Figure 8. Three-dimensional reconstruction of vascular and nerve scaffolds over time

(**A**) UMAP visualization of *PRPH* expression in snRNAseq dataset. (**B**) Spatial plots showing spatial expression of *PRPH* gene in cardiac ganglion (g) on a representative section of 8.4 pcw heart, with highest expression in red and lowest expression in blue. (**C**) Spatial expression of PRPH was detected by immunofluorescence in a 9.0 pcw heart section. PRPH is expressed in the nerves (arrows) and ganglia (g). (**D**) UMAP visualization of *PECAM1* expression in snRNAseq dataset. (**E**) Spatial plots showing spatial expression of *PECAM1* gene on a representative section of 8.4 pcw heart, with highest expression in red and lowest expression in blue. (**F**) Spatial expression of PECAM1 (CD31) was detected by immunofluorescence in a 9.0 pcw heart section. PECAM1 is expressed in the endothelial cells of the vessels (arrows). (**G-L**) Whole-mount immunofluorescence against PECAM-1 (CD31, red (**G**) and turquoise (**I**)) and PRPH (yellow) on 9.0 pcw and 10.0 pcw hearts. (**G,L**) Light-sheet fluorescent microscopy (LSFM) images show parallel networks of vessels (PECAM1+) and nerves (PRPH+) on 9.0 and 10.0 pcw hearts. (**H,K**) Top views of hearts showing in **G** and **J**. (**I,L**) High magnification of LSFM images of the pulmonary valve show the absence of nerves around the valve leaflets. Ao, aorta; AoV, aortic valve; LCA, left coronary artery; LV, left ventricle; Pt, pulmonary trunk; RCA, right coronary artery; RV, right ventricle. Scale bars: 500μm (**G,J**) ; 400μm (**H**); 300μm (**I**); 150μm (**K**,**L**).

Supplementary data

Supplementary figure legends

Supplementary figure 1: Histological analysis of human fetal hearts

(A-I) Hematoxylin and eosin staining on different axial level sections of human fetal heart at 8.7 pcw (A-C), 9.7 pcw (D-F) and 11.0 pcw (G-I). RA, right atrium; LA, left atrium; Ao, aorta; PT, pulmonary trunk; CA, coronary artery; AoV, aortic valve; RV, right ventricle; LV, left ventricle; S, interventricular septum; sb, subvalvular domain (septal bridge); lAV, left atrioventricular valve; rAV, right atrioventricular valve; PV, pulmonary valve.

Supplementary figure 2: Cross-referencing annotations improves identity assignment

(**A**) Co-expression of *KCNQ5* and *INPP4B* in a UMAP plot of integrated snRNA-seq data of all nuclei from 8.4, 9.0 and 10.7 pcw hearts, showing enrichment of cells transcribing both genes in the CmG2M and CmV1 subpopulations of Class I cells (cardiomyocytes). (**B**) Enrichr-KG analysis (Evangelista et al., 2023) of Class VI clusters converges for cluster 18 on Schwann cells, glia and axon guidance, leading us to annotate it as Schwann cells and their precursors, while in (**C**) analysis of cluster 19 with the same parameters highlights different visceral neuron and synapse annotations. Combined with the knowledge that neuronal nuclei at the level of cardiac ganglia are parasympathetic, we annotated this cluster as “neuroendocrine” (as it also expressed *CHGA*, *CHGB*, *SST* and the Descartes term “chromaffin cells in Adrenal”, a neural crest endocrine derivative, was significantly enriched [$p= 3.43 \times 10^{-14}$]).

Supplementary figure 3: Determination of similarities across annotated datasets

(**A**) Similarity between cell types across datasets, calculated as 1/(1 + distance). The distance between cell types was determined using the Mehalanobis-like distance metric in Slingshot (Street et al., 2018) based on the top 30 principal components (PCs) of the integrated dataset. Color bars indicate cell type classes. (**B**) Cell type mapping across datasets based on the similarity in panel A. Cell types were matched between adjacent developmental stages.

Initially, the best-matching cell type was linked if similarity > 0.06. A secondary match was linked if the similarity z-score across all cell types was > 1 and similarity > 0.06. (**C**) PHATE embedding of the integrated dataset, with cells colored by developmental stages. (**D**) Cells from Asp et al., 2019 in the PHATE embedding of our integrated dataset, showing cell types as originally annotated.

Supplementary figure 4: Spatial transcriptomics of 8.4 pcw A90 heart section

(**A**) Seurat UMAP plot of spatial transcriptomic data from 8.4 pcw A90 heart section. (**B**) Spatial dimensionality plot displaying spatial distribution and cluster annotation. (**C**) Spatial plots showing spatial gene expression, with highest expression in red and lowest expression in blue. Example of marker genes of the myocardium, atrial and ventricular myocardium, trabeculae, aorta and valve are displayed. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta.

Supplementary figure 5: Spatial transcriptomics of 8.4 pcw A91 heart section

(**A**) Seurat UMAP plot of spatial transcriptomic data from 8.4 pcw A91 heart section. (**B**) Spatial dimensionality plot displaying spatial distribution and annotation of clusters. (**C**) Spatial plots showing spatial gene expression, with highest expression in red and lowest expression in blue. Example of marker genes of the myocardium, atrial and ventricular myocardium, trabeculae, aorta and valve are displayed. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta.

Supplementary figure 6: Spatial transcriptomics of 9.7 pcw A17 heart section

(**A**) Seurat UMAP plot of spatial transcriptomic data from 9.7 pcw A17 heart section. (**B**) Spatial dimensionality plot displaying spatial distribution and annotation of clusters. (**C**) Spatial plots showing spatial gene expression, with highest expression in red and lowest expression in blue. Example of marker genes of the myocardium, compact myocardium, atrial and ventricular

myocardium, trabeculae, aorta and valve are shown. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta.

Supplementary figure 7: Spatial transcriptomics of 9.7 pcw A18 heart section

(**A**) Seurat UMAP plot of spatial transcriptomic data from 9. 7 pcw heart section number 2. (**B**) Spatial dimensionality plot displaying spatial distribution and annotation of clusters. (**C**) Spatial plots showing spatial gene expression, with highest expression in red and lowest expression in blue. Example of marker genes of the myocardium, compact myocardium, atrial and ventricular myocardium, trabeculae, aorta and valve are shown. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta.

Supplementary figure 8: Spatial transcriptomics of 9.7 pcw A19 heart section

(**A**) Seurat UMAP plot of spatial transcriptomic data from 9.7 pcw heart section number 3. (**B**) Spatial dimensionality plot displaying spatial distribution and annotation of clusters. (**C**) Spatial plots showing spatial gene expression, with highest expression in red and lowest expression in blue. Example of marker genes of the myocardium, atrial and ventricular myocardium, trabeculae, aorta and valve are shown. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta.

Supplementary figure 9: Spatial transcriptomics of 9.7 pcw A20 heart section

(**A**) Seurat UMAP plot of spatial transcriptomic data from 9.7 pcw heart section number 4. (**B**) Spatial dimensionality plot displaying spatial distribution and cluster annotation. (**C**) Spatial plots showing spatial gene expression, with highest expression in red and lowest expression in blue. Example of marker genes of the myocardium, atrial and ventricular myocardium, trabeculae, aorta and valve are displayed. RV, right ventricle; LV, left ventricle; RA, right atrium; LA, left atrium; Ao, aorta.

Supplementary figure 10: Split UMAP plots of integrated spatial transcriptomic data, from two 8.4 pcw and four 9.7 pcw sections with cluster annotations.

Split UMAP plots of the integrated spatial transcriptomic data in Figure 3 according to section, with cluster annotations on the right.

Supplementary figure 11: Deconvolution of 8.4 pcw A90 heart section spatial transcriptomic data with integrated snRNAseq data as reference.

(**A**) Deconvolution of spatial transcriptomic data from 8.4 pcw A90 heart section with RCTD software. The proportion of each cell type identified in integrated snRNAseq data (Figure 1) is estimated and represented in pie charts for each spatial transcriptomic pixel. (**B**) Estimated cell type proportion for each snRNAseq cluster on each spatial pixel of 8.4 pcw A90 heart section, with highest estimated proportion in red and lower in blue. RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle.

Supplementary figure 12: Deconvolution of 8.4 pcw A91 heart section spatial transcriptomic data with integrated snRNAseq data as reference.

(**A**) Deconvolution of spatial transcriptomic data from 8.4 pcw A91 heart section with RCTD software. The proportion of each cell type identified in integrated snRNAseq data (Figure 1) is estimated and represented in pie charts for each spatial transcriptomic pixel. (**B**) Estimated cell type proportion for each snRNAseq cluster on each spatial pixel of 8.4 pcw A91 heart section, with highest estimated proportion in red and lower in blue. RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle.

Supplementary figure 13: Deconvolution of 9.7 pcw A17 heart section spatial transcriptomic data with integrated snRNAseq data as reference.

(**A**) Deconvolution of spatial transcriptomic data from 9.7 pcw heart section number 1 with RCTD software. The proportion of each cell type identified in integrated snRNAseq data (Figure 1) is estimated and represented in pie charts for each spatial transcriptomic pixel. (**B**) Estimated cell type proportion for each snRNAseq cluster on each spatial pixel of 9.7 pcw A17 heart section, with highest estimated proportion in red and lower in blue. RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle.

Supplementary figure 14: Deconvolution of 9.7 pcw A18 heart section spatial transcriptomic data with integrated snRNAseq data as reference.

(A) Deconvolution of spatial transcriptomic data from 9.7 pcw heart section number 2 with RCTD software. The proportion of each cell type identified in integrated snRNAseq data (Figure 1) is estimated and represented in pie charts for each spatial transcriptomic pixel. **(B)** Estimated cell type proportion for each snRNAseq cluster on each spatial pixel of 9.7 pcw A18 heart section, with highest estimated proportion in red and lower in blue. RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle.

Supplementary figure 15: Deconvolution of 9.7 pcw A19 heart section spatial transcriptomic data with integrated snRNAseq data as reference.

(A) Deconvolution of spatial transcriptomic data from 9.7 pcw heart section number 3 with RCTD software. The proportion of each cell type identified in integrated snRNAseq data (Figure 1) is estimated and represented in pie charts for each spatial transcriptomic pixel. **(B)** Estimated cell type proportion for each snRNAseq cluster on each spatial pixel of 9.7 pcw A19 heart section, with highest estimated proportion in red and lower in blue. RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle.

Supplementary figure 16: Deconvolution of 9.7 pcw A20 heart section spatial transcriptomic data with integrated snRNAseq data as reference.

(A) Deconvolution of spatial transcriptomic data from 9.7 pcw heart section number 4 with RCTD software. The proportion of each cell type identified in integrated snRNAseq data (Figure 1) is estimated and represented in pie charts for each spatial transcriptomic pixel. **(B)** Estimated cell type proportion for each snRNAseq cluster on each spatial pixel of 9.7 pcw A20 heart section, with highest estimated proportion in red and lower in blue. RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle.

Supplementary figure 17: Valve markers in spatial transcriptomic sections A91 and A17.

Spatial transcriptomic plots showing expression of valve marker genes on 8.4 pcw A91 heart **(A-G)** and 9.7 pcw A17 **(H-K)** sections, with highest expression in red and lowest in blue.

Supplementary figure 18: Parasympathetic ganglionic marker gene detection in spatial transcriptomic section A90.

Spatial transcriptomic plots showing expression of parasympathetic ganglia marker genes in the 8.4 pcw A90 heart section, with highest expression in red and lowest in blue.

Supplementary figure 19: Use-cases of 3D protein expression data

Segmentation of right and left coronary arteries **(A, B)** from 3D-reconstructed stacks of a 9.1 pcw male fetal heart (F0495) after anti-PECAM1 immunofluorescence, tissue clearing and light-sheet confocal microscopy, shown in **(C)** dorsal and **(D)** ventral views. Ventricles are heavily vascularized with strongly PECAM1-immunoreactive vessels, in particular the coronary arteries (arrows). **(E)** One confocal section from the same heart showing coronary arteries in section at the top of the ventricles and in the ventricular sulcus near the apex, arrowheads. **(F)** PRPH immunofluorescence at 6.5-7 pcw in a whole heart (EE3383). Innervation progresses along the great arteries, circumnavigates the atrioventricular junctions and coverges at the level of the atrioventricular node at the base of the right atrium before extending distally along the ventricles. **(G)** FABP4 is strongly immunofluorescent in the developing lymphatic vessels of the heart and, in this 7 pcw sample (F0004), the pericardium. (**H**) MYH6 immunofluorescence strongly labels the atria in this dorsal view of sample F0010 at 8.0 pcw. (**I**) The same heart was simultaneously stained with an antibody against POSTN, a mesenchymal marker strongly expressed at the atrioventricular valve hinge points (arrows) above the *annulus fibrosus* among other tissues, including the tracheal rings (t). Bar = 500 µm in (**A**-**H)**; in (**I**) the same bar length is 275 µm.

Supplementary Movies

Movie 1.

Heart F0495, 9.1 pcw male; immunofluorescence of PECAM1 (pink, vascular endothelium) and PRPH (yellow, peripheral nerves) followed by modified iDISCO tissue clearing. Reconstruction of signal across 1288 optical sections acquired on a Miltenyi Blaze light-sheet confocal microscope using Imaris software.

Movie 2.

Heart F0514, 9.1 pcw female; immunofluorescence of PECAM1 (pink, vascular endothelium) and PRPH (yellow, peripheral nerves) followed by modified iDISCO tissue clearing. Reconstruction of signal across 1332 optical sections acquired on a Miltenyi Blaze light-sheet confocal microscope using Imaris software.

Movie 3.

Heart F0570, 10.1 pcw female; immunofluorescence of PECAM1 (pink, vascular endothelium) and PRPH (yellow, peripheral nerves) followed by modified iDISCO tissue clearing. Reconstruction of signal across 1340 optical sections acquired on a Miltenyi Blaze light-sheet confocal microscope using Imaris software.

Movie 4.

Heart F0570, 10.1 pcw female; immunofluorescence of PECAM1 (pink, vascular endothelium) and PRPH (yellow, peripheral nerves) followed by modified iDISCO tissue clearing. Segmentation by clipping planes in Imaris software to show PECAM1-stained cusps of aortic valve and departure of left and right coronary arteries.

Supplementary table legends

Supplementary Materials and Methods

Summary of samples, stages, unique identifiers when available, analyses conducted and information about transcriptomics per sample in the summary tab. Photos of whole hearts, when available, are linked from the HuDeCA ID to the gross anatomy tab.

Supplementary Table 1

Data pertaining to 49,227 nuclear transcriptomes from three first-trimester human hearts. **First tab**: Index with hyperlinks to other tabs in table. Within all other tabs, a hyperlinked “Index” button will return to this tab. **Second tab**: Cluster names and cell statistics, broken down by source material. **Third tab**: All genes differentially expressed per cluster, relative to all others. **Fourth tab**: as previous but featuring top 20 differentially expressed genes per cluster, relative to all others. **Subsequent tabs**: All genes differentially expressed per class or cluster, relative

to all other classes or to other cluster(s) in same tab. Abbreviations: *p_val*, p-value; *avg_log2FC*, average log(2) fold-change; *pct.1*, percentage of cluster cells where the gene is detected; *pct.2*, percentage of all other cells where the gene is detected; *p_val_adj*, p-values adjusted for multiple hypothesis testing (based on a Bonferroni correction); *Cluster.no*, local snRNAseq cluster number; *gene*, official human gene name.

Supplementary Table 2

Marker genes and statistics of clusters from integrated snRNA-seq and scRNA-seq data, referring to Figure 2.

Supplemental Table 3

Marker genes and statistics of clusters from individual spatial transcriptomic analyses of A90, A91, A17, A18, A19 and A20 heart section data.

Supplemental Table 4

Marker genes and statistics of spatial transcriptomic clusters integrated from section A90, A91, A17, A18, A19 and A20 data.

Supplemental Table 5

Average proportions of each cell type from integrated snRNAseq data within each spatial transcriptomic cluster after deconvolution, per section.

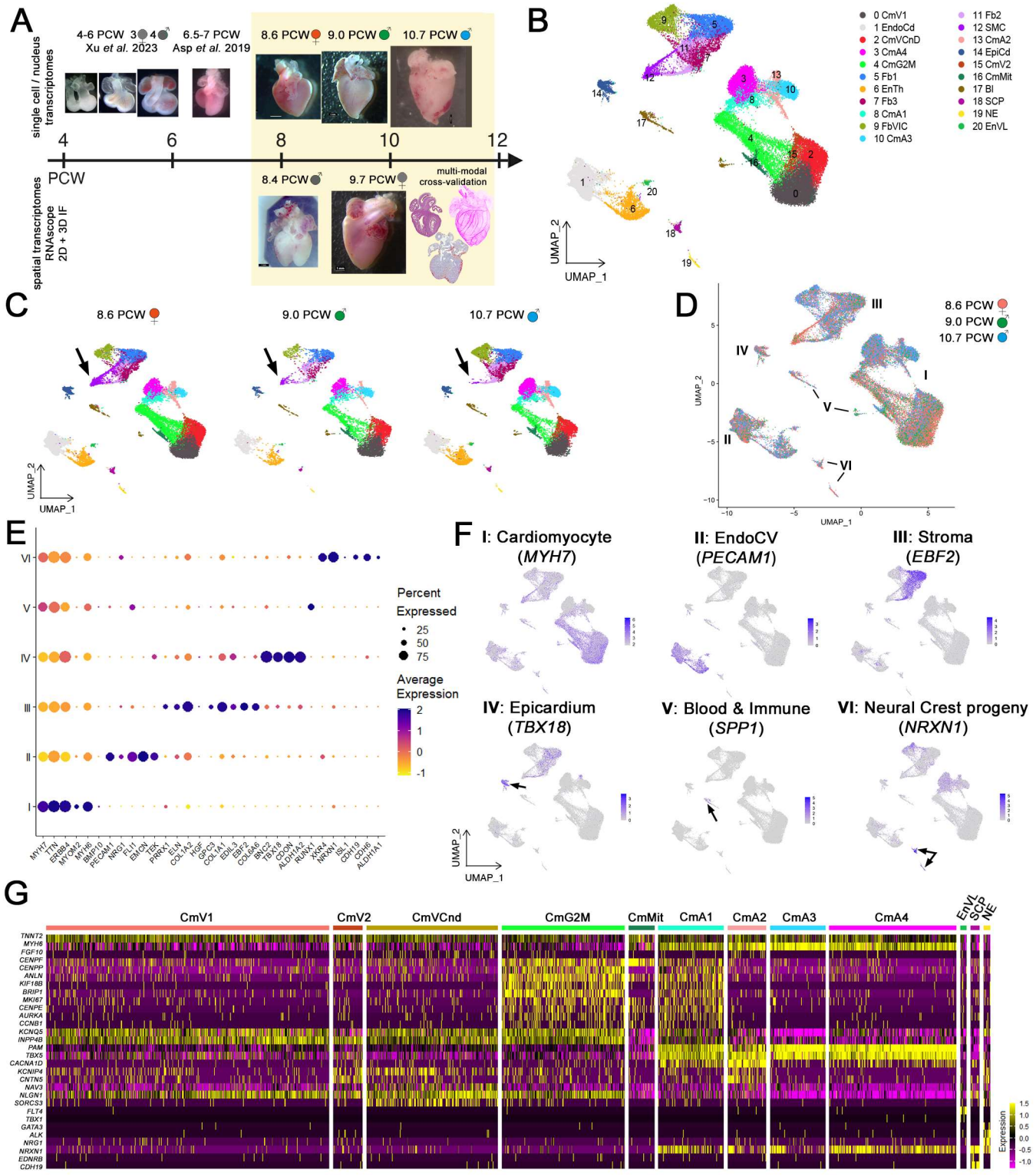
A
4-6 PCW
Xu et al. 2023
6.5-7 PCW
Asp et al. 2019
8.6 PCW
9.0 PCW
10.7 PCW
single cell / nucleus transcriptomes
4
6
8
10
12
PCW
8.4 PCW
9.7 PCW
multi-modal cross-validation
spatial transcriptomes
RNAscope
2D + 3D IF
B
0 CmV1
1 EndoCd
2 CmVCnD
3 CmA4
4 CmG2M
5 Fb1
6 EnTh
7 Fb3
8 CmA1
9 FbVIC
10 CmA3
11 Fb2
12 SMC
13 CmA2
14 EpiCd
15 CmV2
16 CmMit
17 BI
18 SCP
19 NE
20 EnVL
UMAP_2
UMAP_1
C
8.6 PCW
9.0 PCW
10.7 PCW
UMAP_2
UMAP_1
D
8.6 PCW
9.0 PCW
10.7 PCW
I
II
III
IV
V
VI
UMAP_2
UMAP_1
E
VI
V
IV
III
II
I
Percent Expressed
25
50
75
Average Expression
2
1
0
-1
MYH7 TTN ERBB4 MYOM2 MYH6 BMP10 PECAM1 NRG1 FLI1 EMCN TEK PRRX1 ELN COL1A2 HGF GPC3 COL1A1 EDIL3 EBF2 COL6A6 BNC2 TBX18 CDON ALDH1A2 RUNX1 XKR4 NRXN1 ISL1 CDH19 CDH6 ALDH1A1
F
I: Cardiomyocyte (MYH7)
II: EndoCV (PECAM1)
III: Stroma (EBF2)
IV: Epicardium (TBX18)
V: Blood & Immune (SPP1)
VI: Neural Crest progeny (NRXN1)
UMAP_2
UMAP_1
G
CmV1
CmV2
CmVCnd
CmG2M
CmMit
CmA1
CmA2
CmA3
CmA4
EnVL
SCP
NE
TNNT2
MYH6
FGF10
CENPF
CENPP
ANLN
KIF18B
BRIP1
MKI67
CENPE
AURKA
CCNB1
KCNQ5
INPP4B
PAM
TBX5
CACNA1D
KCNIP4
CNTN5
NAV3
NLGN1
SORCS3
FLT4
TBX1
GATA3
ALK
NRG1
NRXN1
EDNRB
CDH19
Expression
1.5
1.0
0.5
0.0
-0.5
-1.0

Figure 1

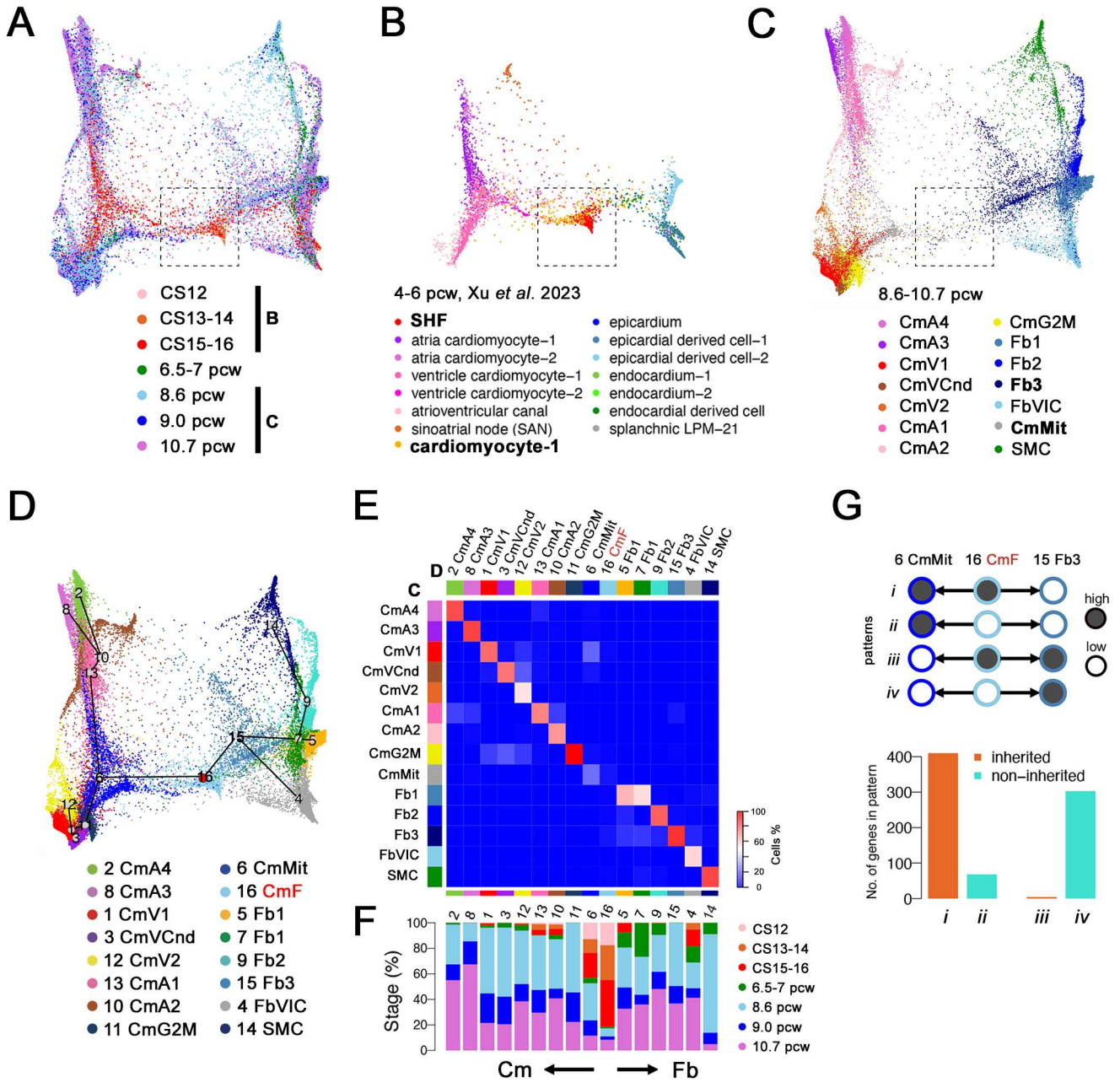
A
CS12
CS13-14
CS15-16
6.5-7 pcw
8.6 pcw
9.0 pcw
10.7 pcw
B
C
B
4-6 pcw, Xu et al. 2023
SHF
atria cardiomyocyte–1
atria cardiomyocyte–2
ventricle cardiomyocyte–1
ventricle cardiomyocyte–2
atrioventricular canal
sinoatrial node (SAN)
cardiomyocyte-1
epicardium
epicardial derived cell–1
epicardial derived cell–2
endocardium–1
endocardium–2
endocardial derived cell
splanchnic LPM–21
C
8.6-10.7 pcw
CmA4
CmA3
CmV1
CmVCnd
CmV2
CmA1
CmA2
CmG2M
Fb1
Fb2
Fb3
FbVIC
CmMit
SMC
D
2 CmA4
8 CmA3
1 CmV1
3 CmVCnd
12 CmV2
13 CmA1
10 CmA2
11 CmG2M
6 CmMit
16 CmF
5 Fb1
7 Fb1
9 Fb2
15 Fb3
4 FbVIC
14 SMC
E
D
C
Cells %
F
Stage (%)
Cm
Fb
CS12
CS13-14
CS15-16
6.5-7 pcw
8.6 pcw
9.0 pcw
10.7 pcw
G
6 CmMit
16 CmF
15 Fb3
patterns
high
low
inherited
non–inherited
No. of genes in pattern

Figure 2

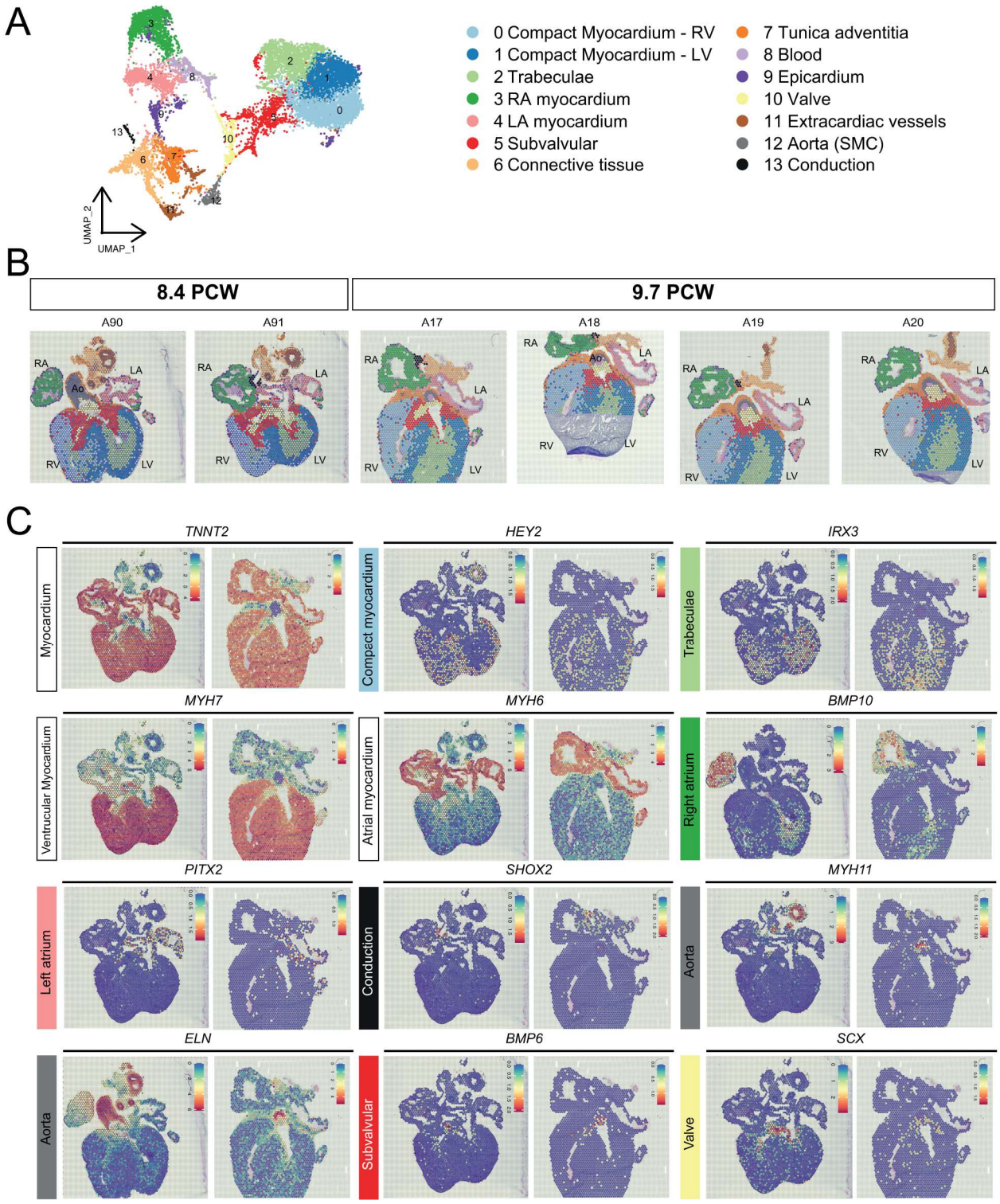
A
0 Compact Myocardium - RV
1 Compact Myocardium - LV
2 Trabeculae
3 RA myocardium
4 LA myocardium
5 Subvalvular
6 Connective tissue
7 Tunica adventitia
8 Blood
9 Epicardium
10 Valve
11 Extracardiac vessels
12 Aorta (SMC)
13 Conduction
UMAP_1
UMAP_2
B
8.4 PCW
9.7 PCW
A90
A91
A17
A18
A19
A20
RA
LA
Ao
RV
LV
C
TNNT2
HEY2
IRX3
MYH7
MYH6
BMP10
PITX2
SHOX2
MYH11
ELN
BMP6
SCX
Myocardium
Compact myocardium
Trabeculae
Ventrucular Myocardium
Atrial myocardium
Right atrium
Left atrium
Conduction
Aorta
Aorta
Subvalvular
Valve

Figure 3

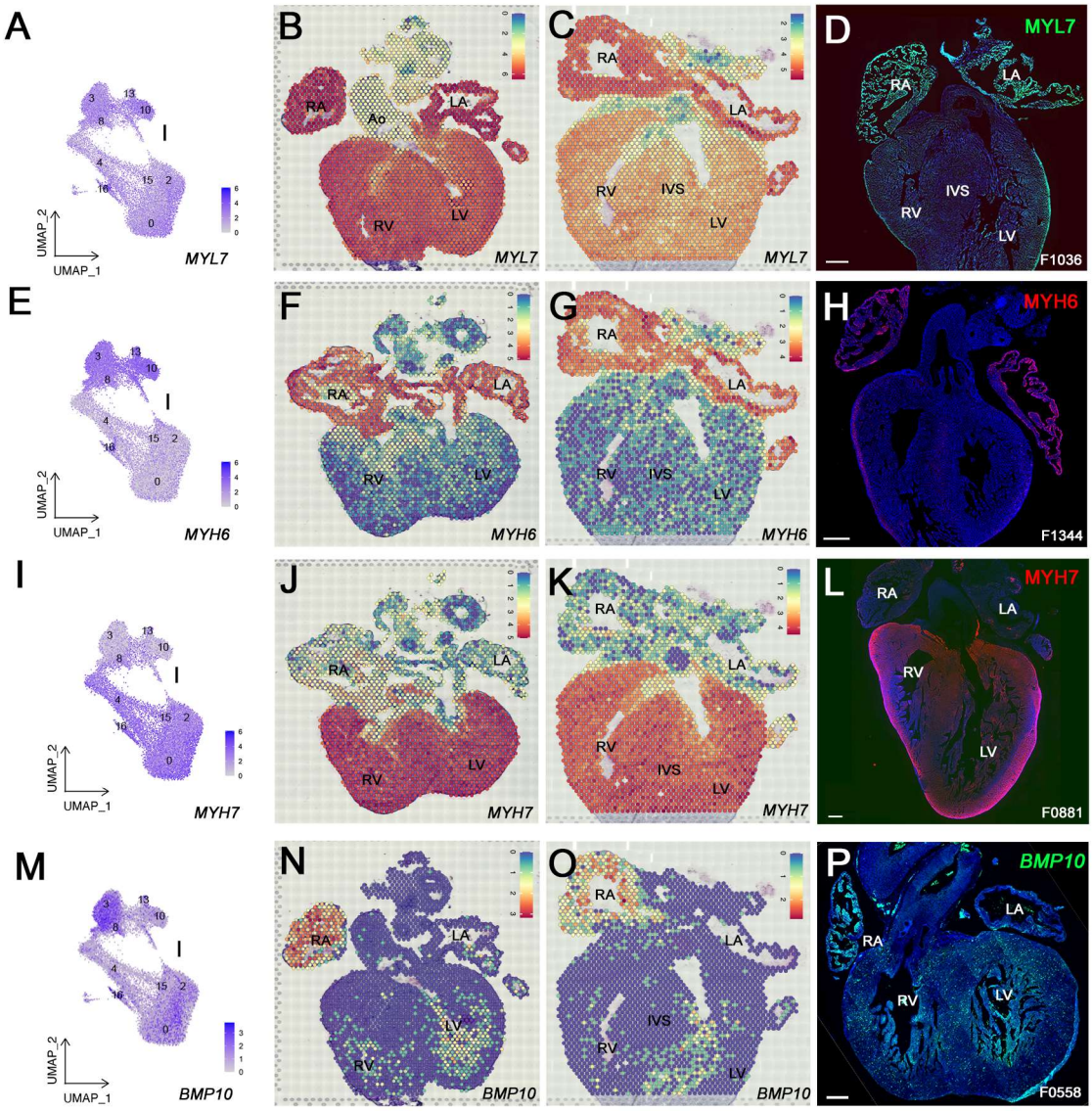
A
B
C
D
MYL7
RA
LA
Ao
RV
LV
IVS
F1036
E
F
G
H
MYH6
F1344
I
J
K
L
MYH7
F0881
M
N
O
P
BMP10
F0558
UMAP_1
UMAP_2

Figure 4

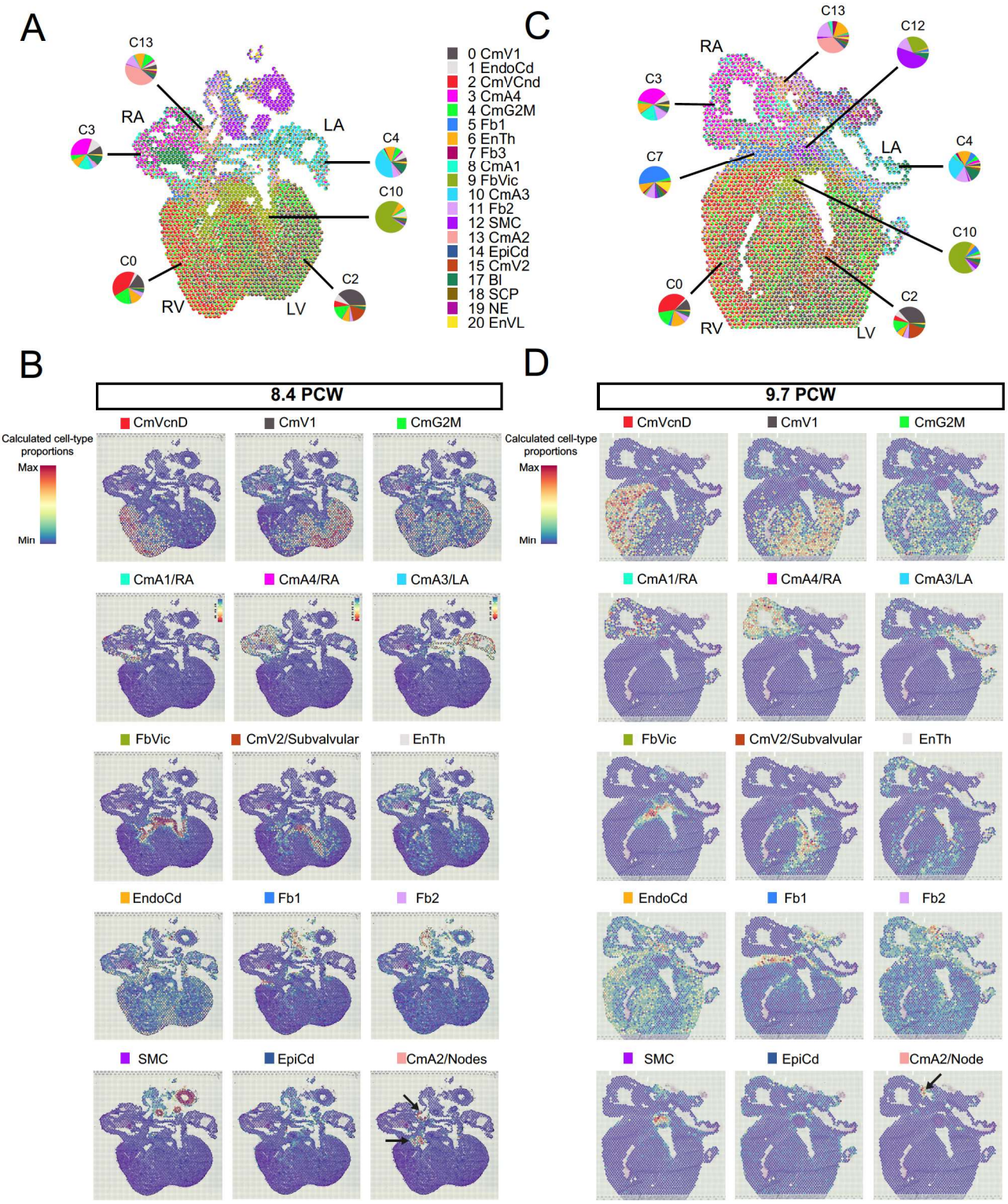
A
C13
RA
C3
LA
C4
C10
C0
C2
RV
LV
0 CmV1
1 EndoCd
2 CmVCnd
3 CmA4
4 CmG2M
5 Fb1
6 EnTh
7 Fb3
8 CmA1
9 FbVic
10 CmA3
11 Fb2
12 SMC
13 CmA2
14 EpiCd
15 CmV2
17 Bl
18 SCP
19 NE
20 EnVL
C
C13
C12
RA
C3
C7
LA
C4
C10
C0
C2
RV
LV
B
8.4 PCW
Calculated cell-type proportions
Max
Min
CmVcnD
CmV1
CmG2M
CmA1/RA
CmA4/RA
CmA3/LA
FbVic
CmV2/Subvalvular
EnTh
EndoCd
Fb1
Fb2
SMC
EpiCd
CmA2/Nodes
D
9.7 PCW
Calculated cell-type proportions
Max
Min
CmVcnD
CmV1
CmG2M
CmA1/RA
CmA4/RA
CmA3/LA
FbVic
CmV2/Subvalvular
EnTh
EndoCd
Fb1
Fb2
SMC
EpiCd
CmA2/Node

Figure 5

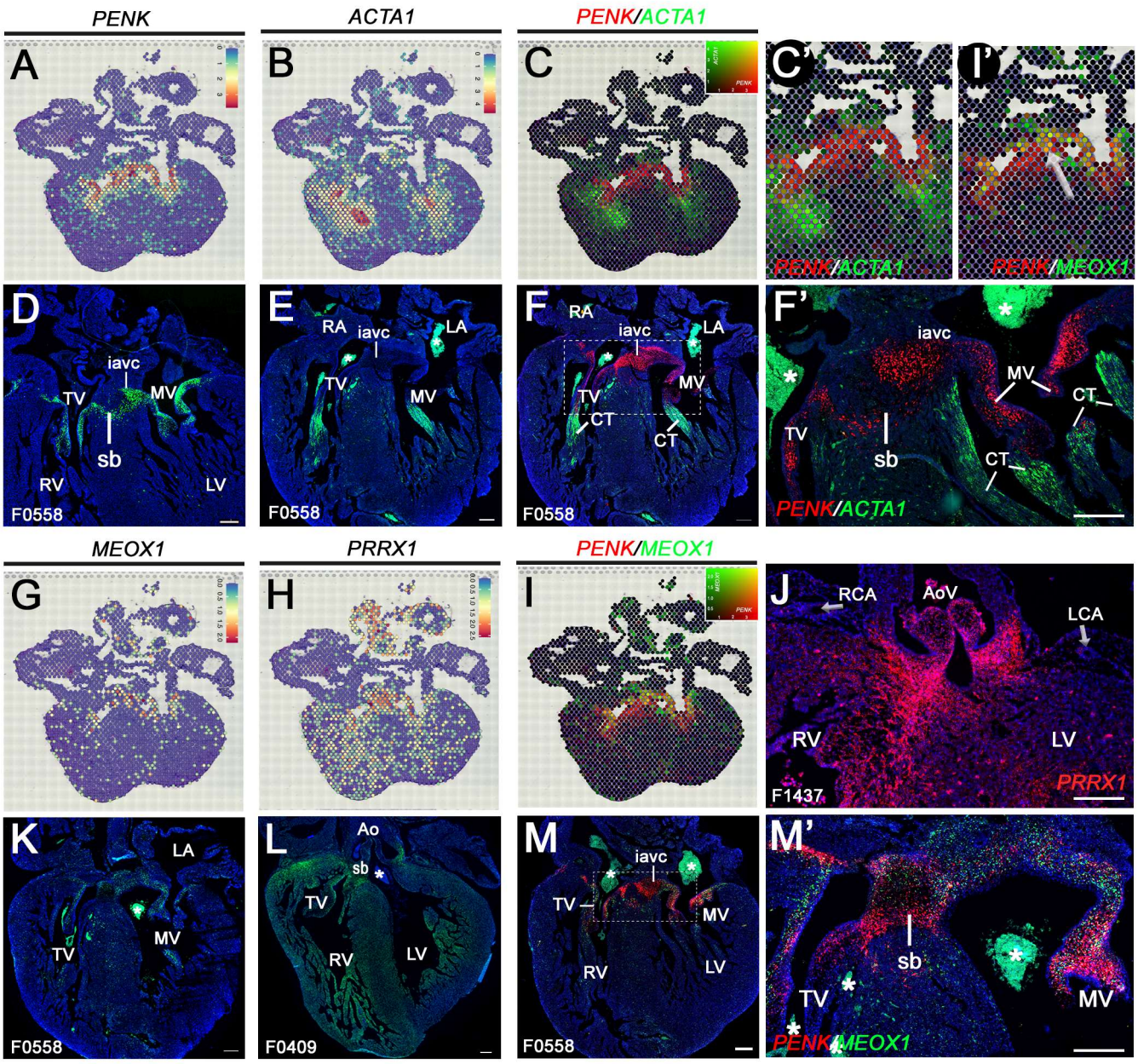
PENK
ACTA1
PENK/ACTA1
A
B
C
C'
I'
PENK/ACTA1
PENK/MEOX1
D
E
F
F'
RA
iavc
LA
TV
MV
CT
sb
RV
LV
F0558
PENK/ACTA1
MEOX1
PRRX1
PENK/MEOX1
G
H
I
J
RCA
AoV
LCA
F1437
PRRX1
K
L
M
M'
Ao
F0409
PENK/MEOX1

Figure 6

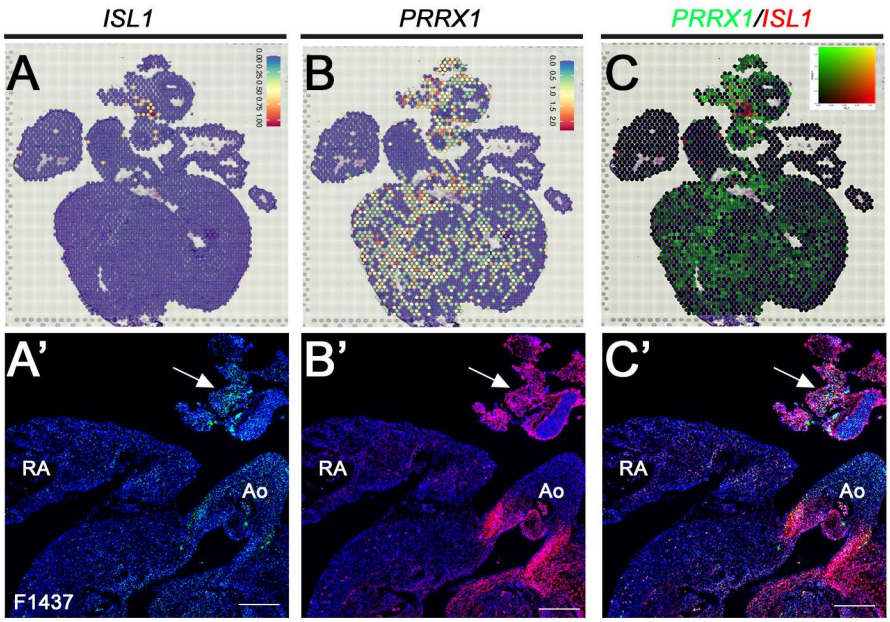

Figure 7

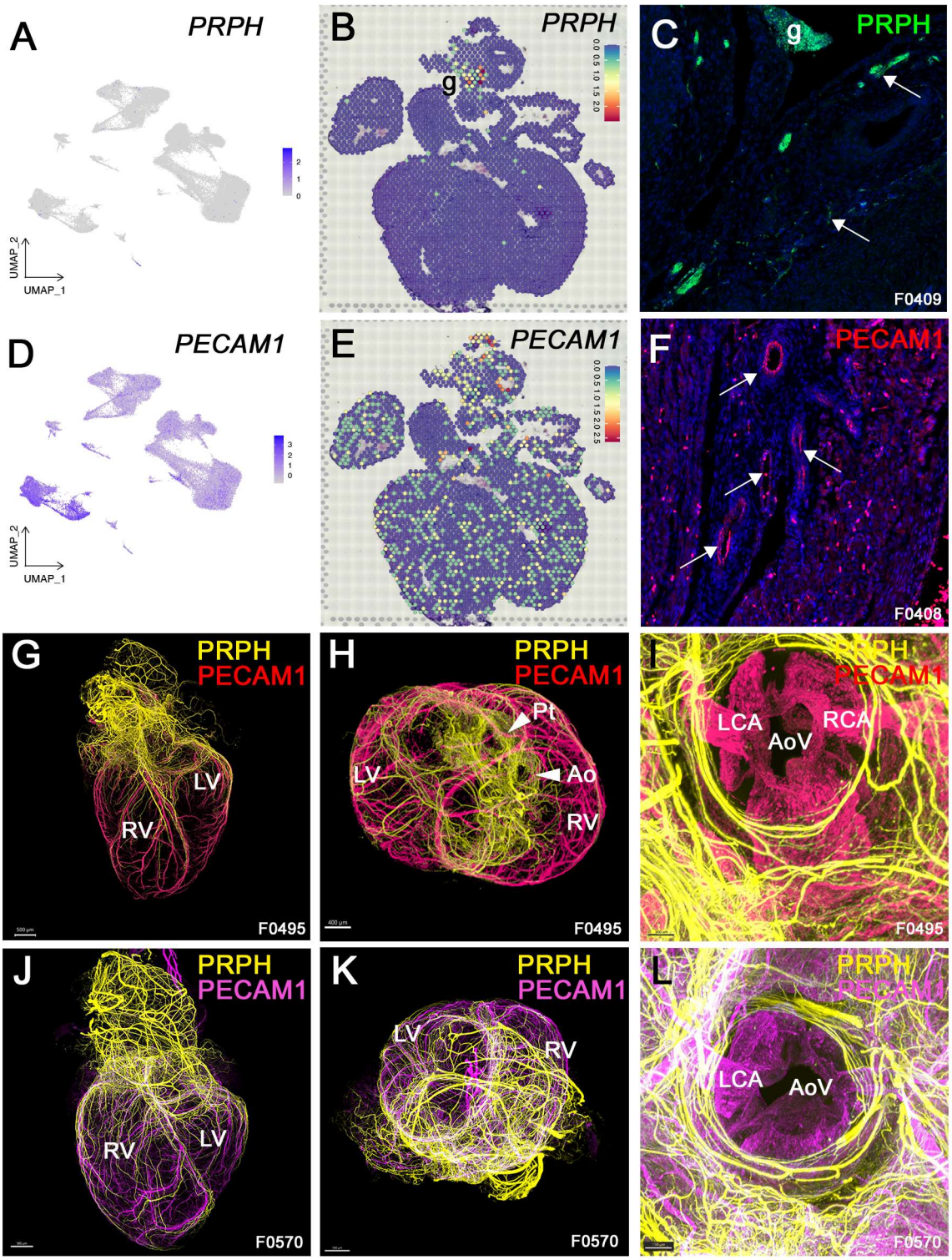
A
PRPH
UMAP_2
UMAP_1
B
PRPH
g
C
g
PRPH
F0409
D
PECAM1
UMAP_2
UMAP_1
E
PECAM1
F
PECAM1
F0408
G
PRPH
PECAM1
LV
RV
F0495
H
PRPH
PECAM1
Pt
LV
Ao
RV
F0495
I
PRPH
PECAM1
LCA
AoV
RCA
F0495
J
PRPH
PECAM1
RV
LV
F0570
K
PRPH
PECAM1
LV
RV
F0570
L
PRPH
PECAM1
LCA
AoV
F0570

Figure 8