STEM CELLS AND REGENERATION

RESEARCH ARTICLE

Dissecting mechanisms of chamber-specific cardiac differentiation and its perturbation following retinoic acid exposure

David M. Gonzalez[1,2,3,4,*], Nadine Schrode[5,*], Tasneem A. M. Ebrahim[1,2,3,4], Nicolas Broguiere[6], Giuliana Rossi[6], Lika Drakhlis[7], Robert Zweigerdt[7], Matthias P. Lutolf[6,7], Kristin G. Beaumont[5,8,‡], Robert Sebra[3,5,9,‡] and Nicole C. Dubois[1,2,3,4,‡]

ABSTRACT

The specification of distinct cardiac lineages occurs before chamber formation and acquisition of bona fide atrial or ventricular identity. However, the mechanisms underlying these early specification events remain poorly understood. Here, we performed single cell analysis at the murine cardiac crescent, primitive heart tube and heart tube stages to uncover the transcriptional mechanisms underlying formation of atrial and ventricular cells. We find that progression towards differentiated cardiomyocytes occurs primarily based on heart field progenitor identity, and that progenitors contribute to ventricular or atrial identity through distinct differentiation mechanisms. We identify new candidate markers that define such differentiation processes and examine their expression dynamics using computational lineage trajectory methods. We further show that exposure to exogenous retinoic acid causes defects in ventricular chamber size, dysregulation in FGF signaling and a shunt in differentiation towards orthogonal lineages. Retinoic acid also causes defects in cell-cycle exit resulting in formation of hypomorphic ventricles. Collectively, our data identify, at a single cell level, distinct lineage trajectories during cardiac specification and differentiation, and the precise effects of manipulating cardiac progenitor patterning via retinoic acid signaling.

KEY WORDS: Cardiac development, Cardiac organoid, Cardiac progenitor, Retinoic acid, Single cell RNA-sequencing, Mouse

INTRODUCTION

The process of early heart development is complex, involving multiple progenitors that contribute differentiated progeny to distinct regions of the heart (Kelly et al., 2014; Meilhac and Buckingham, 2018; Meilhac et al., 2014; Yao et al., 2021; Zhang et al., 2021). Shortly after gastrulation, *Mesp1+* cells from the lateral plate mesoderm (LPM) form two distinct progenitor populations termed the first heart field (FHF) and the second heart field (SHF) (Chabab et al., 2016; Chiapparo et al., 2016; Devine et al., 2014; Lescroart et al., 2014; Saga et al., 1996). Together, these form the cardiac crescent (CC), which later balloons anteriorly into the primitive heart tube (PHT) and undergoes looping (Meilhac et al., 2014). The *Isl1+* SHF forms posteriorly to the FHF and persists as a progenitor population, continuing to proliferate and contribute cells to the developing heart (Cai et al., 2003; Laugwitz et al., 2005). Derivatives from the FHF and SHF are distributed asymmetrically within the four chambers of the heart (Cai et al., 2003; Liang et al., 2013; Meilhac et al., 2014; Mjaatvedt et al., 2001; Waldo et al., 2001; Zaffran et al., 2004) and dysregulation in these progenitors is thought to underlie a number of congenital heart defects (CHDs) (Bruneau, 2008, 2020; Kloesel et al., 2016; Pierpont et al., 2018).

Defects such as atrial/ventricular chamber size imbalances are not adequately described by the existing FHF/SHF model, raising questions on whether specific subpopulations of progenitors are responsible for these phenotypes and about the precise mechanisms governing atrial and ventricular specification. Retinoic acid (RA) acts as a morphogen driving atrial differentiation (Bernheim and Meilhac, 2020; Niederreither et al., 2001; Perl and Waxman, 2020; Stefanovic and Zaffran, 2017). However, the transcriptional mechanisms downstream of this signaling pathway remain elusive. A number of efforts aimed at understanding the mechanisms of atrial/ventricular lineage segregation have performed bulk and single cell RNA-sequencing (scRNA-seq) as early as embryonic day (E) 9 (DeLaughter et al., 2016; Li et al., 2016). Although informative, these studies did not capture the key transcriptional processes that occur before chamber morphogenesis. More recent studies have profiled the transcriptome of the developing heart shortly after gastrulation (Ivanovitch et al., 2021; Lescroart et al., 2018; Tyser et al., 2021) or at the onset of chamber morphogenesis (Mantri et al., 2021; de Soysa et al., 2019; Xiong et al., 2019; Zhang et al., 2021). These studies highlight the transcriptional heterogeneity of cardiac progenitors and have identified markers that segregate distinct cell populations in a spatio-temporal manner. However, the specification and differentiation of atrial or ventricular cells, particularly before acquisition of chamber-specific marker expression, remains to be explored further. We and others have recently expanded on work demonstrating that atrial/ventricular lineage segregation occurs early during mammalian development, perhaps before FHF/SHF formation (Bardot et al., 2017a; Chabab et al., 2016; Garcia-Martinez and Schoenwolf, 1993; Keegan et al., 2004; Lescroart et al., 2014; Yutzey and Bader, 1995). We have shown that transient expression of *Foxa2* in the primitive streak labels a population of cardiac mesoderm that migrates at the leading

[1]Department of Cell, Developmental, and Regenerative Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [2]Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [3]Black Family Stem Cell Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [4]Cardiovascular Research Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [5]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. [6]Laboratory of Stem Cell Bioengineering, School of Life Sciences, EPFL, Lausanne CH-1015, Switzerland. [7]Roche Institute for Translational Bioengineering, Roche Pharma Research and Early Development, Basel 4052, Switzerland. [8]Leibniz Research Laboratories for Biotechnology and Artificial Organs (LEBAO), Department of Cardiothoracic, Transplantation and Vascular Surgery (HTTG), REBIRTH–Research Center for Translational Regenerative Medicine, Hannover Medical School, Hannover, Germany. [9]Sema4, a Mount Sinai venture, Stamford, CT 06902, USA.
*These authors contributed equally to this work

‡Authors for correspondence (nicole.dubois@mssm.edu; robert.sebra@mssm.edu; kristin.beaumont@mssm.edu)

K.G.B., 0000-0003-3075-9977; R.S., 0000-0001-9267-2426; N.C.D., 0000-0002-8467-3004

edge of the LPM. *Foxa2* lineage-tracing experiments show that these early migrating cells contribute to the majority of the ventricular, but not atrial, myocardium (Bardot et al., 2017a).

To interrogate the molecular identity of atrial and ventricular progenitors, as well as lineage relationships between progenitors and their differentiated progeny, we performed scRNA-seq of *Foxa2Cre;mTmG* embryos at the CC (E8.25), PHT (E8.75) and heart tube (HT) (E9.25) stages. RNA velocity revealed relationships between distinct progenitors and atrial/ventricular cells and identified putative driver genes of differentiation that segregated over inferred transcriptional time. Through differential expression analysis we identified a number of candidate markers that correlate with atrial or ventricular differentiation over time. Lastly, we found that exogenous RA negatively impacts ventricular development by preventing differentiation from separate progenitors towards a ventricular fate. This led to defects in ventricular size and shifts in differentiation trajectories towards other lineages. In summary, we provide a transcriptional atlas across cardiac progenitor specification and differentiation at single cell level, illustrating deployment of distinct mechanisms along progenitor differentiation trajectories that are disrupted when cardiac development is perturbed by exogenous RA.

RESULTS
Generating a single cell transcriptomic atlas of early cardiac development

We sub-dissected cardiac regions as well as the surrounding areas comprising the gut tube and pharyngeal structures of *Foxa2Cre; mTmG* embryos, which allowed us to identify ventricular-fated progenitors at early stages (Bardot et al., 2017a) (Fig. 1A and Fig. S1A). Three age-matched embryos from the CC, PHT and HT stages each were pooled and a minimum of 10,000 cells sequenced (CC: 14,355; PHT: 13,036; HT: 21,318). We performed clustering and differential gene expression (DGE) analysis to identify multiple mesoderm and cardiac cell types, as well as cells from ectoderm and endodermal lineages (Fig. S1B). Importantly, we found that we can identify *Foxa2* lineage-traced cells through *EGFP* expression without cell-sorting, which might have resulted in loss of rare cell types. We merged the data from the three time points to create a time-resolved atlas of the developing heart and surrounding tissues, identifying 27 clusters that separated broadly according to germ layer identity (Fig. S2A). DGE analysis determined signature genes defining populations of differentiated cardiomyocytes and associated progenitors, pharyngeal endoderm and mesoderm, as well as various definitive ectoderm and endoderm cell types among others (Fig. S2B-D).

To focus on the cardiovascular lineages, we subset and re-clustered cardiomyocytes, cardiac progenitors and mesoderm cells along with neural crest, endocardial, endothelial and epicardial cells at each stage (Fig. S3). We merged these data to resolve 27 clusters including heart field progenitors, cardiomyocytes at various stages of differentiation, as well as other supporting cell types of the heart, which were identified through known markers and candidate genes identified by DGE (Fig. 1B-D). Progenitors and differentiated cells segregated in part by cell cycle state, with *Pdgfra+* mesoderm populations more commonly found in S phase and differentiated *Nkx2-5+* progeny more commonly found in G2 M/G1 phase (Fig. 1E). This was also reflected by the contribution to each cluster from the different samples, with differentiated cardiomyocytes [cluster (C) 4, C2, C17] derived primarily from the PHT and HT stages, whereas the LPM (C26) primarily derived from the CC stage (Fig. 1F). *Pdgfra+/Isl1+* SHF progenitors segregated into *Tbx1+*

anterior and *Aldh1a2+/Hoxb1+/Foxf1+* posterior domains (Hoffmann et al., 2014; Rana et al., 2014; Stefanovic et al., 2020) (Fig. 1C,D). Intriguingly, although we did identify an FHF population expressing *Nkx2-5/Hcn4/Tbx5* when examining the CC stage alone (Figs S3 and S4A), when merged with later time points these cells did not form a discrete population. Instead FHF cells clustered with multiple differentiating cardiomyocyte populations (C7, C12, C14), underscoring the identity of the FHF as a population of cells that most closely resembles differentiating cardiomyocytes. Differentiated cardiomyocytes segregated into several subpopulations including *Kcna5+* atrial and *Myl2+* ventricular cells of left/right identity (Fig. 1C; Fig. S4B,C). Atrial cells did not segregate discretely into left/right populations, though did show markers of left/right identity such as *Pitx2* (Fig. S4D). We also identified populations corresponding to the *Rspo3+* atrioventricular canal (AVC) and *Hcn4+/Shox2+* sinus venosus (SV), which clustered near atrial cardiomyocytes.

In addition to cardiomyocytes, we identified *Sox2+* neural crest cells, as well as *Wt1+* epicardial cells (Fig. 1C,D). Localized distantly to the cardiomyocytes we identified several *Pecam1+* endothelial populations (C9, C13, C16), which clustered closely with *Nfatc1+* endocardial cells (C23). Both populations are proximal to *Kdr+* (C3) and *Etv2+* (C24) clusters that likely represent hematoendothelial cells. Curiously, we also identified a population of cardiomyocytes (C0) and mesoderm cells (C15) which clustered separately from the other heart field progenitors/ cardiomyocytes and associated more closely with endothelial cells. We were unable to determine the exact source of this observation as these cells expressed similar markers to other mesoderm and cardiomyocytes. Positioning of these clusters could suggest that they represent a distinct subset of mesoderm/ cardiomyocytes, or alternatively it could be related to comparatively lower transcript expression in C0/15 and the adjacent endothelial populations.

We were intrigued to find multiple clusters of differentiated *Tnnt2+* cardiomyocytes (C0, C2, C4, C5, C7, C12, C17), only some of which had clear atrial/ventricular or left/right identity. To determine if these represented functional subtypes we performed DGE and GO term analysis to identify differentially regulated processes in each cluster (Fig. 1G; Fig. S4C). We found that processes involved in 'mitochondrial ATP synthesis coupled electron transport' or 'striated muscle hypertrophy' segregated myocardial cells into two broad categories of more or less mature cells. Clusters that did not show enrichment in energetic processes were instead characterized by developmental processes of migration and/or differentiation. Others, such as those in C0, were enriched for terms involved in cardiac muscle contraction and ribosome biogenesis, but few other more advanced processes, leading us to conclude that these represented early myocardial precursors. Cells of the SV (C12) were enriched for genes involved in 'SA nodal action potential', suggesting that they contained a subpopulation of pacemaker cells. When subclustered, the SV cells broadly segregate into an *Nkx2-5* population expressing sinoatrial nodal markers such as *Smoc2*, *Hcn4* and *Shox2* and an *Nkx2-5+* population that instead expressed markers of atrial cardiomyocytes (Fig. S5A). We found processes involved in aortic and pulmonary valve development to be enriched in one of the *Tnnt2+* clusters (C5). Given the proximity to *Isl1/Tbx1*-expressing anterior second heart field (aSHF), we hypothesized that this cluster might represent aSHF-derived cells of both the outflow tract (OFT) and right ventricle (RV). We subclustered C5 and found that the cells separated into three distinct populations consisting of two populations of *Myl2+* cells and a
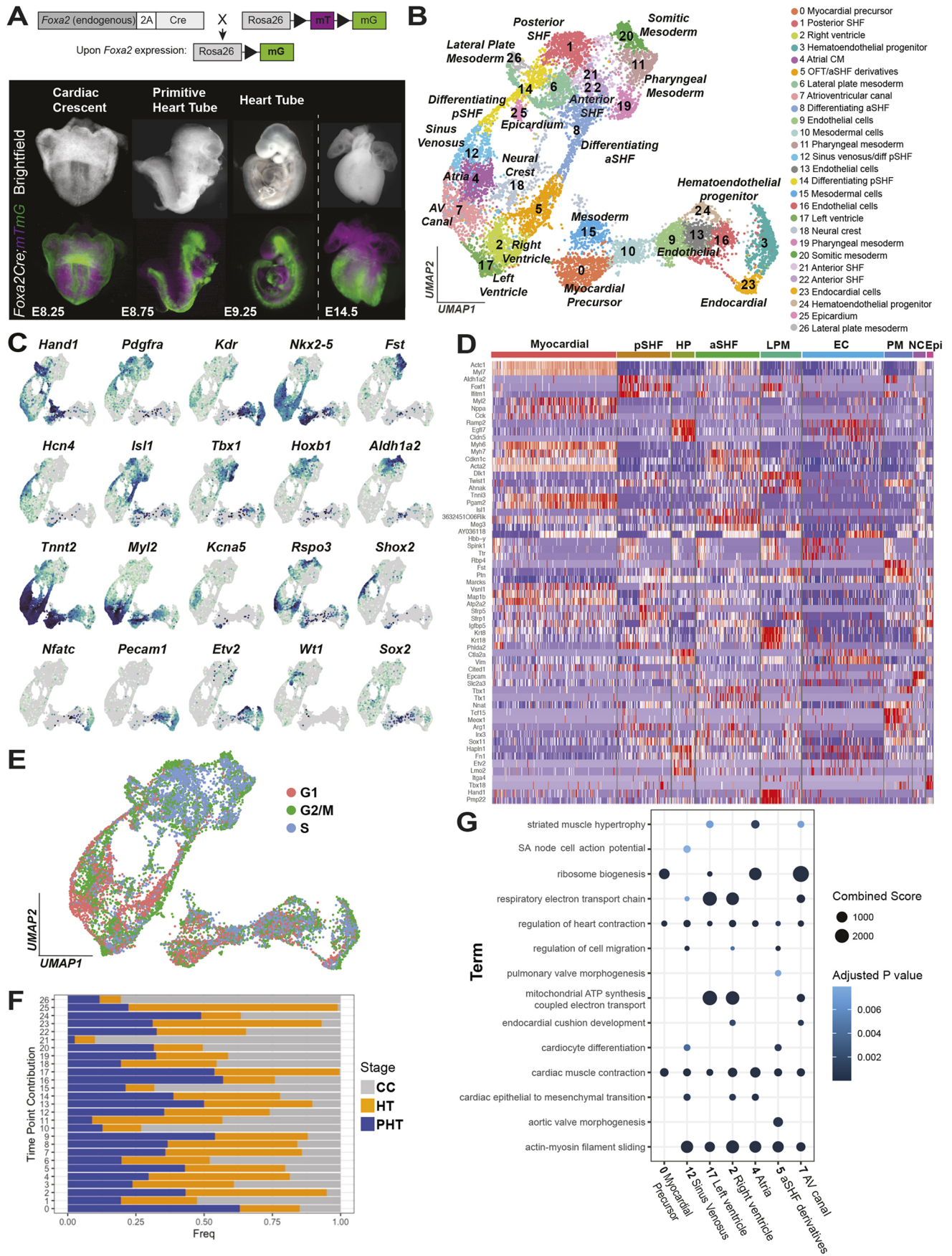
2

Fig. 1. See next page for legend.

Fig. 1. Establishing a transcriptional atlas of early cardiogenesis. (A) Top: schematic showing mating strategy used to generate *Foxa2Cre;R26mTmG* mice. Bottom: lineage-tracing in *Foxa2Cre;R26mTmG* embryos (representative of stages sequenced) and dissected heart shows ventricular-specific labeling of cardiac tissues. (B) UMAP clustering of cardiovascular populations. (C) FeaturePlots of key progenitor, cardiomyocyte, epicardium, neural crest and endothelial cell markers used together with DGE to annotate populations presented in B. (D) Heatmap demonstrating top three most differentially expressed genes for broad categories of cell types. (E) Cell cycle phase scoring projected onto UMAP embedding demonstrating preponderance of cardiac progenitors in S phase, with differentiating cell types primarily in G1 and G2/M phase. (F) Contribution from each stage to individual clusters (gray, CC; blue, PHT; orange, HT). Frequency is calculated as the number of cells in a given cluster from one stage relative to the total number of cells in the sample. (G) GO analysis of differentially regulated processes. Color of dot represents adjusted *P*-value score. Size of dot represents Combined Score given by enrichr. aSHF, anterior second heart field; CM, cardiomyocyte; EC, endothelial/endocardial cell; Epi, epicardium; HP, hematoendothelial progenitor; LPM, lateral plate mesoderm; NC, neural crest; OFT, outflow tract; pSHF, posterior second heart field.

separate population enriched for *Crabp1*, *Rgs5*, *Flna* and *3632451O06Rik* (also known as *Armh4*), markers of the OFT (Fig. S5B). The clustering of developmental intermediates with divergent fates (such as the OFT and RV cardiomyocytes) suggests that, at early stages of differentiation, transcriptional similarity due to a shared origin may predominate over cell type-specific gene expression.

Use of transcriptomic atlas as a reference identifies previously uncovered heterogeneity of *in vitro* organoid models

We next tested the ability of our data to provide more granular detail on cell-type annotation from equivalent populations in other model systems. Recently, Rossi et al. reported the generation of gastruloids from mouse embryonic stem cells that mimic development of cardiac cells and other lineages (Rossi et al., 2021). Using scRNA-seq data from these gastruloids, we subset equivalent populations to those profiled in our data and performed supervised machine learning using a random forest classifier, using our embryo scRNA-seq data as a training dataset (Tan and Cahan, 2019). This allowed us to map the cell types present in cardiac organoids to their most equivalent cell type in the embryo (Fig. 2). We found that pharyngeal and somitic mesoderm, as well as hematoendothelial progenitors in the cardiac organoid, mapped with high accuracy to their equivalent *in vivo* populations. Cardiomyocytes within the organoids, which had previously been annotated as early and late cardiomyocytes, mapped primarily to aSHF and posterior second heart field (pSHF) cells, as well as the developing SV and AVC. Few, if any, cardiomyocytes mapped to differentiated atrial or ventricular cardiomyocytes, consistent with Rossi et al.'s description of the cardiac regions representing early stages of cardiac development. We applied a similar approach to data obtained from human pluripotent stem cell-derived heart forming organoids (HFOs) (Fig. 2) (Drakhlis et al., 2021). Mesenchymal cell types identified previously by Drakhlis et al. were sub-categorized between the LPM and a/pSHF population, indicating that HFOs recapitulate aspects of heart field progenitor formation. Furthermore, populations identified by Drakhlis et al. as cardiomyocytes mapped mostly to myocardial/OFT intermediates derived from the aSHF (C5), and showed little overlap with bona-fide atrial or ventricular cardiomyocytes (Drakhlis et al., 2021). This underscores that myocardial populations formed in current organoid models represent early stage myocardial intermediates rather than chamber-specific cardiomyocytes, as expected, and demonstrates the potential of our *in vivo* transcriptional atlas to provide additional detail on cellular heterogeneity and differentiation state of cells in organoid models.

Identification of lineage relationships and dynamically regulated genes along cardiac differentiation trajectories

Having identified a number of intermediary cell types in our data, we next aimed to define lineage relationships between cardiac progenitors and their progeny, as well as dynamic changes in genes that inform these relationships. Toward this end, we applied scVelo to estimate RNA velocity, which quantifies the relative abundance of spliced and unspliced transcripts for any given gene in each cell (Bergen et al., 2020). For each gene, the ratio of pre-mature (unspliced) and mature (spliced) mRNA counts is fitted to a model of splicing kinetics constituting a constant transcriptional state, where positive velocity indicates upregulation of a given gene (i.e. a higher abundance of unspliced mRNA than expected in a steady state). From this, we calculated a velocity vector representing the direction and speed of movement of individual cells, revealing several trajectories of differentiation (Fig. 3A). We observed two parallel differentiation streams which began from the pSHF and aSHF (C1 and C21/22, respectively) and terminated in differentiated cardiomyocytes. In addition, we observed diverging differentiation of the aSHF toward the pharyngeal mesoderm and cardiac lineages. Derivatives from the aSHF (e.g. RV and OFT) were clustered near populations of differentiated ventricular cells and *Sox2*-expressing neural crest cells. These observations are consistent with important interactions between the neural crest and OFT during septation of the aortic and pulmonary arteries (Kirby et al., 1983). Accordingly, RNA velocity also identified a stream of differentiation within the neural crest cells, the terminal position of which diverged and rested with the AVC. This association is consistent with the role of the neural crest in driving endocardial cushion development (Phillips et al., 2013). To understand the temporal ordering of cells along this trajectory, we calculated a latent time score (a measure of the internal clock of the cell based on transcriptional dynamics) and the velocity length and confidence, measurements of the speed of differentiation and congruency of vector fields, respectively (Fig. 3B-E). This captured the transition from cycling low velocity progenitors towards differentiating intermediates, which increased their velocity length as they express differentiated markers, indicating that heterogeneity observed in SHF progenitors is related to differentiation state (Figs 1B and 3D).

To identify the genes underlying the vector field and inferred lineages, we determined the genes with the highest differential velocity as these may represent putative drivers of cell identity or differentiation. We generated a ranked list of the top 100 dynamic genes for each cluster (Table S1). Furthermore, we subset each parallel differentiation trajectory initiated from the pSHF and aSHF and identified the top 300 most dynamically regulated genes across the differentiation stream (Fig. 3F,I). Not unexpectedly, about half of the top dynamically regulated genes were shared across the LPM/pSHF and aSHF differentiation stream; however, 136 genes were unique to each differentiation stream (Tables S2 and S3). In order to visualize the expression patterns of these rapidly changing genes across the pSHF and aSHF, we plotted the expression patterns of the top 300 across latent time, as well as a number of other candidates of interest (Fig. 3F-K; Fig. S6A,B). This allowed us to determine the ordered cascade of these genes across differentiation, and identify at
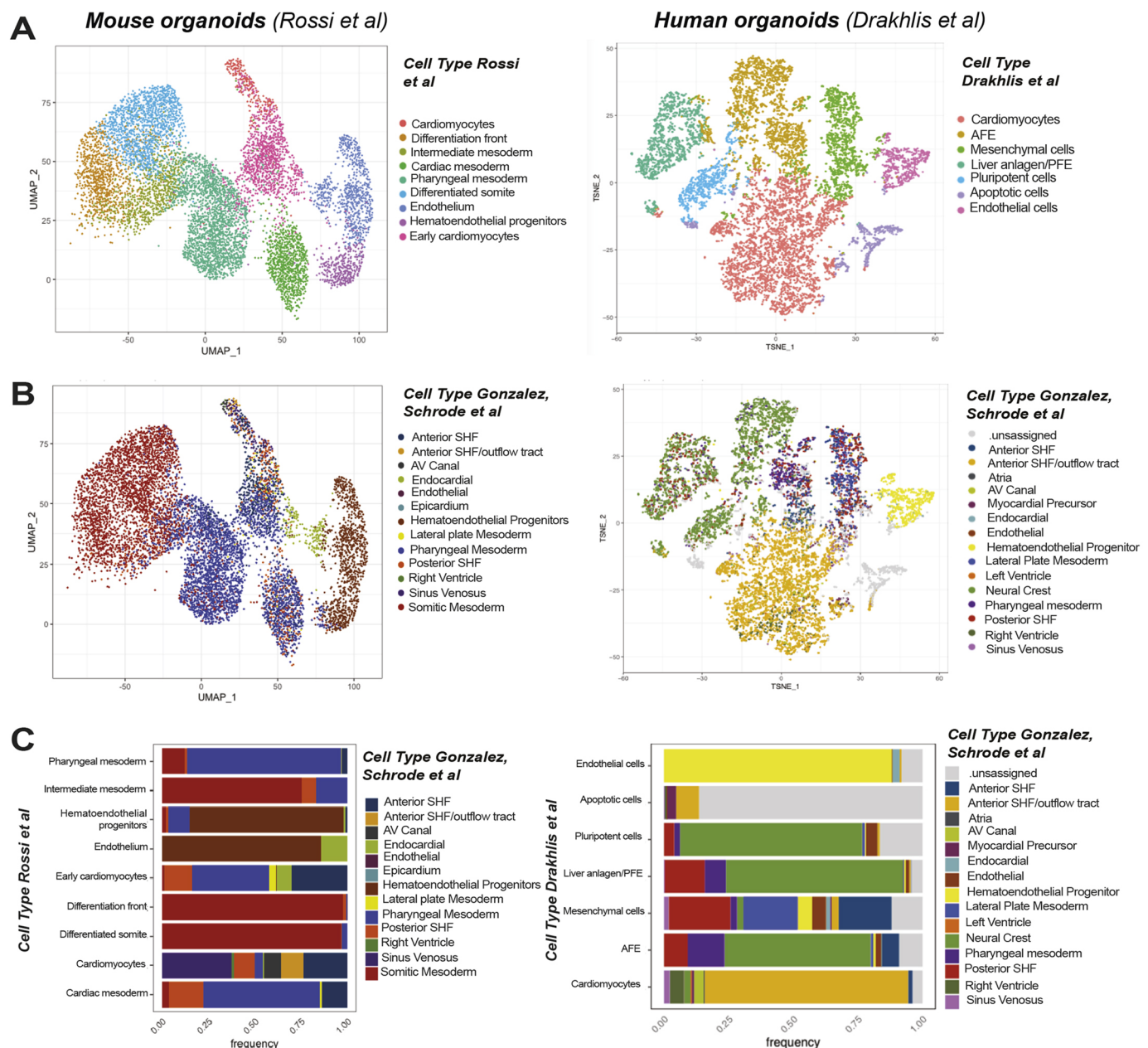
Fig. 2. Using the *in vivo* transcriptional atlas to annotate cells in mouse and human cardiac organoids. (A) Reproduction of UMAP clustering from mouse gastruloids reported by Rossi et al., 2021 (left), and human cardiac organoids reported by Drakhlis et al., 2021 (right) with author's annotation after subsetting cardiac region. (B) UMAP clustering of organoids, re-annotated following use of a random forest classifier trained on embryonic *in vivo* scSeq data from Gonzalez, Schrode et al. (C) Breakdown of composition of cluster identities from Rossi et al., 2021 (left) and Drakhlis et al., 2021 (right) with newly assigned identities based on classifier training in this study. AFE, anterior foregut endoderm; AV, atrioventricular; PFE, posterior foregut endoderm; SHF, second heart field.

which point these dynamically regulated genes were being expressed. By plotting the splicing dynamics and expression over time of individual candidates from this list of 300 genes, we were able to better understand how these patterns varied for individual populations. We found that genes such as *Vsnl1* were most strongly upregulated in cell types found at the earlier/intermediary stages of differentiation (C12, C14) and were later downregulated in atrial/AVC cells (Fig. 3H). In contrast, *Ezh2*, which was also highly dynamically regulated was highly expressed at earlier stages but became downregulated later during differentiation. This is consistent with previous studies demonstrating a role for *Ezh2* in repressing the cardiac progenitor genes in differentiated cardiomyocytes (Delgado-Olguín et al., 2012).

In addition to these examples we identified numerous additional candidates with previously unrevealed cell type-specific regulation across the LPM/pSHF and aSHF differentiation streams. A number of the top dynamically regulated genes identified were specific to one differentiation stream, and had no previously identified role in heart development. These included candidates such as *Prtg*, which plays roles in neurons and the developing tooth germ (Fig. 3K) (Takahashi et al., 2010; Wong et al., 2010). *Prtg* was strongly upregulated in differentiating aSHF cells but had low expression in differentiated cardiomyocytes, similar to another dynamically regulated gene *Meis2*. *Meis2* has been implicated in cardiac and facial defects and is essential for cranial and cardiac neural crest development (Louw et al., 2015; Machon et al., 2015), but its
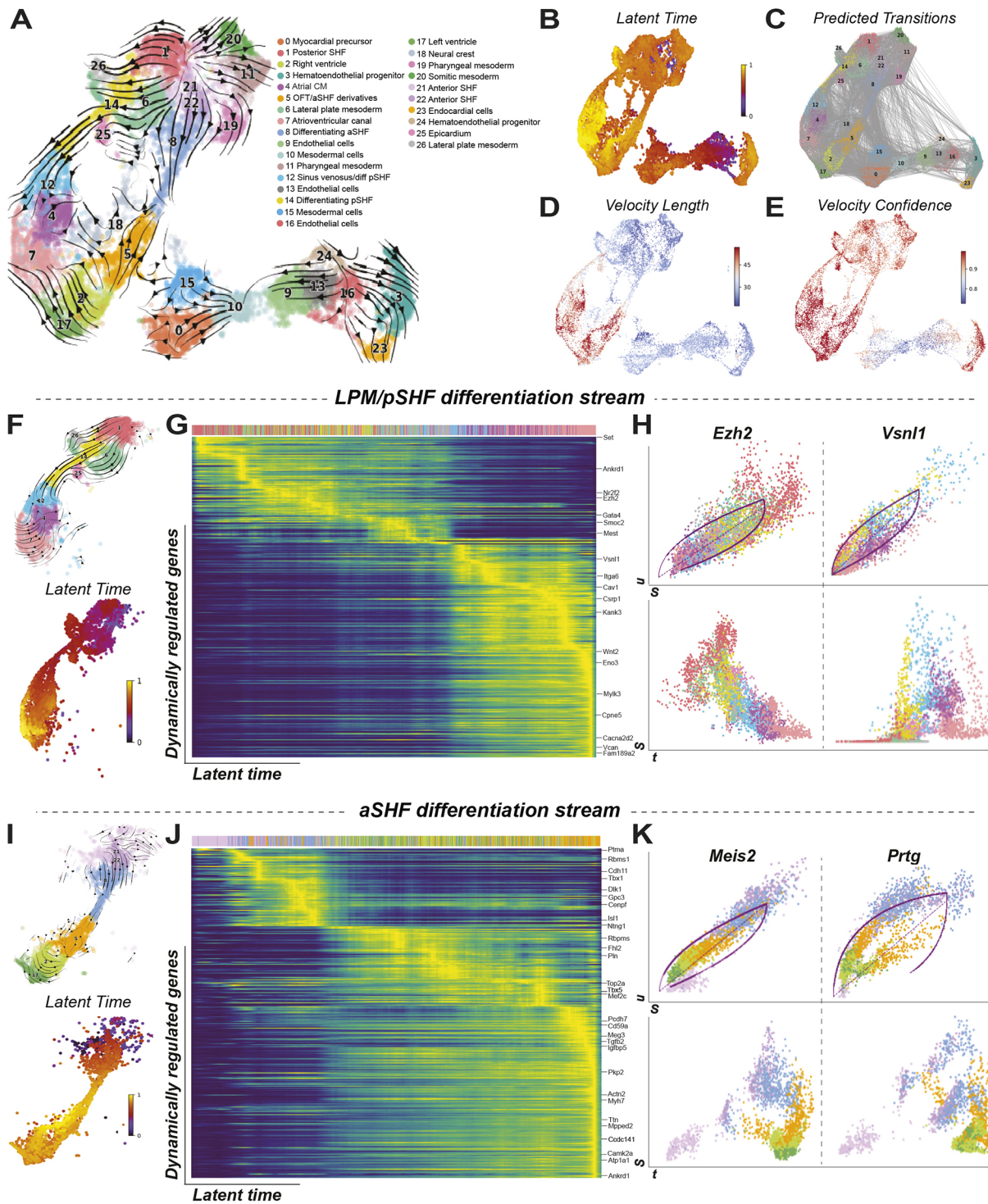
Fig. 3. Characterization of cardiomyocyte subtypes and corresponding lineage relationships to progenitor cells. (A) RNA velocity demonstrating differentiation relationships within cardiac subsets. (B) Latent time scores overlaid on UMAP clusters. Color indicates inferred developmental time. (C-E) Calculation of predicted transitions (C), velocity vector length (D) and velocity confidence (E). Color indicates velocity length and confidence score, measures of the speed of differentiation and the coherence of the vector field (i.e. how the vector correlates with its neighbors), respectively. (F,I) Subset of pSHF/aSHF differentiation streams (top) with overlaid latent time values (bottom). (G,J) Heatmap of top 300 dynamically regulated genes for pSHF/aSHF differentiation streams with selected candidates highlighted. (H,K) Examples of dynamically regulated genes along the pSHF/aSHF differentiation streams. Top: phase portraits demonstrate ratio of unspliced to spliced transcripts, colored according to cluster population in Fig. 1B. Dotted line represents steady state expression. Bottom: expression dynamics along latent time (right) demonstrate activity of genes across time. aSHF, anterior second heart field; CM, cardiomyocytes; LPM, lateral plate mesoderm; NC, neural crest; OFT, outflow tract; pSHF, posterior second heart field; S, spliced transcript; t, time; u, unspliced transcript.

expression pattern within the differentiating aSHF has not previously been recognized, and raises the question of whether observed defects in patients are purely due to a neural crest etiology (Fig. 3I). Also dynamically regulated with a cell type-specific expression was *Smyd1*, which was strongly upregulated in a subset of the differentiated aSHF-derived cells (C5) (Fig. S6D). Smyd1

plays a well-established role in orchestrating early heart development, particularly in the OFT/SHF (Franklin et al., 2016; Rasmussen et al., 2015; Wang et al., 2021), and the selective upregulation of this gene within distinct aSHF derivatives upregulation of this gene within the aSHF derivatives raises interesting questions about whether this occurs in a more granular fashion than previously appreciated. Further studies of these and other candidates identified here may shed new light on the mechanistic role of these genes in a cell type-specific and temporally defined context.

SHF progenitors undergo early differentiation through transcriptionally distinct mechanisms

We next studied the transition from pSHF and aSHF cells toward their progeny at the earliest stage of differentiation, which was captured by C14 and C8, respectively, based on RNA velocity length (Fig. 4A). We unexpectedly found that gene set enrichment analysis uncovered that processes involved in 'left ventricle development' were differentially regulated in C14 (Fig. 4A), despite its presumed origin from the pSHF. When examining distribution of *Foxa2*-lineage traced cells we found that *EGFP+* cells formed a distinct line that segregated from the atrial cell types and extended from the LPM population to ventricular cardiomyocytes (Fig. 4B). This expression pattern is reminiscent of the distribution of *Tbx5+/Hcn4+* cells from the CC that correlated with the FHF, which contributes to the left ventricle (LV) (Fig. 4A). We interrogated the expression of *Stard10* and *Lbh*, markers of early atrial and ventricular identity, respectively, and found *Stard10* expression restricted to atrial and SV cells (C4, C12), whereas *Lbh* was expressed across both differentiation streams, and its expression closely overlapped with *EGFP+* cells (Fig. 4B). These data indicate that the stream of rapid differentiation captured by cluster 14 does not simply describe a population of differentiating atrial cells derived from the pSHF. Rather, it likely represents a mix of pSHF derivatives as well as ventricular-fated cells from the LPM that differentiate through similar mechanisms as those employed by the pSHF. This might suggest that intermediary cells at this stage of development are defined more so by their progenitor of origin and its transcriptomic signature during differentiation rather than eventual atrial/ventricular identity.

Pairwise DGE analysis of the parallel trajectories identified differences in various pathways including Wnt, Hh, TGF-β and FGF signaling (Fig. 4C and Table S4). We observed higher expression of canonical Wnt genes such as *Wnt2*, *Sfrp1* and *Sfrp5* in pSHF cells (C14), whereas the noncanonical Wnt ligand *Wnt5a* was highly expressed in differentiating aSHF cells (C8). Differentiation of aSHF progenitors further involved upregulation of *Bmp4*, as well as *Fgf8*, which was expressed exclusively in the differentiating aSHF, consistent with previously published data (De Zoysa et al., 2020; High et al., 2009; Pradhan et al., 2017). Although some key components of the cardiac gene regulatory network (GRN) such as *Mef2c* and *Gata6* displayed shared expression between C8 and C14, the differentiating LPM/pSHF shows increased expression of *Tbx5* and *Tbx20*, as well as *Gata4* in comparison with differentiation of the aSHF. In contrast, aSHF cells demonstrated increased expression of a number of Irx and Etv family members (Fig. 4C). In this way, our data uncovered both known and new candidates differentially expressed in a/pSHF cells, confirming that differentiation of aSHF and pSHF occurs through activation of shared as well as discrete pathways characteristic for each path.

We next plotted the splicing dynamics and expression over time of the most highly dynamically regulated genes for C8 and C14,

reasoning that these might represent genes governing early transition from the LPM/heart field progenitors towards cardiomyocytes. *Cap2*, a gene implicated in cardiac conduction and sudden cardiac death, was among the top dynamically regulated candidates in C14 (Field et al., 2015; Peche et al., 2013). *Cap2* showed rapid upregulation compared with cycling pSHF cells, suggesting that *Cap2* may play an important early role in development, consistent with its role as a regulator of actin dynamics (Fig. 4D). *Ccdc141* was also among the most highly dynamically regulated genes in C14 and is implicated in heart rate regulation (van den Berg et al., 2017), but no functional studies during heart development have been reported to date. *Nebl* and *Nexn*, two genes encoding components of the Z-disc, were among the top five most dynamically regulated genes for C8 and have been associated with familial cardiomyopathies (Fig. 4E). Similar to many candidates identified here, their role during development has not yet been studied; their upregulation at the earliest stage of the differentiation stream suggests that these may play a key role early on. We also identified dynamic regulation of *Mest*, which is known to be expressed in the trabeculae during development but otherwise has no known function in heart development (King et al., 2002). *Mest* was downregulated in C14 but rapidly upregulated in the aSHF derivative population (C5), perhaps indicating a transient requirement for this gene, or context-specific roles at different stages of development. These examples highlight interesting patterns not only in gene expression levels, but also in the dynamic changes in the regulation of their expression within individual populations and along differentiation trajectories. This is true both for highly dynamic genes identified using scVelo, as well as well-studied genes involved in cardiac differentiation such as *Nkx2-5*, the expression of which was positively regulated during differentiation but was negatively regulated following acquisition of myocardial identity later in the differentiation process (Fig. 4F). We observed that *Myl7* (an atrial marker in adult cardiomyocytes) was widely expressed in differentiating progenitors irrespective of atrial/ventricular identity but became downregulated at later stages of differentiation in ventricular cells. This underscores a previously established notion that some canonical postnatal atrial/ventricular markers are expressed in a non chamber-specific manner early during development, potentially indicating that they perform different functions during early development compared with later stages or alternatively, highlighting processes that are only transiently required in some lineages but remain key throughout life in others. A more thorough understanding of candidate genes with shifting dynamics, in particular those which are unique to LPM/pSHF and aSHF differentiation trajectories may help to further delineate these populations and their distinct differentiation potential.

Identification of markers associated with atrial/ventricular specification across multiple differentiation states

Our data indicate that the differentiation streams captured by the clustering and RNA velocity analysis did not directly capture segregation of atrial/ventricular lineages. To understand the mechanisms governing formation of atrial/ventricular lineages, we performed pairwise DGE analysis of differentiated atrial and ventricular cardiomyocytes (C4 and C2/C17, respectively) (Fig. 5A). We identified 3695 differentially expressed genes, including expected markers such as *Irx4*, *Myl2*, *Mpped2* and *Stard10* and markers not previously associated with atrial/ventricular identity (Fig. 5B; Table S5). To predict potential functional consequences of these atrial/ventricular candidate genes, we performed GO and KEGG pathway enrichment analysis
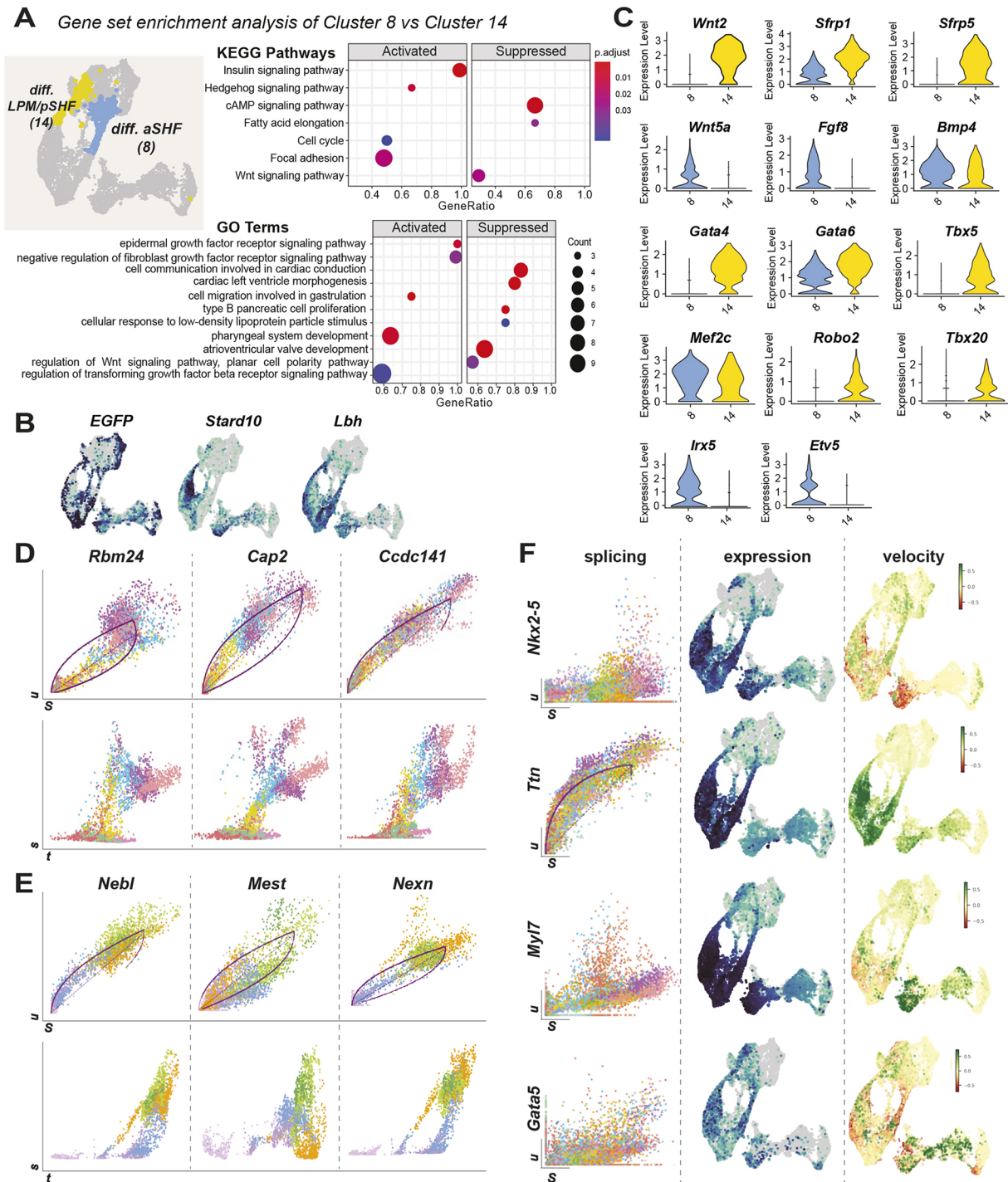
Fig. 4. Differentiation of second heart field subpopulations occurs through transcriptionally distinct mechanisms. (A) Schematic of clusters analyzed [aSHF differentiating cells (8); pSHF differentiating cells (14)] and KEGG pathway (top) and GO term (bottom) analysis of upregulated and downregulated processes between clusters. Color of dot represents adjusted *P*-value; size of dot represents number of genes within GO term uncovered by DGE analysis. (B) Feature plots for *EGFP* and markers of atrial (*Stard10*) and ventricular (*Lbh*) identity. (C) Violin plots of selected differentially expressed genes across C8 and C14. (D,E) Dynamically regulated genes during differentiation and identity across LPM/pSHF differentiation stream (C14, D), and across aSHF differentiation stream (C8, E). Phase portraits (top) demonstrate ratio of unspliced to spliced transcripts, colored according to cluster population in Fig. 1B. Dotted line represents steady state expression. Expression dynamics along latent time (bottom) demonstrate activity of genes across time. (F) Differential regulation of *Nkx2-5*, *Ttn*, *Myl7* and *Gata5*. Panels for each gene show spliced (*x*-axis) versus unspliced (*y*-axis) transcripts (left); expression of specified gene (middle); RNA velocity score based on ratio of unspliced versus spliced transcripts (right), in which positive score (green) indicates upregulation of gene, negative score (red) indicates downregulation of gene. aSHF, anterior second heart field; LPM, lateral plate mesoderm; pSHF, posterior second heart field; S, spliced transcript; t, time; u, unspliced transcript.

(Fig. 5C). We observed upregulation of processes involved in 'regulation of calcium ion transport' and 'mitochondrial respiratory chain complex assembly', and 'cardiac myofibril assembly' within ventricular cells, consistent with the increased load on ventricular compared with atrial cells. We also observed differences in a number of pathways including Hh, JAK/STAT and Ras signaling, as
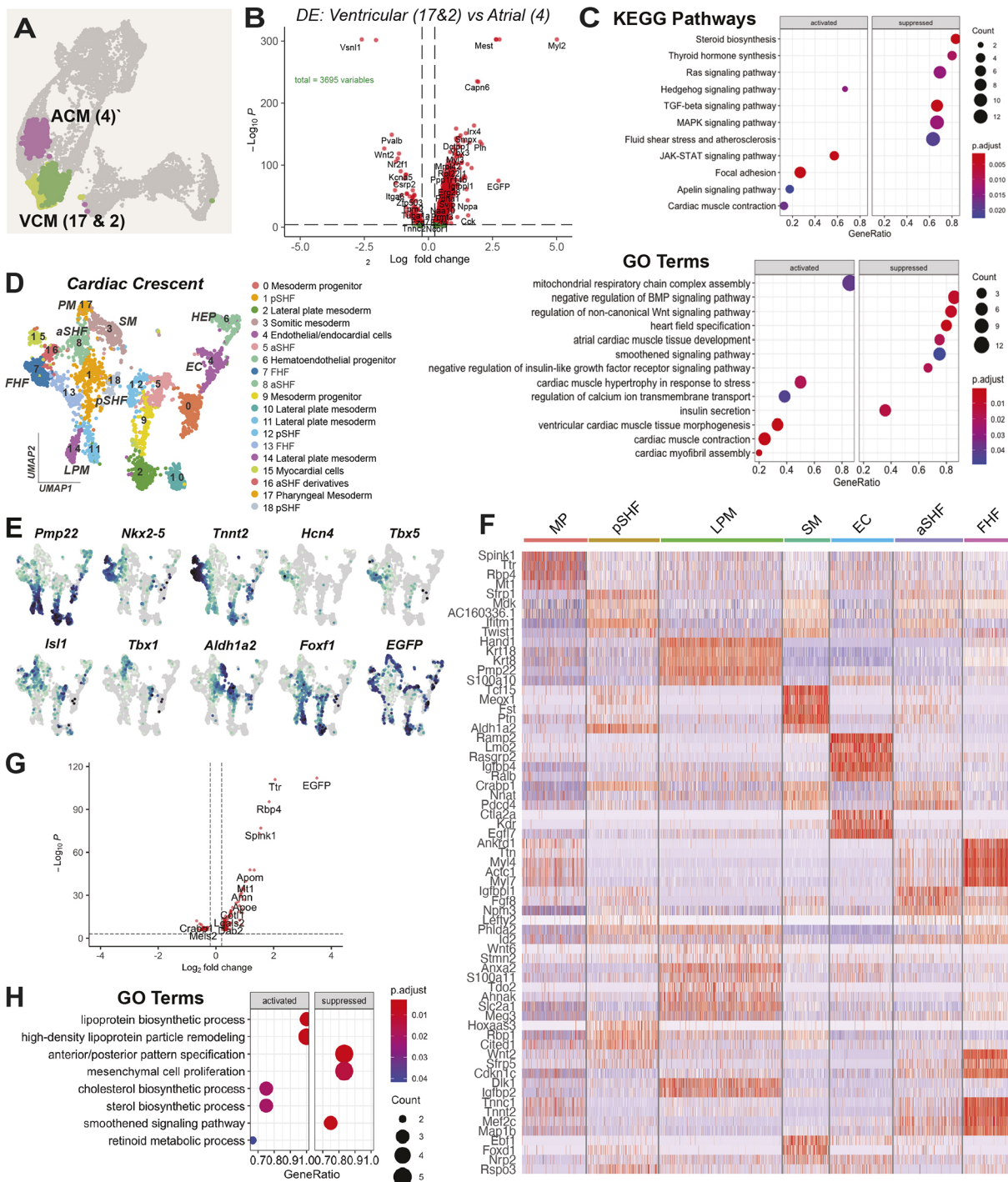
Fig. 5. Identification of differentially expressed genes and regulatory networks of atrial and ventricular cells at multiple developmental stages.
(A) Schematic of atrial (C4) and ventricular (C2, C17) populations used for pairwise DGE analysis. (B) Volcano plot demonstrating differential expression of ventricular (C17, C2) versus atrial (C4) cells. DGE analysis was performed using negative binomial distribution test in Seurat with $P<0.01$. (C) Selected upregulated/downregulated GO terms and KEGG pathways in ventricular compared with atrial cells. (D) UMAP clustering of cardiac cell types at the CC stage. (E) Feature plots for key markers. (F) Heatmap demonstrating top five most differentially expressed genes for each cell type. Differentially expressed genes were identified using Wilcoxon rank-sum test with cutoff P-value <0.01. (G) Volcano plot of differential expression across *EGFP*+/− cells within the CC. (H) Selected upregulated/downregulated GO terms between *EGFP*+/− cells at the CC. In C and H, color of dot represents adjusted P-value; size of dot represents number of genes within GO term represented in the DGE analysis. ACM, atrial cardiomyocyte; aSHF, anterior second heart field; EC, endothelial/endocardial cell; FHF, first heart field; MP, mesoderm progenitor; pSHF, posterior second heart field; SM, somitic mesoderm; VCM, ventricular cardiomyocyte.

well as a number of metabolic processes such as 'thyroid hormone synthesis' and 'steroid biosynthesis'.

To investigate mechanisms underlying atrial/ventricular progenitor segregation at the earliest stages, before expression of chamber-specific genes, we examined the CC stage alone, which was composed of multiple progenitors including the LPM, aSHF and the pSHF, as well as a transcriptionally distinct *Hcn4+/Tbx5+* population representing the FHF (Fig. 5D-F). DGE and GO

enrichment analysis between *EGFP+* and *EGFP−* cells identified 127 differentially expressed genes (Fig. 5G; Table S6) and uncovered processes involved in 'lipoprotein biosynthesis and remodeling' (Fig. 5H). A number of lipoproteins such as *Apoe* and *Apom* were upregulated in the *EGFP+* cells. We had also observed genes such as *Apobec*2, a member of the apoB mRNA editing enzyme, to be among the top differentially expressed genes between ventricular and atrial cells (Fig. 5B; Table S5). The spatiotemporal distribution of these lipoproteins at early stages of cardiac development has not been characterized extensively, opening the question as to whether metabolic differences may be implicated in the earliest segregation of atrial and ventricular cell types during development, potentially as early as the CC.

In summary, we identified candidate genes of atrial/ventricular identity within differentiated cardiomyocytes and early progenitors, identifying a putative class of molecules that may characterize the ventricular lineage before chamber formation. Alternatively, candidates enriched in *Foxa2* lineage-traced cells at the CC stage may represent components relevant in early differentiating cells, as opposed to later differentiating cells, equally offering interesting opportunities to further dissect mechanisms of early cardiac specification. Experimental validation of the expression pattern and loss-of-function experiments of such candidates will determine their spatiotemporal distribution within the developing heart and their relevance for early heart specification.

Exogenous RA causes defects in head and heart development

Previous studies have shown that an anterior-posterior gradient of RA directs differentiation of cardiac precursors towards an atrial fate (Bernheim and Meilhac, 2020; Perl and Waxman, 2019, 2020; Waxman and Yelon, 2009). Accordingly, loss of the RA synthesis enzyme *Raldh2* (*Aldh1a2*) in the heart leads to loss of the atrial chambers (Niederreither et al., 2001). Exposure to exogenous RA *in utero* in turn is associated with a number of cardiac and facial defects, indicating that retinoids can act as teratogens (Piersma et al., 2017). Despite the well-known importance for RA signaling during development, the comprehensive transcriptional effects induced by exogenous exposure to RA are incompletely understood in the context of chamber specification.

We and others have shown that manipulations in RA signaling during cardiac development lead to defects in FHF/SHF morphology as well as dysregulation of key components of the cardiac GRN (Bardot et al., 2017a; Bertrand et al., 2011; De Bono et al., 2018; Hochgreb et al., 2003; Lin et al., 2010; Rana et al., 2014; Stefanovic and Zaffran, 2017; Xavier-Neto et al., 1999). It is well established that the effects of RA are dose-dependent, and that different concentrations can even have opposite effects (Perl and Waxman, 2019). To test the effect of RA levels we exposed embryos to different concentrations of RA at E7.25 (Fig. 6A,B). As expected, exogenous RA caused patterning defects across the entire embryo in a dose-dependent manner (Fig. 6B). Intraperitoneal injection of 65 mg/kg RA caused severe defects in the head and pharyngeal structures, along with stunted formation of caudal structures. Exposure to 16.25 mg/kg RA caused milder defects, with grossly intact development of the tail and limb bud and milder, yet nonetheless hypomorphic, effects on head and pharyngeal structures (Fig. 6B). We also observed dysregulation in heart chamber size and overall reduced embryo size. Low concentrations of RA (3.25 mg/kg) led to decreased ventricular size at E10.5 and mild craniofacial abnormalities, leading us to conclude this was more suitable for understanding selective effects of exogenous RA on early heart development.

Embryos from pregnant dams injected with 3.25 mg/kg RA intraperitoneally at E7.25 were collected at equivalent stages as in control embryos (Fig. 6C). We merged samples from CC-, PHT- and HT-stage embryos across control and RA-injected embryos and annotated cluster identity and sample contribution as previously, establishing similar UMAP projections compared with the control samples of the earlier analysis, consistent with the mild morphological changes observed (Fig. 6D-G; Fig. S7A,B).

Consistent with our morphological observations the frequency of cells contributing to ventricular cardiomyocytes (C3, C24) was decreased in RA-injected embryos relative to control (Fig. 6E). We thus first asked whether exogenous RA affected the contribution of cells to the aSHF or pSHF populations. We observed no differences in the relative proportion of cells from RA-injected embryos to aSHF, pSHF and LPM populations (C8/C17, C5/C7 and C12, respectively) (Fig. 6E). We next asked whether exogenous RA influenced the allocation of ventricular-fated progenitors at early stages of cardiogenesis. Using a quantitative whole-mount immunofluorescence approach (Bardot et al., 2017b), we quantified the percentage of the *Nkx2-5+* CC that was composed of *Foxa2* lineage-traced cells in control and RA-injected embryos at E8.25 (Fig. 6H; Fig. S7E). We found no significant differences between both groups, indicating that changes in ventricular chamber size were not due to a shift in chamber-specific fate of progenitors and instead suggested that the observed defects might originate in aberrant differentiation of progenitors.

Defects in differentiation of aSHF progenitors toward cardiomyocytes causes a shunt toward pharyngeal mesoderm formation

In order to identify trajectories of differentiation, we visualized RNA velocity across the various populations from both control and RA-injected embryos (Fig. 7A), noting dysconnectivity between the aSHF cells (C8, C17) and the differentiating aSHF derivatives (C10), which cluster adjacent to ventricular cells (Fig. 6D). This prompted us to ask whether defects in differentiation of aSHF cells was affected by exogenous RA. A neighboring mesenchymal cell cluster (C6) was strongly over-represented in RA embryos, suggesting that exogenous RA may cause defects in the normal lineage trajectory of aSHF progenitors resulting in a shunt of these cells toward a pharyngeal/mesenchymal cell type, consistent with the multipotent nature of the aSHF (Fig. 7A).

This led us to interrogate transcriptional changes in the aSHF as a result of RA exposure, finding that exogenous RA caused a downregulation in processes involved in 'myotube development', 'ventricular cardiac muscle development' and 'cardiac myofibril assembly' (Fig. 7B). These were accompanied by an increase in processes involved in 'head morphogenesis' and 'head development', consistent with the subtle head deformities observed previously. We further found increased negative regulation of FGF signaling, which is known to be required for differentiation of aSHF cells towards cardiomyocytes (Pradhan et al., 2017; de Soysa et al., 2019). Consistent with this, aSHF cells from RA-exposed embryos had decreased expression of cardiomyocyte markers such as *Myl2* and *Nkx2-5*, with downregulation of *Bmp4*, which plays a key role in differentiation of the aSHF progenitors. These changes were accompanied by increases in *Pbx1* and *Sox9*, both genes involved in development of the pharyngeal lineages (Cheung and Briscoe, 2003; Welsh et al., 2018). These data suggest that exogenous RA causes defects in different signaling pathways that subsequently may result in a block in differentiation of aSHF towards the ventricular lineage and a
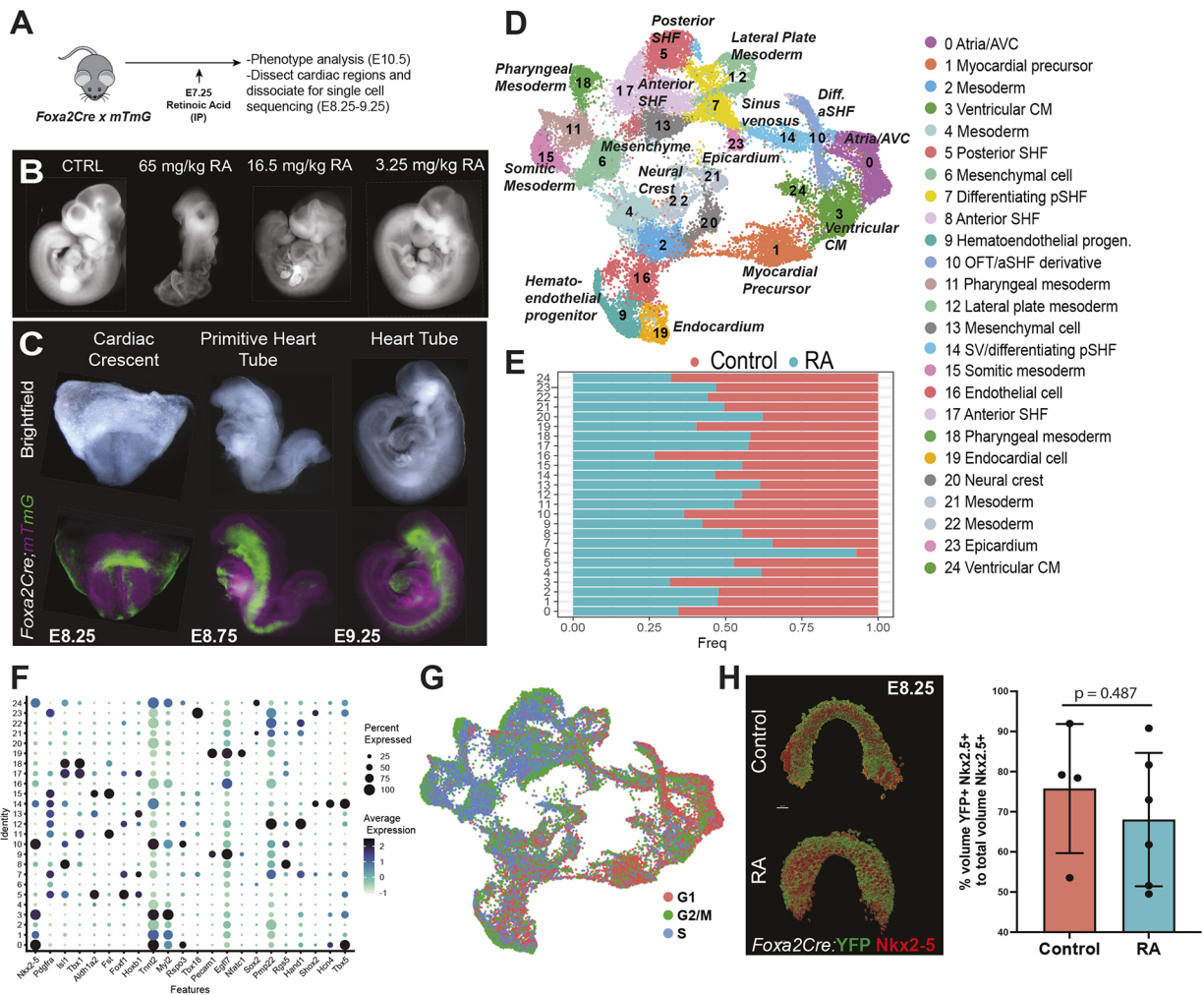
Fig. 6. scRNA-seq following *in utero* exposure to RA reveals defects in ventricular development and lineage relationships between aSHF and ventricular differentiation. (A) Schematic of RA-exposure model and experimental workflow. (B) Brightfield images of control and *in utero*-exposed embryos demonstrating dose-dependent effects on heart and head development. (C) Brightfield (top) and lineage-tracing information (bottom) of representative embryos at stages sequenced showing the morphology of RA-induced defect at each stage. (D) UMAP clustering of merged dataset comprising control and RA-exposed embryos. (E) Contribution to each cluster from control (red) and RA samples (teal). Frequency is calculated as total number of cells within cluster, relative to total number of cells within a treatment condition. (F) Bubble plot of selected markers across individual clusters. Color of dot represents the average expression level in each cluster; size of dot represents the percentage of cells in each cluster expressing the gene of interest. (G) Cell cycle score for cell types in control and RA-exposed samples. (H) Left: representative image of surface volume used for visualization of YFP (*Foxa2Cre;R26RYFP*) population (green) within Nkx2-5 population (red). Right: quantification of percentage YFP volume. Data are mean±s.d. *P*-value is calculated using the Mann-Whitney test. aSHF, anterior second heart field; AVC, atrioventricular canal; CM, cardiomyocytes; IP, intraperitoneal; LPM, lateral plate mesoderm; OFT, outflow tract; pSHF, posterior second heart field; SHF, second heart field; SV, sinus venosus.

shunt towards other lineages within its multipotent capacity, such as the pharyngeal mesoderm.

In order to better visualize this shunt in differentiation, we subset the aSHF populations as well as the ventricular cardiomyocytes and mesenchymal/pharyngeal mesoderm populations and reclustered them (Fig. 7C,D). We observed diverging fates from the aSHF toward the ventricular cardiomyocytes and the pharyngeal/mesenchymal cell types (Fig. 7E). In order to better visualize these transition states between populations, we constructed trajectories in URD, using the aSHF as a starting population and the mesenchymal/ventricular cardiomyocytes/pharyngeal mesoderm clusters as the terminal population. We observed that cells from RA-treated embryos had impaired differentiation toward a *Tnnt2+* ventricular cardiomyocyte identity at later time points, and were over-represented in the differentiation arm towards mesenchymal cell types, suggesting again that RA-treatment had caused impaired

differentiation toward cardiomyocyte identity and instead resulted in the formation of mesenchymal/pharyngeal cell types (Fig. 7F-H).

When examining the effect of RA on other relevant populations, we discovered increased expression of processes involved in biosynthetic machinery such as 'tRNA metabolism' and 'amino acid synthesis' within early cardiomyocyte progenitors (C1) (Fig. 7I). This was coupled with decreased sarcomere organization and differentiation toward cardiomyocytes accompanied by dysregulation in Wnt signaling, potentially suggesting that these cells were arrested during differentiation towards a cardiac lineage. This was further supported by RNA velocity projections, which demonstrated cycling cardiomyocyte progenitors that may be inhibited in their ability to contribute further to differentiated cardiomyocytes (Fig. 7A). The progenitors in RA-exposed embryos demonstrated decreased *Nkx2-5* and *Tnnt2* expression but interestingly had increased expression of
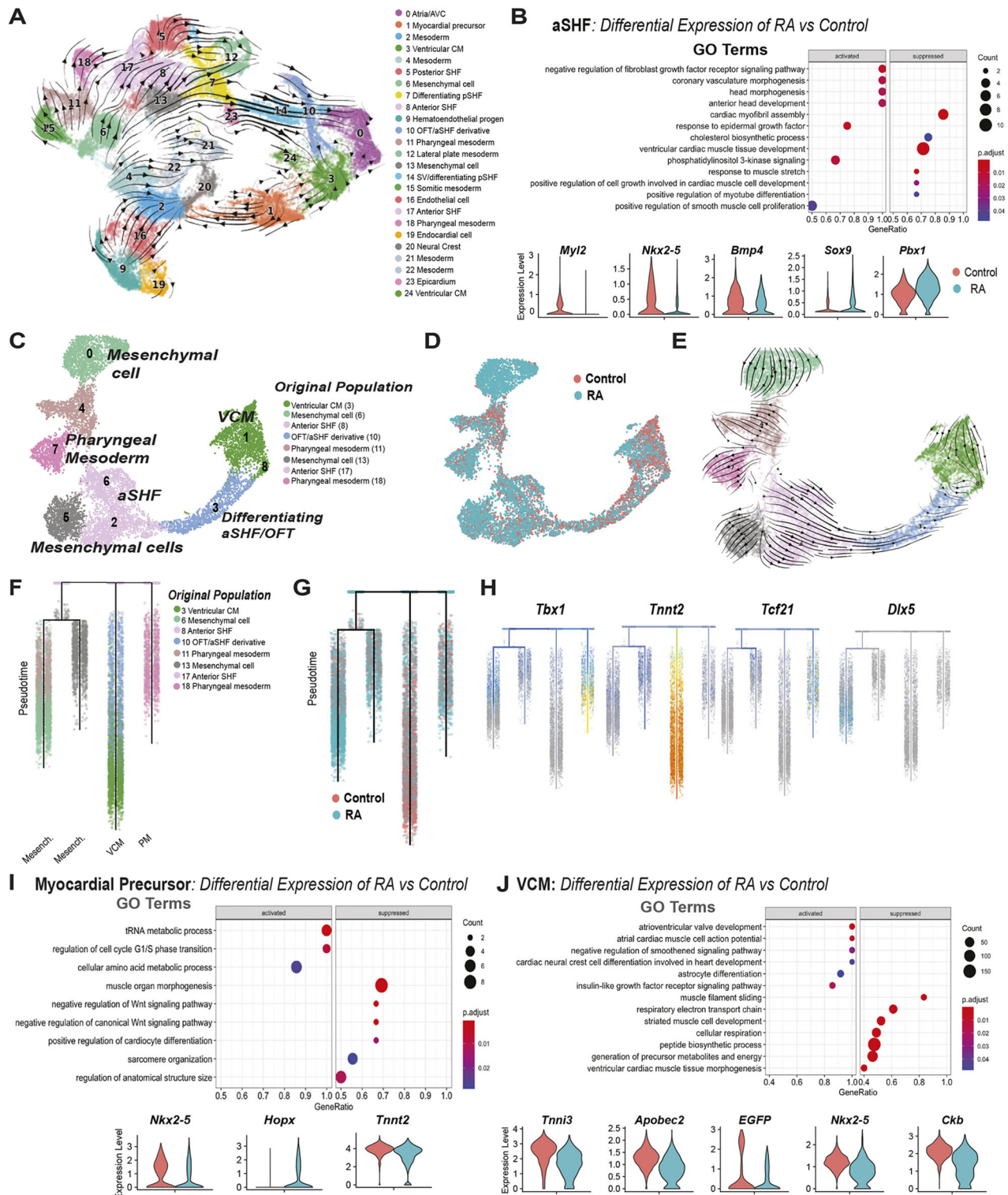
Fig. 7. Differential expression analysis of ventricular cells and associated progenitors shows dysregulation in processes related to ventricular development and differentiation. (A) RNA velocity plot demonstrating directionality of differentiation across cell types. (B) Top: selected up- and downregulated GO terms within aSHF. Bottom: selected differentially expressed candidates between control (red) and RA (teal) cells within aSHF. (C) Subclustering of aSHF cell types and derivatives, labeled with original population in A. (D) Contribution of Control and RA-treated embryos to UMAP projection. (E) RNA velocity map of aSHF derivative UMAP projection. (F,G) URD trajectories demonstrating cell state transitions, colored by cell type (F) and condition (G). (H) Expression of characteristic markers across URD trajectory. (I,J) Top: selected up- and downregulated GO terms within myocardial precursor (I) and ventricular cardiomyocytes (J). Bottom: selected differentially expressed candidates between control (red) and RA (teal) cells within myocardial precursor (I) and ventricular cardiomyocytes (J). In B, I and J, color of dot represents adjusted *P*-value and size of dot represents number of genes within the GO term uncovered by the DGE analysis. aSHF, anterior second heart field; AVC, atrioventricular canal; CM, cardiomyocytes; OFT, outflow tract; pSHF, posterior second heart field; SHF, second heart field; SV, sinus venosus; VCM, ventricular cardiomyocytes.

DEVELOPMENT

Hopx, perhaps indicative of an arrest of differentiation at the cardiomyoblast stage that could not be overcome to form functional cardiomyocytes (Jain et al., 2015).

Finally, we investigated changes in RA-exposed differentiated ventricular cardiomyocytes, where we uncovered decreases in processes involved in 'striated muscle development' and 'ventricular muscle morphogenesis', but also found that RA caused defects in metabolic and energetic components such as the respiratory electron transport chain (Fig. 7J). Other genes that were downregulated in the RA condition included sarcomeric genes and cardiac transcription factors such as *Nkx2-5* and the *Foxa2* lineage-tracing marker *EGFP*, as well as ventricular markers such as *Ckb* (Fig. 7J). Collectively, this indicated that not only were ventricular cells present at lower frequency due to the improper differentiation of progenitors, but those that did exist may also have decreased fitness due to perturbations in their maturation and metabolism.

DISCUSSION

Here, we have generated a single cell transcriptomic atlas to interrogate dynamic changes of differentiation over time, building on existing work defining the heterogeneity of heart field progenitors (Ivanovitch et al., 2021; de Soysa et al., 2019; Tyser et al., 2021; Zhang et al., 2021). We uncovered broadly similar cell populations and investigated in depth their relationship to each other through a variety of computational lineage trajectory tools. We show that FHF cells integrated with differentiating cardiomyocytes when analyzed across real developmental time, adding to a growing understanding of the FHF representing a differentiating early cardiomyocyte population that is altogether different from its SHF progenitor counterpart, which undergoes replication and retains multipotent capacity. We were intrigued to find clusters of SHF progenitors at multiple stages of differentiation, including their 'stable' replicative state as well as intermediary states during differentiation. Visualization of gene expression across differentiation time uncovered comprehensive information about the temporal ordering of genes involved in cardiac differentiation and the dynamics of their expression in a cell type-specific manner. This identified several candidates previously described in genetic mutants – for example, the upregulation of *Fgf8* and *Bmp4* in differentiation of the aSHF toward the OFT/RV (De Zoysa et al., 2020; Hutson et al., 2010; Ilagan et al., 2006; McCulley et al., 2008; Park et al., 2006; Watanabe et al., 2012) – but also uncovered new concepts such as the differences in canonical/non-canonical Wnt signaling across aSHF and LPM/pSHF differentiation. Although the importance of Wnt signaling in anterior/posterior patterning and heart formation has long been recognized (Ai et al., 2007; Gessert and Kühl, 2010; Marvin et al., 2001; Sinha et al., 2015; Yamaguchi, 2001), we have shown here that different components of Wnt signaling are used in distinct patterns during differentiation from the arterial/venous poles of the heart. Our analysis also showed that differentiation of the pSHF and aSHF progenitors involves deployment of separate but sometimes overlapping components of the cardiac GRN such as *Tbx5/Tbx20* and Robo/Slit (Bruneau, 2013; Paige et al., 2015; Srivastava, 2006). This raises questions about how different cardiac progenitors might use discrete components of the GRN during differentiation and what upstream mechanisms are driving these differences.

We identified the genes with highest velocity for each cluster and across the course of differentiation from aSHF and pSHF/LPM progenitors to cardiomyocytes. The highly dynamic regulation of these genes suggests that some members of this list may act as putative drivers of differentiation, though follow-up studies would

be needed to determine this experimentally. These included a number of candidates for which the role in cardiac development has not previously been shown, such as *Prtg*, or candidates that have been implicated in cardiac disease but have an unclear role in heart development, such as *Cap2* and *NebI/Nexn*. Importantly, our data revealed not only the ordered cascade of these dynamically regulated genes and other known markers across cardiac differentiation, it also showed how the dynamic regulation of these genes can occur in a cell type-specific context and in intermediate populations during development. This information should be used to inform the design of functional follow-up studies of these candidates, for example in the selection of genes for conditional deletion/knock out studies, or the stages that should be profiled to capture any potential phenotypes.

We further characterized the atrial and ventricular cardiomyocyte populations and identified putative candidates of atrial/ventricular fate and identity. Use of *Foxa2* lineage tracing identified genes and processes divergent at the earliest stages of differentiation toward a ventricular lineage, including a number of lipoproteins such as *Apom* and *Apoe* and mRNA binding proteins such as *Apobec2*. These have not been studied in depth during cardiac development but play roles in myoblast differentiation in skeletal muscle and regulation of left-right axis specification in other systems (Ohtsubo et al., 2017; Vonica et al., 2011). The differences in metabolic components between atrial and ventricular cells may be related to expected differences in hemodynamic load between these chambers and the increased energetic demands on the ventricle as it continues to develop, and mirrors changes in metabolism that drive maturation of ventricular cells at later stages of development (Chattergoon, 2019; Murphy et al., 2021; Wickramasinghe et al., 2022). Alternatively, these may represent genes involved in early differentiation, as our previous work has indicated that isolation of *Foxa2* lineage-traced cells during migration of cardiac mesoderm can identify early migrating cells that experience different signaling environments than their posterior counterparts (Bardot et al., 2020 preprint).

In-depth single cell data from developmentally relevant stages promise to provide reliable and easy-to-compare-to blueprints for newly emerging *in vitro* models. When compared with our embryo data, organoid-derived cardiomyocytes mapped across several progenitor subtypes but primarily overlapped with intermediate cell types of differentiating pSHF/aSHF as opposed to chamber-specific cardiomyocytes. These data highlight the utility of organoids for capturing transient differentiation events during early development, which are thought to be particularly affected in CHDs (Bruneau, 2008, 2020; Kloesel et al., 2016). It remains unclear, however, whether the mapping to discrete progenitors also informs the distribution of more differentiated cardiomyocytes that may arise at later stages in culture. This analysis will be important for understanding both the limitations of these models as well as providing strategies for their further development. In addition, this underscores the power of cross-model comparisons for a deeper understanding of these questions, and should be further explored across data from other animal models (Asp et al., 2019; Cao et al., 2020; Cui et al., 2019; Holowiecki et al., 2020; Mantri et al., 2021; Wang et al., 2019).

The use of scRNA-seq data to understand cell type-specific effects of genetic defects has been explored previously (Kathiriya et al., 2021; de Soysa et al., 2019). We extended this to an *in utero* exposure model of RA, as RA has been well-characterized as a teratogen (Lammer et al., 1985; Piersma et al., 2017). The hypomorphic ventricular phenotype observed in RA-exposed

embryos is consistent with classical studies in the chick (Osmond et al., 1991; Yutzey et al., 1994), zebrafish (Stainier and Fishman, 1992) and mouse (Xavier-Neto et al., 1999), demonstrating that exogenous RA results in changes to atrial and ventricular chamber size in a concentration-dependent fashion (Bernheim and Meilhac, 2020; Perl and Waxman, 2019; Waxman and Yelon, 2009). Further studies have shown that depletion of RA receptors results in a low increase in RA concentration and an increase in the number of atrial cardiomyocytes, and combining a Cyp26a1 depletion with it results in an intermediate increase in RA concentration, with reduced ventricular cardiomyocytes and no effect on atrial cells (D'Aniello et al., 2013). The work presented here builds on these studies and aims to determine the transcriptional effects induced directly by RA signaling in a cell type-specific manner (Hochgreb et al., 2003; Lin et al., 2010; Perl and Waxman, 2020). We demonstrate that *in utero* exposure to RA at E7.5 did not affect the distribution of heart field progenitors between control and RA-exposed conditions, which was interesting as previous work has shown a role for RA in restriction of the cardiac progenitor pool (Duong et al., 2021; Keegan et al., 2004, 2005). This may be because we elected to inject at a time point following migration of the LPM, in an effort to focus on the effect of RA on cardiac mesodermal cell differentiation and expansion. Future work should examine whether earlier injection of RA at similar concentrations might recapitulate previously observed defects in progenitor size, and determine whether this occurs in an atrial/ventricular-specific manner.

Our observation that aSHF differentiation towards a myocardial fate was negatively affected by RA exposure is consistent with previous work demonstrating posterior expansion of *Fgf8+* and *Isl1+* SHF progenitors and defects in OFT septation in RA-deficient mouse mutants. Antagonism of Fgf8 and RA has been well-characterized in the literature in various contexts, placing RA at the top of an important signaling cascade driving growth and differentiation of the developing heart (del Corral et al., 2003; Cunningham et al., 2013; Pasini et al., 2012). Interestingly, recent work has indicated that RA-deficient zebrafish embryos display defects in differentiation of SHF-derived cardiomyocytes. It is unclear whether this represents differences across species, or perhaps concentration differences between KO studies and our low-dose RA exposure model. The increase in processes involved in head morphogenesis is also consistent with the role of RA as a teratogen, and a cause for craniofacial malformations and microcephaly (Petrelli et al., 2019). Future work should determine whether an RA concentration-dependent mechanism fine tunes control of SHF progenitor expansion/differentiation across various lineages, or whether the increase in pharyngeal mesoderm specification is an otherwise aberrant differentiation event driven by restriction of cardiomyocyte fate.

MATERIALS AND METHODS
Mice
Foxa2Cre mice (C57BL/6J) mice were shared with us by Dr Heiko Lickert (Uetzmann et al., 2008) and maintained on a mixed background. *Rosa26-mTmG* mice were obtained from the Jackson Laboratory. All murine embryo scRNA-seq studies were performed on *Foxa2Cre;R26R: mTmG* embryos derived from crossing homozygous *Foxa2Cre* males with homozygous *Rosa26-mTmG* reporter females between 2-3 months of age. For timed matings, the day of plug identification was considered to be E0.5. All animals were housed in the Center for Comparative Medicine and Surgery (CCMS) facilities at Icahn School of Medicine and all experiments were conducted in accordance with the guidelines and approval of the Institutional Animal Care and Use Committee at Icahn School of Medicine at Mount Sinai. For studies on the effect of RA, pregnant *mTmG* females previously crossed with *Foxa2Cre* males were injected intraperitoneally at

E7.25 with 65 mg/kg, 16.25 mg/kg or 3.25 mg/kg of RA dissolved in corn oil. Single cell analysis on RA-exposed embryos was performed following injection of 3.25 mg/kg RA.

Dissection and isolation of murine cardiac tissues
Foxa2Cre;mTmG embryos were dissected at E8.25, E8.75 and E9.25. The cardiac region was sub-dissected at each stage (Fig. S1), including the surrounding endoderm, head folds and pharyngeal structures. This was carried out in order to ensure complete dissection of the cardiac structures and also to include transcriptional information within key tissues that co-develop with the cardiac structures at each stage. Embryos were dissected in cold PBS with 0.1% bovine serum albumin (BSA) and kept on ice until dissociation. Three stage-matched littermates of unknown sex were incubated separately in 200 µl of 0.25% Trypsin-EDTA in a 37°C water bath for 5 min or until the tissue dissociated completely, vortexing once halfway. Trypsin solution was quenched with 1 ml DMEM with 10% fetal bovine serum (FBS), centrifuged at 300 *g* for 5 min, and resuspended in 50 µl PBS with 0.1% BSA for scRNA-seq preparation.

Whole-mount lineage trace quantification
Quantification of the lineage-traced *Foxa2Cre:YFP* contributions to the CC was performed using previously published methods (Bardot et al., 2017b). Briefly, whole E8.25 embryos were stained for GFP and Nkx2-5 and mounted for confocal microscopy [Nkx2-5 antibody: Santa Cruz Biotechnology, sc8697, 1:100 (IF); GFP antibody: Abcam, ab13970, 1:500]. Masks were generated for the Nkx2-5 and YFP regions in Imaris Viewer and volumes were calculated for each region. We then quantified the percentage of the Nkx2-5+ volume that was occupied by YFP+ cells and compared these across control and RA-injected embryos.

Single cell cDNA library preparation and scRNA-seq
Individual samples for each time point were hashed using the Chromium platform in order to determine that downstream analysis was not affected by biological variability between embryos. After hashing, libraries were generated using the Chromium platform with the 3′ gene expression (3′ GEX) V3 kit, using an input of ∼10,000 cells. Gel-Bead in Emulsions (GEMs) were generated on the sample chip in the Chromium controller. Barcoded cDNA was extracted from the GEMs using Post-GEM RT-cleanup and amplified for 12 cycles. Amplified cDNA was then fragmented and subjected to end-repair, poly A-tailing, adapter ligation and 10x-specific sample indexing following the manufacturer's protocol. Libraries were quantified using Bioanalyzer and QuBit analysis. Libraries were sequenced in paired end mode on a NovaSeq instrument targeting a depth of 50,000 reads per cell.

Processing of sequencing reads
Sequencing data were aligned and quantified using the Cell Ranger Single-Cell Software Suite (version 3.0, 10x Genomics) against the provided GRCm38 (mm10) mouse reference genome. Lineage tracing information was determined by further alignment to a custom reference for *EGFP* using provided maps of constructs used in creation of *mTmG* mice from Jackson Laboratory. Complementary Tomato expression information could not be obtained owing to increased distance from the polyA tail compared with *EGFP*.

Clustering and differential gene expression analysis
Downstream differential expression and clustering analysis was performed using the Seurat V.4.0 package, as described in the tutorials (http://satijalab.org/seruat/). cellRanger matrices were imported for each sample, and distributions of nCountRNA, nFeatures and percentage mitochondrial gene expression were examined for each sample in order to filter out doublets or cells of low quality. Cells with greater than 20% of genes coming from mitochondrial genes were selected against, as well as those with fewer than 200 genes.

We then normalized the resulting subset Seurat objects using the scTransform workflow and further scaled and normalized the RNA assay in order to perform downstream differential expression analysis. We

performed principal component analysis using the highly variable genes for each sample. The most significant principal components (20-30 depending on sample) were used for graph-based unsupervised clustering (FindClusters and FindNeighbor functions). Uniform manifold approximation and projection (UMAP) was performed using a standard resolution parameter of 0.5 and iteratively modified after performing marker gene expression and examining expression of key markers.

In order to generate a combined Seurat object encompassing all three time points, we visualized expression of cardiac and mesoderm genes (*Pdgfra*, *Nkx2-5*), endothelial/endocardial (*Pecam1*, *Nfatc1*), endoderm (*Epcam*, *Sox17*) and ectoderm (*Sox2*, *Pou3f1*). We subset clusters with positive expression of mesoderm/cardiac markers, and more completely annotated clusters that did not express these markers but clustered closely through GO/KEGG analysis and DGE analysis. These represented related lineages such as the neural crest or early/related mesoderm progenitors. These were subset together with the cardiac lineages for each stage.

Seurat objects for the CC, PHT and HT stages were labeled with their original sample ID and merged together using the scTransform-based integration workflow (Hafemeister and Satija, 2019). For analyses involving comparison of control and RA-injected embryos, the above steps were repeated by combining objects from both conditions either for corresponding stages or across all samples. Downstream clustering and dimensionality reduction was performed using similar methods described above. Cell cCycle quantification and annotation was performed on merged samples using the CellCycleScoring vignette available from the Satija lab (https://satijalab.org/). Quantification of time point, condition or sample contribution to clusters was determined by normalizing the frequency of a given metadata classification to the total cell number from each sample in order to control for cell number differences across samples.

In most cases, differential expression was determined using the FindMarkers function on the 'RNA' assay, using a negative binomial model with parameters for min.pct=0.2 and a *P*-value cut off of <0.01 either by comparing a single cluster with all others for annotation, or in pairwise fashion for clusters of interest. Differential expression of *Foxa2*-lineage-traced positive and negative cells was analyzed by classifying each cell with a binary variable based on expression of *EGFP*>1.0 and then performing differential expression on subsets of cells across the *EGFP*+/− metadata information. A similar strategy was employed to query differentially expressed genes from embryos post RA injection on a cluster-by-cluster basis across control and RA condition metadata information.

Gene set enrichment and GO/KEGG analysis

Differential gene expression was performed in Seurat through the FindMarkers function in Seurat using a negative binomial test using a *P*-value cutoff of <0.05. Gene set enrichment for GO terms and KEGG pathways were performed using the ClusterProfiler tool (https://guangchuangyu.github.io/software/clusterProfiler/) using a cutoff *P*-value of <0.05.

RNA velocity analysis

RNA velocity was calculated using Velocyto and scVelo according to the developer's manual (Bergen et al., 2020; La Manno et al., 2018). Briefly, .loom files containing exon and intron information of all samples were created using the Velocyto python package 'velocyto run10x()' command and merged using 'loompy.combine()' in the loompy package. Using the scVelo python package 'scv.pp.filter_and_normalize()' and 'scv.pp.moments()' functions, the data was normalized and moments computed. Subsequently the 'scv.tl.velocity(stochastic)' and 'scv.tl.velocity_graph()' functions were used to estimate RNA velocity, which was then projected onto Seurat-derived UMAP coordinates in conjunction with the 'scv.pl.velocity_embedding_stream()' function. To analyze the dynamics of individual genes of interest, phase portraits, gene-specific RNA velocity UMAP projections and expression plots were created through 'scv.pl.velocity()'. The length and coherence of the velocity vectors, which indicate differentiation speed and directional confidence, respectively, were calculated using 'scv.tl.velocity_confidence()'. To visualize individual cell connections in more detail, high confidence transitions were plotted through 'scv.pl.velocity_graph()' with a threshold of 0.5. A dynamical model was applied to analyze transcriptional states and cell-internal latent time by running 'scv.tl.recover_dynamics()' and subsequently recomputing RNA velocities ['scv.tl.velocity(dynamic)' and 'scv.tl.velocity_graph()' functions]. The latent time, based solely on the transcriptional dynamics of a cell, was then determined using 'scv.tl.latent_time()' and plotted with 'scv.pl.scatter()'. Finally, the dynamical model also uncovered putative driver genes, as these show increased dynamic behavior ['scv.tl.rank_dynamical_genes()'].

URD trajectory analysis

To reconstruct branching trajectory trees of the aSHF subset (Fig. 7C) in the control/RA merged dataset we used URD (v1.1.0). Seurat information was imported into URD, including relevant count matrices and metadata information using custom functions. We calculated a diffusion map using the calcDM function from URD with knn=200 and sigma.use=10. As the root, we assigned both aSHF cell clusters from the original dataset (C8, C17). Cells were then ordered in pseudotime by simulating diffusion from the root to calculate the distance of each cell from the root. For this, we used the floodPseudotime function with *n*=10 (number of simulations) and minimum.cells.flooded=2. In total, 200 simulations were performed. Mesenchymal cells, ventricular cardiomyocytes and pharyngeal mesoderm populations were defined as tips for the RNA full tree, using the RNA velocity information to guide choice of endpoints between related populations. To apply URD, we used pseudotimeWeightTransitionMatrix with parameters optimal.cells.forward=40 and max.cells.back=80 to determine the slope and inflection point of the logistic function used to bias the transition probabilities. We simulated random walks on the cell–cell graph from each tip to the root using processRandomWalks function from URD. In total, 35,000 random walks were performed per tip for the RNA full tree. Finally, trees were built using buildTree function, which starts from each tip and joins trajectories that visited the same cells and compares all predefined tips in a pairwise manner. The following parameters were used to build the tree: bins.per.pseudotime.window=8, cells.per.pseudotime.bin=80, divergence.method='preference', p.thresh=0.0.

Random forest classifier

Machine learning was used to compare our dataset to others by training a random forest classifier (singleCellNet R package) on fully cell type-annotated external data (Rossi et al., 2021) and assessing it on a withheld subset (Tan and Cahan, 2019). The classifier was then applied to the current dataset and a matching cell type was predicted for each cell. In turn, a classifier was trained on our data and applied to the external datasets.

Author contributions

Conceptualization: D.M.G., N.S., K.G.B., R.S., N.C.D.; Methodology: D.M.G., N.S., T.A.M.E., K.G.B., R.S., N.C.D.; Validation: D.M.G., T.A.M.E.; Formal analysis: D.M.G., N.S., T.A.M.E.; Investigation: D.M.G., T.A.M.E.; Resources: N.B., G.R., L.D., R.Z., M.P.L., K.G.B., R.S., N.C.D.; Data curation: D.M.G., N.S.; Writing - original draft: D.M.G. Writing - review & editing: D.M.G, N.S., T.A.M.E, K.G.B., R.S., N.C.D.; Visualization: D.M.G., N.S., T.A.M.E, N.C.D.; Supervision: K.G.B., R.S., N.C.D.; Project administration: K.G.B, R.S., N.C.D.; Funding acquisition: D.M.G, R.S., K.G.B., N.C.D.

References

Ai, D., Fu, X., Wang, J., Lu, M.-F., Chen, L., Baldini, A., Klein, W. H. and Martin, J. F. (2007). Canonical Wnt signaling functions in second heart field to promote right ventricular growth. *Proc. Natl. Acad. Sci. USA* **104**, 9319-9324. doi:10.1073/pnas.0701212104

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. doi:10.1016/j.cell.2019.11.025

Bardot, E., Calderon, D., Santoriello, F., Han, S., Cheung, K., Jadhav, B., Burtscher, I., Artap, S., Jain, R., Epstein, J. et al. (2017a). Foxa2 identifies a cardiac progenitor population with ventricular differentiation potential. *Nat. Commun.* **8**, 14428. doi:10.1038/ncomms14428

Bardot, E., Tzavaras, N., Benson, D. L. and Dubois, N. C. (2017b). Quantitative whole-mount immunofluorescence analysis of cardiac progenitor populations in mouse embryos. *J. Vis. Exp.* **128**, e56446. doi:10.3791/56446

Bardot, E. S., Jadhav, B., Wickramasinghe, N., Rezza, A., Rendl, M., Sharp, A. J. and Dubois, N. C. (2020). Notch signaling commits mesoderm to the cardiac lineage. *BioRxiv 2020.02.20.958348*. doi:10.1101/2020.02.20.958348

Bergen, V., Lange, M., Peidli, S., Wolf, F. A. and Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408-1414. doi:10.1038/s41587-020-0591-3

Bernheim, S. and Meilhac, S. M. (2020). Mesoderm patterning by a dynamic gradient of retinoic acid signalling. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **375**, 20190556. doi:10.1098/rstb.2019.0556

Bertrand, N., Roux, M., Ryckebüsch, L., Niederreither, K., Dollé, P., Moon, A., Capecchi, M. and Zaffran, S. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. *Dev. Biol.* **353**, 266-274. doi:10.1016/j.ydbio.2011.02.029

Bruneau, B. G. (2008). The developmental genetics of congenital heart disease. *Nature* **451**, 943-948. doi:10.1038/nature06801

Bruneau, B. G. (2013). Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harb. Perspect. Biol.* **5**, a008292. doi:10.1101/cshperspect.a008292

Bruneau, B. G. (2020). The developing heart: from The Wizard of Oz to congenital heart disease. *Development* **147**, dev194233. doi:10.1242/dev.194233

Cai, C.-L., Liang, X., Shi, Y., Chu, P.-H., Pfaff, S. L., Chen, J. and Evans, S. (2003). Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev. Cell* **5**, 877-889. doi:10.1016/s1534-5807(03)00363-0

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. doi:10.1126/science.aba7721

Chabab, S., Lescroart, F., Rulands, S., Mathiah, N., Simons, B. D. and Blanpain, C. (2016). Uncovering the number and clonal dynamics of Mesp1 progenitors during heart morphogenesis. *Cell Rep.* **14**, 1-10. doi:10.1016/j.celrep.2015.12.013

Chattergoon, N. N. (2019). Thyroid hormone signaling and consequences for cardiac development. *J. Endocrinol.* **242**, T145-T160. doi:10.1530/JOE-18-0704

Cheung, M. and Briscoe, J. (2003). Neural crest development is regulated by the transcription factor Sox9. *Development* **130**, 5681-5693. doi:10.1242/dev.00808

Chiapparo, G., Lin, X., Lescroart, F., Chabab, S., Paulissen, C., Pitisci, L., Bondue, A. and Blanpain, C. (2016). Mesp1 controls the speed, polarity, and directionality of cardiovascular progenitor migration. *J. Cell Biol.* **213**, 463-477. doi:10.1083/jcb.201505082

Cui, Y., Zheng, Y., Liu, X., Yan, L., Fan, X., Yong, J., Hu, Y., Dong, J., Li, Q., Wu, X. et al. (2019). Single-cell transcriptome analysis maps the developmental track of the human heart. *Cell Rep.* **26**, 1934-1950.e5. doi:10.1016/j.celrep.2019.01.079

Cunningham, T. J., Zhao, X., Sandell, L. L., Evans, S. M., Trainor, P. A. and Duester, G. (2013). Antagonism between retinoic acid and fibroblast growth factor signaling during limb development. *Cell Rep.* **3**, 1503-1511. doi:10.1016/j.celrep.2013.03.036

D'Aniello, E., Rydeen, A. B., Anderson, J. L., Mandal, A. and Waxman, J. S. (2013). Depletion of retinoic acid receptors initiates a novel positive feedback mechanism that promotes teratogenic increases in retinoic acid. *PLoS Genet.* **9**, e1003689. doi:10.1371/journal.pgen.1003689

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. doi:10.1093/hmg/ddy266

de Soysa, T. Y., Ranade, S. S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H. T., Schricker, A., Del Sol, A., Gifford, C. A. and Srivastava, D. (2019). Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. *Nature* **572**, 120-124. doi:10.1038/s41586-019-1414-x

De Zoysa, P., Liu, J., Toubat, O., Choi, J., Moon, A., Gill, P. S., Duarte, A., Sucov, H. M. and Kumar, S. R. (2020). Delta-like ligand 4-mediated Notch signaling controls proliferation of second heart field progenitor cells by regulating Fgf8 expression. *Development* **147**, dev185249. doi:10.1242/dev.185249

del Corral, R. D., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M. and Storey, K. (2003). Opposing FGF and retinoid pathways control ventral neural pattern, neuronal differentiation, and segmentation during body axis extension. *Neuron* **40**, 65-79. doi:10.1016/S0896-6273(03)00565-8

DeLaughter, D. M., Bick, A. G., Wakimoto, H., McKean, D., Gorham, J. M., Kathiriya, I. S., Hinson, J. T., Homsy, J., Gray, J., Pu, W. et al. (2016). Single-cell resolution of temporal gene expression during heart development. *Dev. Cell* **39**, 480-490. doi:10.1016/j.devcel.2016.10.001

Delgado-Olguín, P., Huang, Y., Li, X., Christodoulou, D., Seidman, C. E., Seidman, J. G., Tarakhovsky, A. and Bruneau, B. G. (2012). Epigenetic repression of cardiac progenitor gene expression by Ezh2 is required for postnatal cardiac homeostasis. *Nat. Genet.* **44**, 343-347. doi:10.1038/ng.1068

Devine, W. P., Wythe, J. D., George, M., Koshiba-Takeuchi, K. and Bruneau, B. G. (2014). Early patterning and specification of cardiac progenitors in gastrulating mesoderm. *Elife* **3**, e03848. doi:10.7554/eLife.03848

Drakhlis, L., Biswanath, S., Farr, C.-M., Lupanow, V., Teske, J., Ritzenhoff, K., Franke, A., Manstein, F., Bolesani, E., Kempf, H. et al. (2021). Human heart-forming organoids recapitulate early heart and foregut development. *Nat. Biotechnol.* **39**, 737-746. doi:10.1038/s41587-021-00815-9

Duong, T. B., Holowiecki, A. and Waxman, J. S. (2021). Retinoic acid signaling restricts the size of the first heart field within the anterior lateral plate mesoderm. *Dev. Biol.* **473**, 119-129. doi:10.1016/j.ydbio.2021.02.005

Field, J., Ye, D. Z., Shinde, M., Liu, F., Schillinger, K. J., Lu, M., Wang, T., Skettini, M., Xiong, Y., Brice, A. K. et al. (2015). CAP2 in cardiac conduction, sudden cardiac death and eye development. *Sci. Rep.* **5**, 17256. doi:10.1038/srep17256

Franklin, S., Kimball, T., Rasmussen, T. L., Rosa-Garrido, M., Chen, H., Tran, T., Miller, M. R., Gray, R., Jiang, S., Ren, S. et al. (2016). The chromatin-binding protein Smyd1 restricts adult mammalian heart growth. *Am. J. Physiol. Heart Circ. Physiol.* **311**, H1234-H1247. doi:10.1152/ajpheart.00235.2016

Garcia-Martinez, V. and Schoenwolf, G. C. (1993). Primitive-streak origin of the cardiovascular system in avian embryos. *Dev. Biol.* **159**, 706-719. doi:10.1006/dbio.1993.1276

Gessert, S. and Kühl, M. (2010). The multiple phases and faces of Wnt signaling during cardiac differentiation and development. *Circ. Res.* **107**, 186-199. doi:10.1161/CIRCRESAHA.110.221531

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. doi:10.1186/s13059-019-1874-1

High, F. A., Jain, R., Stoller, J. Z., Antonucci, N. B., Lu, M. M., Loomes, K. M., Kaestner, K. H., Pear, W. S. and Epstein, J. A. (2009). Murine Jagged1/Notch signaling in the second heart field orchestrates Fgf8 expression and tissue-tissue interactions during outflow tract development. *J. Clin. Invest.* **119**, 1986-1996. doi:10.1172/JCI38922

Hochgreb, T., Linhares, V. L., Menezes, D. C., Sampaio, A. C., Yan, C. Y. I., Cardoso, W. V., Rosenthal, N. and Xavier-Neto, J. (2003). A caudorostral wave of RALDH2 conveys anteroposterior information to the cardiac field. *Development* **130**, 5363-5374. doi:10.1242/dev.00750

Hoffmann, A. D., Yang, X. H., Burnicka-Turek, O., Bosman, J. D., Ren, X., Steimle, J. D., Vokes, S. A., McMahon, A. P., Kalinichenko, V. V. and Moskowitz, I. P. (2014). Foxf genes integrate tbx5 and hedgehog pathways in the second heart field for cardiac septation. *PLoS Genet.* **10**, e1004604. doi:10.1371/journal.pgen.1004604

Holowiecki, A., Linstrum, K., Ravisankar, P., Chetal, K., Salomonis, N. and Waxman, J. S. (2020). Pbx4 limits heart size and fosters arch artery formation by partitioning second heart field progenitors and restricting proliferation. *Development* **147**, dev185652. doi:10.1242/dev.185652

Hutson, M. R., Zeng, X. L., Kim, A. J., Antoon, E., Harward, S. and Kirby, M. L. (2010). Arterial pole progenitors interpret opposing FGF/BMP signals to proliferate or differentiate. *Development* **137**, 3001-3011. doi:10.1242/dev.051565

Ilagan, R., Abu-Issa, R., Brown, D., Yang, Y.-P., Jiao, K., Schwartz, R. J., Klingensmith, J. and Meyers, E. N. (2006). Fgf8 is required for anterior heart field development. *Development* 133, 2435-2445. doi:10.1242/dev.02408

Ivanovitch, K., Soro-Barrio, P., Chakravarty, P., Jones, R. A., Bell, D. M., Mousavy Gharavy, S. N., Stamataki, D., Delile, J., Smith, J. C. and Briscoe, J. (2021). Ventricular, atrial, and outflow tract heart progenitors arise from spatially and molecularly distinct regions of the primitive streak. *PLoS Biol.* 19, e3001200. doi:10.1371/journal.pbio.3001200

Jain, R., Li, D., Gupta, M., Manderfield, L. J., Ifkovits, J. L., Wang, Q., Liu, F., Liu, Y., Poleshko, A., Padmanabhan, A. et al. (2015). Integration of Bmp and Wnt signaling by Hopx specifies commitment of cardiomyoblasts. *Science* 348, aaa6071. doi:10.1126/science.aaa6071

Kathiriya, I. S., Rao, K. S., Iacono, G., Devine, W. P., Blair, A. P., Hota, S. K., Lai, M. H., Garay, B. I., Thomas, R., Gong, H. Z. et al. (2021). Modeling human TBX5 haploinsufficiency predicts regulatory networks for congenital heart disease. *Dev. Cell* 56, 292-309.e9. doi:10.1016/j.devcel.2020.11.020

Keegan, B. R., Meyer, D. and Yelon, D. (2004). Organization of cardiac chamber progenitors in the zebrafish blastula. *Development* 131, 3081-3091. doi:10.1242/dev.01185

Keegan, B. R., Feldman, J. L., Begemann, G., Ingham, P. W. and Yelon, D. (2005). Retinoic acid signaling restricts the cardiac progenitor pool. *Science* 307, 247-249. doi:10.1126/science.1101573

Kelly, R. G., Buckingham, M. E. and Moorman, A. F. (2014). Heart fields and cardiac morphogenesis. *Cold Spring Harb. Perspect. Med.* 4, a015750. doi:10.1101/cshperspect.a015750

King, T., Bland, Y., Webb, S., Barton, S. and Brown, N. A. (2002). Expression of Peg1 (Mest) in the developing mouse heart: involvement in trabeculation. *Dev. Dyn.* 225, 212-215. doi:10.1002/dvdy.10142

Kirby, M. L., Gale, T. F. and Stewart, D. E. (1983). Neural crest cells contribute to normal aorticopulmonary septation. *Science* 220, 1059-1061. doi:10.1126/science.6844926

Kloesel, B., DiNardo, J. A. and Body, S. C. (2016). Cardiac embryology and molecular mechanisms of congenital heart disease – a primer for anesthesiologists. *Anesth. Analg.* 123, 551-569. doi:10.1213/ANE.0000000000001451

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A. et al. (2018). RNA velocity of single cells. *Nature* 560, 494-498. doi:10.1038/s41586-018-0414-6

Lammer, E. J., Chen, D. T., Hoar, R. M., Agnish, N. D., Benke, P. J., Braun, J. T., Curry, C. J., Fernhoff, P. M., Grix, A. W., Lott, I. T. et al. (1985). Retinoic acid embryopathy. *N. Engl. J. Med.* 313, 837-841. doi:10.1056/NEJM198510033131401

Laugwitz, K.-L., Moretti, A., Lam, J., Gruber, P., Chen, Y., Woodard, S., Lin, L.-Z., Cai, C.-L., Lu, M. M., Reth, M. et al. (2005). Postnatal isl1+ cardioblasts enter fully differentiated cardiomyocyte lineages. *Nature* 433, 647-653. doi:10.1038/nature03215

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. doi:10.1038/ncb3024

Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sànchez-Dànes, A., Moignard, V., Dubois, C., Paulissen, C., Kinston, S. et al. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science* 359, 1177-1181. doi:10.1126/science.aao4174

Li, G., Xu, A., Sim, S., Priest, J. R., Tian, X., Khan, T., Quertermous, T., Zhou, B., Tsao, P. S., Quake, S. R. et al. (2016). Transcriptomic profiling maps anatomically patterned subpopulations among single embryonic cardiac cells. *Dev. Cell* 39, 491-507. doi:10.1016/j.devcel.2016.10.014

Liang, X., Wang, G., Lin, L., Lowe, J., Zhang, Q., Bu, L., Chen, Y., Chen, J., Sun, Y. and Evans, S. M. (2013). HCN4 dynamically marks the first heart field and conduction system precursors. *Circ. Res.* 113, 399-407. doi:10.1161/CIRCRESAHA.113.301588

Lin, S.-C., Dollé, P., Ryckebüsch, L., Noseda, M., Zaffran, S., Schneider, M. D. and Niederreither, K. (2010). Endogenous retinoic acid regulates cardiac progenitor differentiation. *Proc. Natl. Acad. Sci. USA* 107, 9234-9239. doi:10.1073/pnas.0910430107

Louw, J. J., Corveleyn, A., Jia, Y., Hens, G., Gewillig, M. and Devriendt, K. (2015). MEIS2 involvement in cardiac development, cleft palate, and intellectual disability. *Am. J. Med. Genet. A* 167, 1142-1146. doi:10.1002/ajmg.a.36989

Machon, O., Masek, J., Machonova, O., Krauss, S. and Kozmik, Z. (2015). Meis2 is essential for cranial and cardiac neural crest development. *BMC Dev. Biol.* 15, 40. doi:10.1186/s12861-015-0093-6

Mantri, M., Scuderi, G. J., Abedini-Nassab, R., Wang, M. F. Z., McKellar, D., Shi, H., Grodner, B., Butcher, J. T. and De Vlaminck, I. (2021). Spatiotemporal single-cell RNA sequencing of developing chicken hearts identifies interplay between cellular differentiation and morphogenesis. *Nat. Commun.* 12, 1771. doi:10.1038/s41467-021-21892-z

Marvin, M. J., Di Rocco, G., Gardiner, A., Bush, S. M. and Lassar, A. B. (2001). Inhibition of Wnt activity induces heart formation from posterior mesoderm. *Genes Dev.* 15, 316-327. doi:10.1101/gad.855501

McCulley, D. J., Kang, J.-O., Martin, J. F. and Black, B. L. (2008). BMP4 is required in the anterior heart field and its derivatives for endocardial cushion remodeling, outflow tract septation, and semilunar valve development. *Dev. Dyn.* 237, 3200-3209. doi:10.1002/dvdy.21743

Meilhac, S. M. and Buckingham, M. E. (2018). The deployment of cell lineages that form the mammalian heart. *Nat. Rev. Cardiol.* 15, 705-724. doi:10.1038/s41569-018-0086-9

Meilhac, S. M., Lescroart, F., Blanpain, C. and Buckingham, M. E. (2014). Cardiac cell lineages that form the heart. *Cold Spring Harb. Perspect. Med.* 4, a013888. doi:10.1101/cshperspect.a013888

Mjaatvedt, C. H., Nakaoka, T., Moreno-Rodriguez, R., Norris, R. A., Kern, M. J., Eisenberg, C. A., Turner, D. and Markwald, R. R. (2001). The outflow tract of the heart is recruited from a novel heart-forming field. *Dev. Biol.* 238, 97-109. doi:10.1006/dbio.2001.0409

Murphy, S. A., Miyamoto, M., Kervadec, A., Kannan, S., Tampakakis, E., Kambhampati, S., Lin, B. L., Paek, S., Andersen, P., Lee, D.-I. et al. (2021). PGC1/PPAR drive cardiomyocyte maturation at single cell level via YAP1 and SF3B2. *Nat. Commun.* 12, 1648. doi:10.1038/s41467-021-21957-z

Niederreither, K., Vermot, J., Messaddeq, N., Schuhbaur, B., Chambon, P. and Dolle, P. (2001). Embryonic retinoic acid synthesis is essential for heart morphogenesis in the mouse. *Development* 128, 1019-1031. doi:10.1242/dev.128.7.1019

Ohtsubo, H., Sato, Y., Suzuki, T., Mizunoya, W., Nakamura, M., Tatsumi, R. and Ikeuchi, Y. (2017). APOBEC2 negatively regulates myoblast differentiation in muscle regeneration. *Int. J. Biochem. Cell Biol.* 85, 91-101. doi:10.1016/j.biocel.2017.02.005

Osmond, M. K., Butler, A. J., Voon, F. C. and Bellairs, R. (1991). The effects of retinoic acid on heart formation in the early chick embryo. *Development* 113, 1405-1417. doi:10.1242/dev.113.4.1405

Paige, S. L., Plonowska, K., Xu, A. and Wu, S. M. (2015). Molecular regulation of cardiomyocyte differentiation. *Circ. Res.* 116, 341-353. doi:10.1161/CIRCRESAHA.116.302752

Park, E. J., Ogden, L. A., Talbot, A., Evans, S., Cai, C.-L., Black, B. L., Frank, D. U. and Moon, A. M. (2006). Required, tissue-specific roles for Fgf8 in outflow tract formation and remodeling. *Development* 133, 2419-2433. doi:10.1242/dev.02367

Pasini, A., Manenti, R., Rothbächer, U. and Lemaire, P. (2012). Antagonizing retinoic acid and FGF/MAPK pathways control posterior body patterning in the invertebrate chordate ciona intestinalis. *PLoS One* 7, e46193. doi:10.1371/journal.pone.0046193

Peche, V. S., Holak, T. A., Burgute, B. D., Kosmas, K., Kale, S. P., Wunderlich, F. T., Elhamine, F., Stehle, R., Pfitzer, G., Nohroudi, K. et al. (2013). Ablation of cyclase-associated protein 2 (CAP2) leads to cardiomyopathy. *Cell. Mol. Life Sci.* 70, 527-543. doi:10.1007/s00018-012-1142-y

Perl, E. and Waxman, J. S. (2019). Reiterative mechanisms of retinoic acid signaling during vertebrate heart development. *J. Dev. Biol.* 7, 11. doi:10.3390/jdb7020011

Perl, E. and Waxman, J. S. (2020). Retinoic acid signaling and heart development. *Subcell. Biochem.* 95, 119-149. doi:10.1007/978-3-030-42282-0_5

Petrelli, B., Bendelac, L., Hicks, G. G. and Fainsod, A. (2019). Insights into retinoic acid deficiency and the induction of craniofacial malformations and microcephaly in fetal alcohol spectrum disorder. *Genesis* 57, e23278. doi:10.1002/dvg.23278

Phillips, H. M., Mahendran, P., Singh, E., Anderson, R. H., Chaudhry, B. and Henderson, D. J. (2013). Neural crest cells are required for correct positioning of the developing outflow cushions and pattern the arterial valve leaflets. *Cardiovasc. Res.* 99, 452-460. doi:10.1093/cvr/cvt132

Pierpont, M. E., Brueckner, M., Chung, W. K., Garg, V., Lacro, R. V., McGuire, A. L., Mital, S., Priest, J. R., Pu, W. T., Roberts, A. et al. (2018). Genetic basis for congenital heart disease: revisited: a scientific statement from the American Heart Association. *Circulation* 138, e653-e711. doi:10.1161/CIR.0000000000000606

Piersma, A. H., Hessel, E. V. and Staal, Y. C. (2017). Retinoic acid in developmental toxicology: teratogen, morphogen and biomarker. *Reprod. Toxicol.* 72, 53-61. doi:10.1016/j.reprotox.2017.05.014

Pradhan, A., Zeng, X.-X. I., Sidhwani, P., Marques, S. R., George, V., Targoff, K. L., Chi, N. C. and Yelon, D. (2017). FGF signaling enforces cardiac chamber identity in the developing ventricle. *Development* 144, 1328-1338. doi:10.1242/dev.143719

Rana, M. S., Théveniau-Ruissy, M., De Bono, C., Mesbah, K., Francou, A., Rammah, M., Domínguez, J. N., Roux, M., Laforest, B., Anderson, R. H. et al. (2014). Tbx1 coordinates addition of posterior second heart field progenitor cells to the arterial and venous poles of the heart. *Circ. Res.* 115, 790-799. doi:10.1161/CIRCRESAHA.115.305020

Rasmussen, T. L., Ma, Y., Park, C. Y., Harriss, J., Pierce, S. A., Dekker, J. D., Valenzuela, N., Srivastava, D., Schwartz, R. J., Stewart, M. D. et al. (2015).

Smyd1 facilitates heart development by antagonizing oxidative and ER stress responses. *PLoS One* **10**, e0121765. doi:10.1371/journal.pone.0121765

Rossi, G., Broguiere, N., Miyamoto, M., Boni, A., Guiet, R., Girgin, M., Kelly, R. G., Kwon, C. and Lutolf, M. P. (2021). Capturing cardiogenesis in gastruloids. *Cell Stem Cell* **28**, 230-240.e6. doi:10.1016/j.stem.2020.10.013

Saga, Y., Hata, N., Kobayashi, S., Magnuson, T., Seldin, M. F. and Taketo, M. M. (1996). MesP1: a novel basic helix-loop-helix protein expressed in the nascent mesodermal cells during mouse gastrulation. *Development* **122**, 2769-2778. doi:10.1242/dev.122.9.2769

Sinha, T., Li, D., Théveniau-Ruissy, M., Hutson, M. R., Kelly, R. G. and Wang, J. (2015). Loss of Wnt5a disrupts second heart field cell deployment and may contribute to OFT malformations in DiGeorge syndrome. *Hum. Mol. Genet.* **24**, 1704-1716. doi:10.1093/hmg/ddu584

Srivastava, D. (2006). Making or breaking the heart: from lineage determination to morphogenesis. *Cell* **126**, 1037-1048. doi:10.1016/j.cell.2006.09.003

Stainier, D. Y. and Fishman, M. C. (1992). Patterning the zebrafish heart tube: acquisition of anteroposterior polarity. *Dev. Biol.* **153**, 91-101. doi:10.1016/0012-1606(92)90094-w

Stefanovic, S. and Zaffran, S. (2017). Mechanisms of retinoic acid signaling during cardiogenesis. *Mech. Dev.* **143**, 9-19. doi:10.1016/j.mod.2016.12.002

Stefanovic, S., Laforest, B., Desvignes, J.-P., Lescroart, F., Argiro, L., Maurel-Zaffran, C., Salgado, D., Plaindoux, E., De Bono, C., Pazur, K. et al. (2020). Hox-dependent coordination of mouse cardiac progenitor cell patterning and differentiation. *ELife* **9**, e55124. doi:10.7554/eLife.55124

Takahashi, K. F., Kiyoshima, T., Kobayashi, I., Xie, M., Yamaza, H., Fujiwara, H., Ookuma, Y., Nagata, K., Wada, H., Sakai, T. et al. (2010). Protogenin, a new member of the immunoglobulin superfamily, is implicated in the development of the mouse lower first molar. *BMC Dev. Biol.* **10**, 115. doi:10.1186/1471-213X-10-115

Tan, Y. and Cahan, P. (2019). SingleCellNet: a computational tool to classify single cell RNA-Seq data across platforms and across species. *Cell Syst.* **9**, 207-213.e2. doi:10.1016/j.cels.2019.06.004

Tyser, R. C. V., Ibarra-Soria, X., McDole, K., Arcot Jayaram, S., Godwin, J., van den Brand, T. AH.,, Miranda, A. M. A., Scialdone, A., Keller, P. J., Marioni, J. C. et al. (2021). Characterization of a common progenitor pool of the epicardium and myocardium. *Science* **371**, eabb2986. doi:10.1126/science.abb2986

Uetzmann, L., Burtscher, I. and Lickert, H. (2008). A mouse line expressing Foxa2-driven Cre recombinase in node, notochord, floorplate, and endoderm. *Genesis* **46**, 515-522. doi:10.1002/dvg.20410

van den Berg, M. E., Warren, H. R., Cabrera, C. P., Verweij, N., Mifsud, B., Haessler, J., Bihlmeyer, N. A., Fu, Y.-P., Weiss, S., Lin, H. J. et al. (2017). Discovery of novel heart rate-associated loci using the Exome Chip. *Hum. Mol. Genet.* **26**, 2346-2363. doi:10.1093/hmg/ddx113

Vonica, A., Rosa, A., Arduini, B. L. and Brivanlou, A. H. (2011). APOBEC2, a selective inhibitor of TGFβ signaling, regulates left-right axis specification during early embryogenesis. *Dev. Biol.* **350**, 13-23. doi:10.1016/j.ydbio.2010.09.016

Waldo, K. L., Kumiski, D. H., Wallis, K. T., Stadt, H. A., Hutson, M. R., Platt, D. H. and Kirby, M. L. (2001). Conotruncal myocardium arises from a secondary heart field. *Development* **128**, 3179-3188. doi:10.1242/dev.128.16.3179

Wang, W., Niu, X., Stuart, T., Jullian, E., Mauck, W. M., Kelly, R. G., Satija, R. and Christiaen, L. (2019). A single-cell transcriptional roadmap for cardiopharyngeal fate diversification. *Nat. Cell Biol.* **21**, 674-686. doi:10.1038/s41556-019-0336-z

Wang, Z., Schwartz, R. J., Liu, J., Sun, F., Li, Q. and Ma, Y. (2021). Smyd1 orchestrates early heart development through positive and negative gene regulation. *Front. Cell Dev. Biol.* **9**, 654682. doi:10.3389/fcell.2021.654682

Watanabe, Y., Zaffran, S., Kuroiwa, A., Higuchi, H., Ogura, T., Harvey, R. P., Kelly, R. G. and Buckingham, M. (2012). Fibroblast growth factor 10 gene regulation in the second heart field by Tbx1, Nkx2–5, and Islet1 reveals a genetic switch for down-regulation in the myocardium. *Proc. Natl. Acad. Sci. USA* **109**, 18273-18280. doi:10.1073/pnas.1215360109

Waxman, J. S. and Yelon, D. (2009). Increased Hox activity mimics the teratogenic effects of excess retinoic acid signaling. *Dev. Dyn.* **238**, 1207-1213. doi:10.1002/dvdy.21951

Welsh, I. C., Hart, J., Brown, J. M., Hansen, K., Rocha Marques, M., Aho, R. J., Grishina, I., Hurtado, R., Herzlinger, D., Ferretti, E. et al. (2018). Pbx loss in cranial neural crest, unlike in epithelium, results in cleft palate only and a broader midface. *J. Anat.* **233**, 222-242. doi:10.1111/joa.12821

Wickramasinghe, N. M., Sachs, D., Shewale, B., Gonzalez, D. M., Dhanan-Krishnan, P., Torre, D., LaMarca, E., Raimo, S., Dariolli, R., Serasinghe, M. N. et al. (2022). PPARdelta activation induces metabolic and contractile maturation of human pluripotent stem-cell-derived cardiomyocytes. *Cell Stem Cell* **29**, 559-576.e7. doi:10.1016/j.stem.2022.02.011

Wong, Y.-H., Lu, A.-C., Wang, Y.-C., Cheng, H.-C., Chang, C., Chen, P.-H., Yu, J.-Y. and Fann, M.-J. (2010). Protogenin defines a transition stage during embryonic neurogenesis and prevents precocious neuronal differentiation. *J. Neurosci.* **30**, 4428-4439. doi:10.1523/JNEUROSCI.0473-10.2010

Xavier-Neto, J., Neville, C. M., Shapiro, M. D., Houghton, L., Wang, G. F., Nikovits, W., Stockdale, F. E. and Rosenthal, N. (1999). A retinoic acid-inducible transgenic marker of sino-atrial development in the mouse heart. *Development* **126**, 2677-2687. doi:10.1242/dev.126.12.2677

Xiong, H., Luo, Y., Yue, Y., Zhang, J., Ai, S., Li, X., Wang, X., Zhang, Y.-L., Wei, Y., Li, H.-H. et al. (2019). Single-cell transcriptomics reveals chemotaxis-mediated intraorgan crosstalk during cardiogenesis. *Circ. Res.* **125**, 398-410. doi:10.1161/CIRCRESAHA.119.315243

Yamaguchi, T. P. (2001). Heads or tails: Wnts and anterior-posterior patterning. *Curr. Biol.* **11**, R713-R724. doi:10.1016/s0960-9822(01)00417-1

Yao, Y., Marra, A. N. and Yelon, D. (2021). Pathways regulating establishment and maintenance of cardiac chamber identity in zebrafish. *J. Cardiovasc. Dev. Dis.* **8**, 13. doi:10.3390/jcdd8020013

Yutzey, K. E. and Bader, D. (1995). Diversification of cardiomyogenic cell lineages during early heart development. *Circ. Res.* **77**, 216-219. doi:10.1161/01.RES.77.2.216

Yutzey, K. E., Rhee, J. T. and Bader, D. (1994). Expression of the atrial-specific myosin heavy chain AMHC1 and the establishment of anteroposterior polarity in the developing chicken heart. *Development* **120**, 871-883. doi:10.1242/dev.120.4.871

Zaffran, S., Kelly, R. G., Meilhac, S. M., Buckingham, M. E. and Brown, N. A. (2004). Right ventricular myocardium derives from the anterior heart field. *Circ. Res.* **95**, 261-268. doi:10.1161/01.RES.0000136815.73623.BE

Zhang, Q., Carlin, D., Zhu, F., Cattaneo, P., Ideker, T., Evans, S. M., Bloomekatz, J. and Chi, N. C. (2021). Unveiling complexity and multipotentiality of early heart fields. *Circ. Res.* **129**, 474-487. doi:10.1161/CIRCRESAHA.121.318943
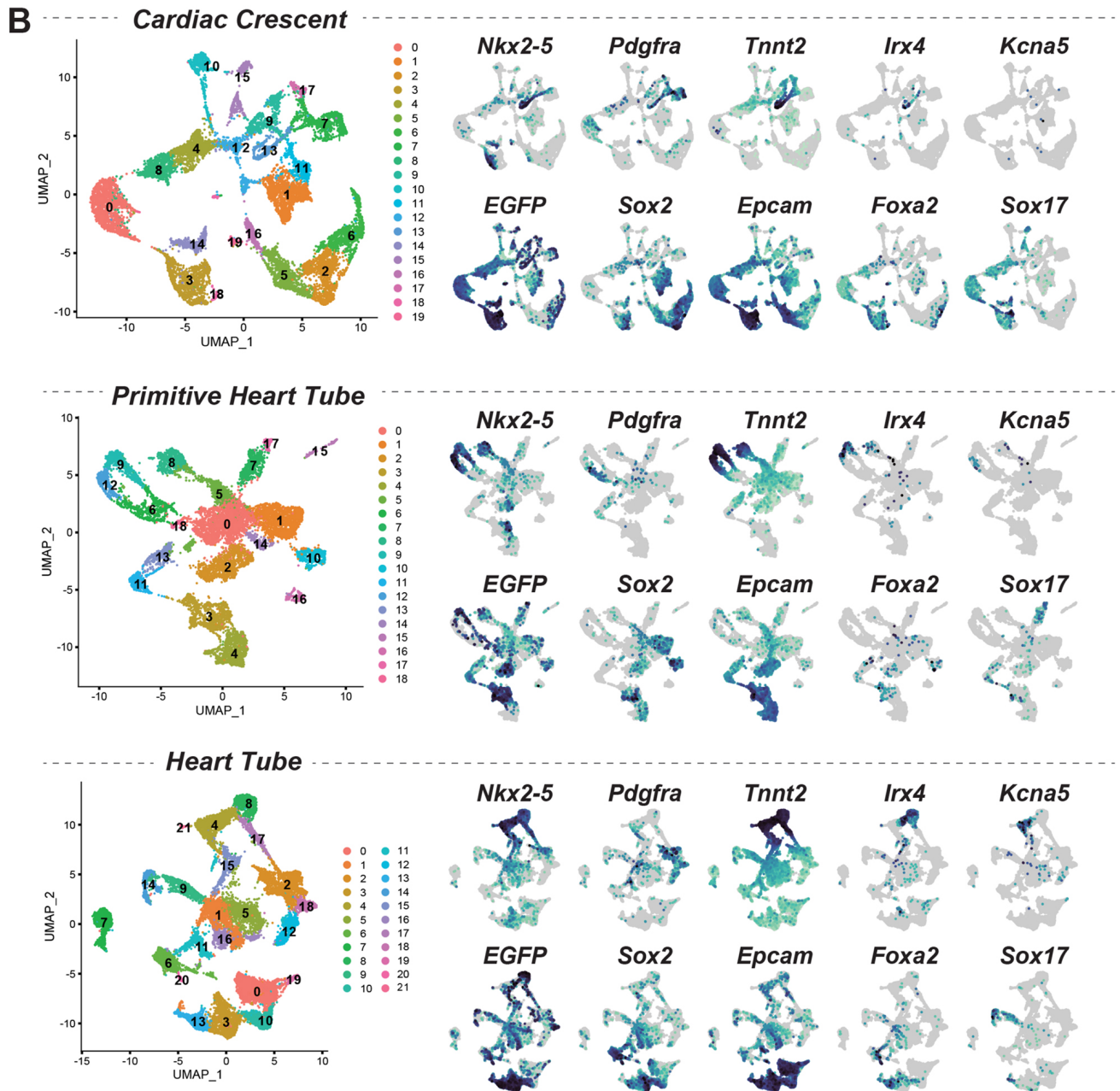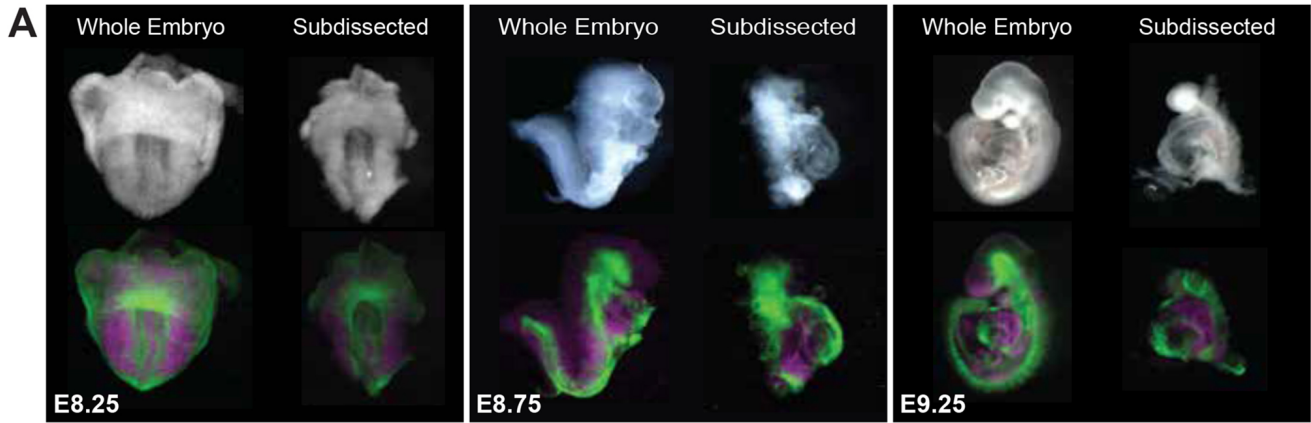
DEVELOPMENT

Fig. S1. Representative Dissection Examples and scRNAseq of Individual Stages.

A) Representative images showing brightfield (top) and fluorescence (bottom) in *Foxa2Cre;R26R:mTmG* embryos for each stage sequenced. Left: whole embryo; right: representative example of dissected cardiac regions. **B)** UMAP clustering and feature plots for markers of mesoderm, endoderm, and ectoderm cell types, as well as atrial/ventricular cells at CC, PHT and HT stage.
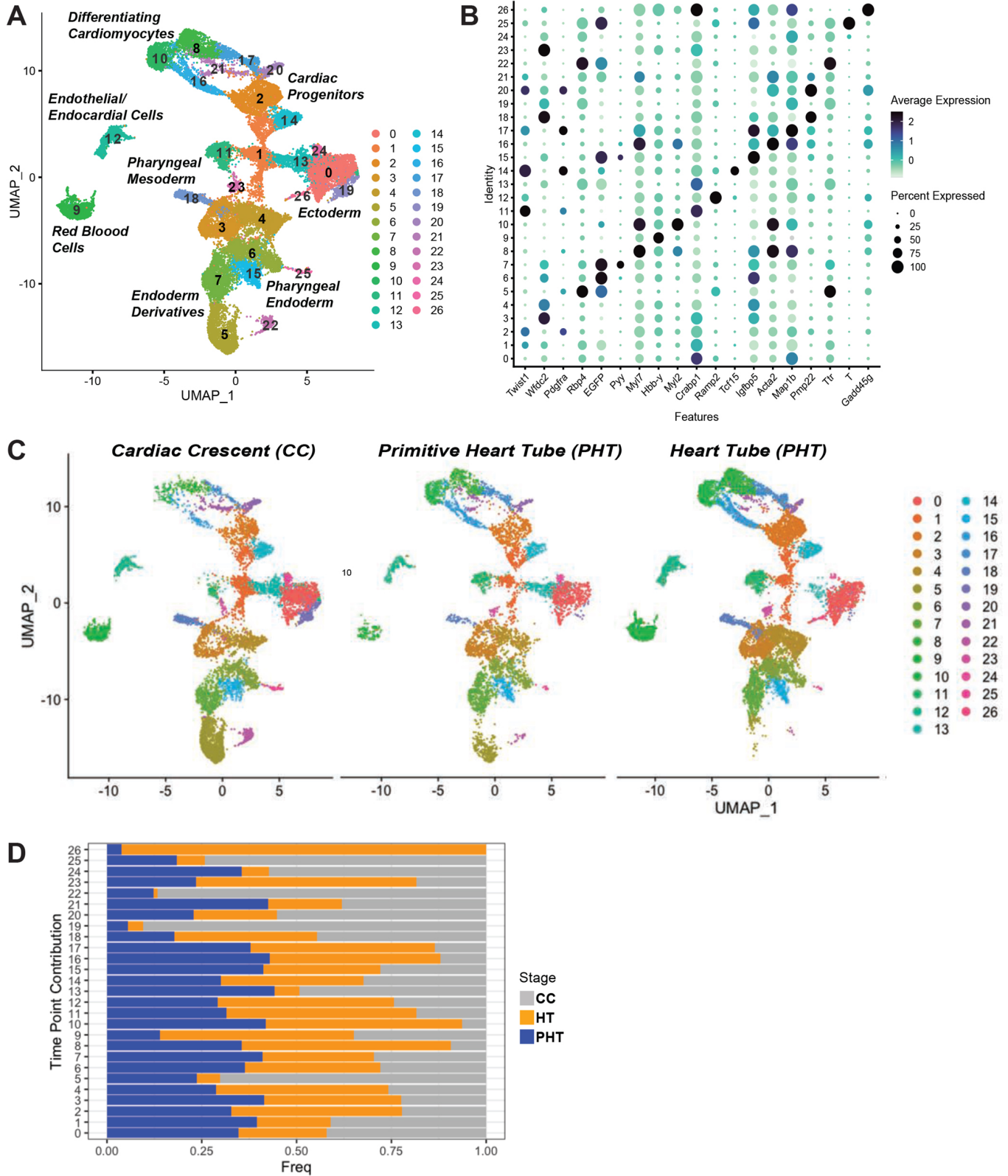
Fig. S2. Merge of Individual Time Point Samples Reveals a Changing Landscape of Cardiac Development and Surrounding Tissues

A) UMAP clustering of all cells sequenced, demonstrating presence of differentiated cardiomyocytes, cardiac progenitors, pharyngeal cell types, endoderm derivatives, and endothelial/endocardial cells. **B)** Bubble plot of top differentially expressed gene for each cluster. Differentially expressed genes for each cluster were calculated using a Wilcoxon rank-sum test with p-value <0.01. **C)** UMAP of all cells sequenced showing contribution from each time point individually. **D)** Quantification of overall contribution of cells from any one particular embryonic time point. Frequency is calculated as the number of cells in a given cluster relative to the total number of cells in the sample.
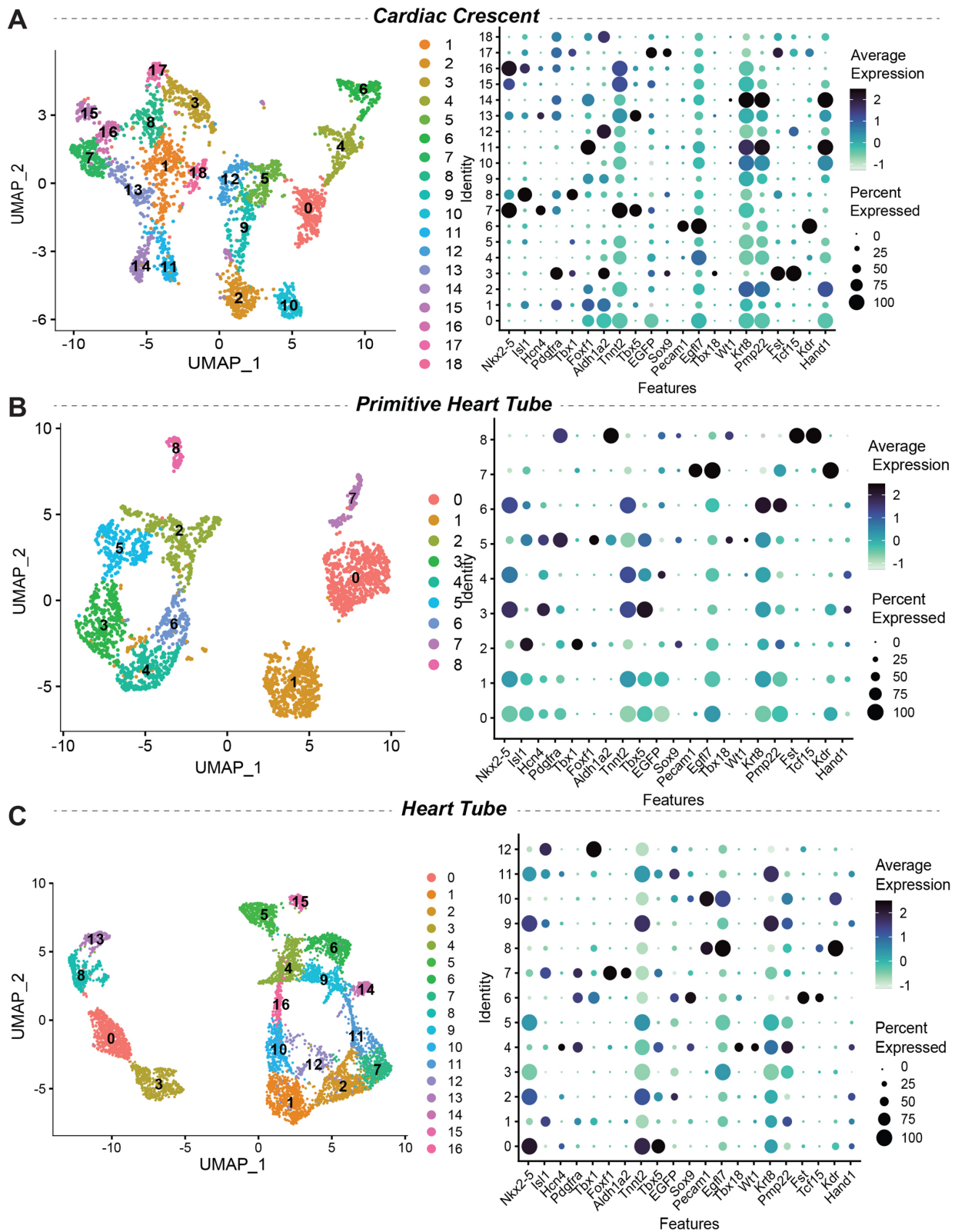
Fig. S3. UMAP Clustering and Selection of Key Markers for Cardiac Subpopulations at Individual Stages

A-C) UMAP clustering (left) and dot plot of selected genes of interest (right) of subclustered cardiac cells at the CC stage (A), the PHT stage (B) and the HT stage (C).

Fig. S4. Characterization of Cardiomyocyte Subpopulations

A) Feature plots for *Hcn4, Tbx5,* and *Tnnt2* split separately across CC, PHT and HT samples. **B)** Feature plots for expression of right (*Cited1/Cck)* and left (*Hand1/Pln)* ventricular identity. **C)** Heatmap demonstrating top 5 most differentially expressed genes for each cluster, identified using FindMarkers function in Seurat. **D)** UMAP clustering of Cluster 4 following re-clustering (bottom left) and Feature plots for key atrial markers (right). **E)** UMAP showing contribution from individual time points to combined cardiac UMAP projection.
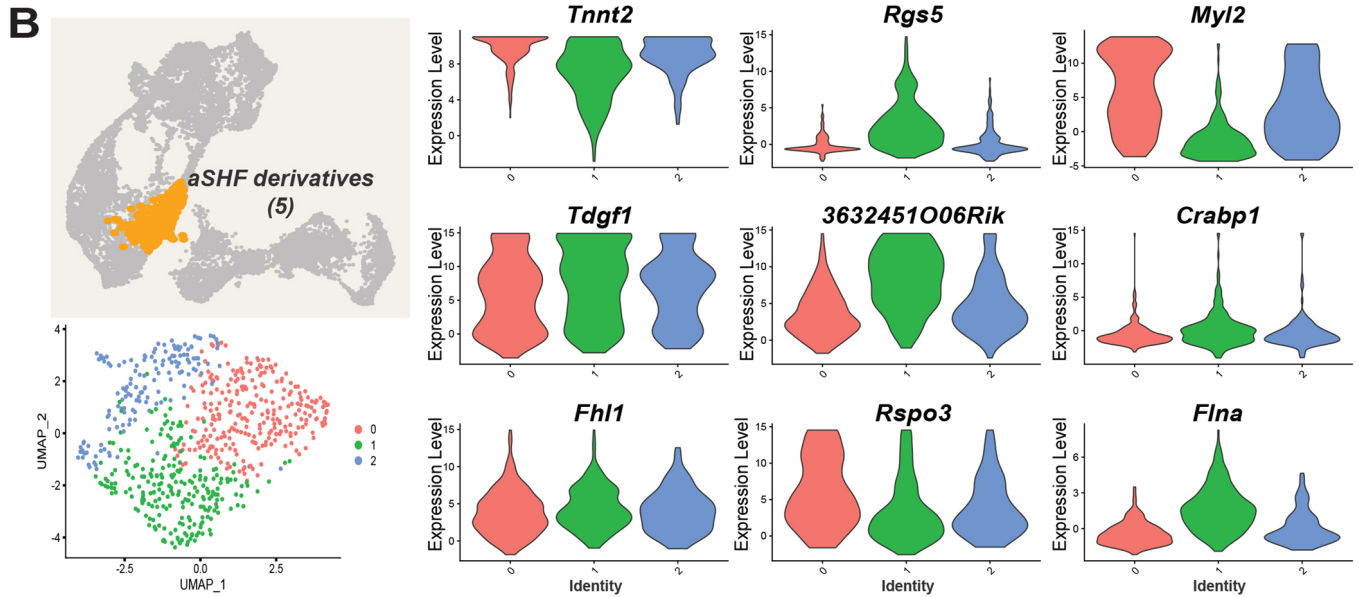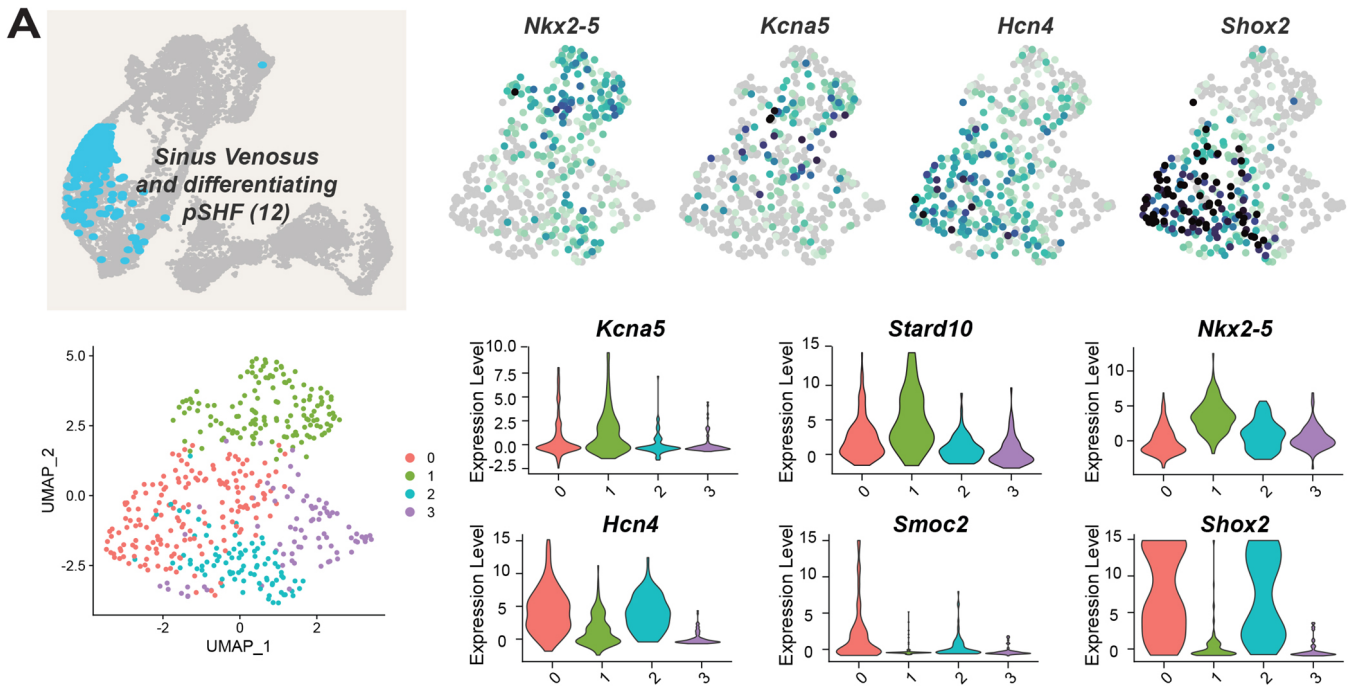
Fig. S5. Mixed Populations Cluster Together Based on Progenitor of Origin

A) *Top left:* Schematic of cluster investigated (Cluster 12)**.** *Bottom left:* UMAP clustering of Cluster 12 alone following re-clustering. *Bottom left:* Feature plots for markers of atrial and sinoatrial nodal identity. *Bottom right:* Violin plots for markers of ventricular and OFT lineages across individual clusters. **B)** *Top left:* Schematic of cluster investigated (Cluster 5)**.** *Bottom left:* UMAP clustering of Cluster 5 alone following re-clustering. *Right:* Violin plots for markers of ventricular and OFT lineages across individual clusters.
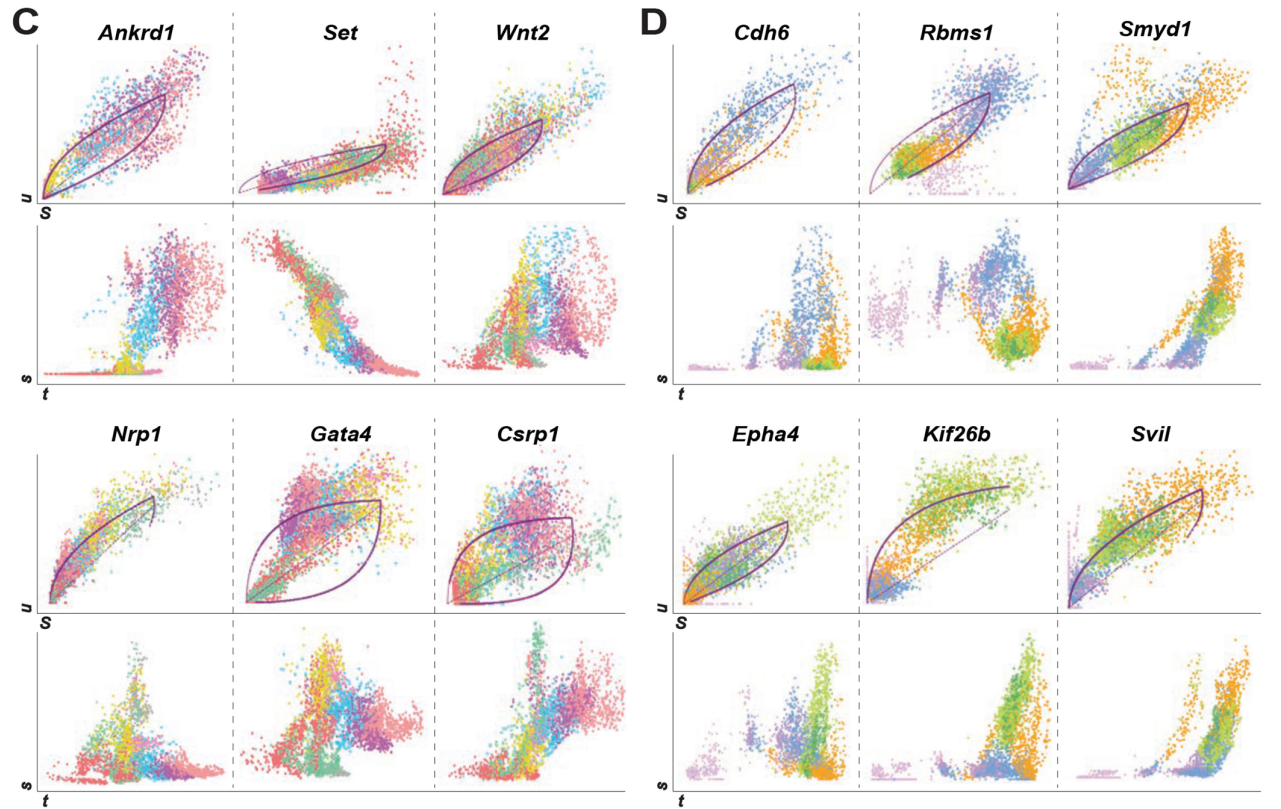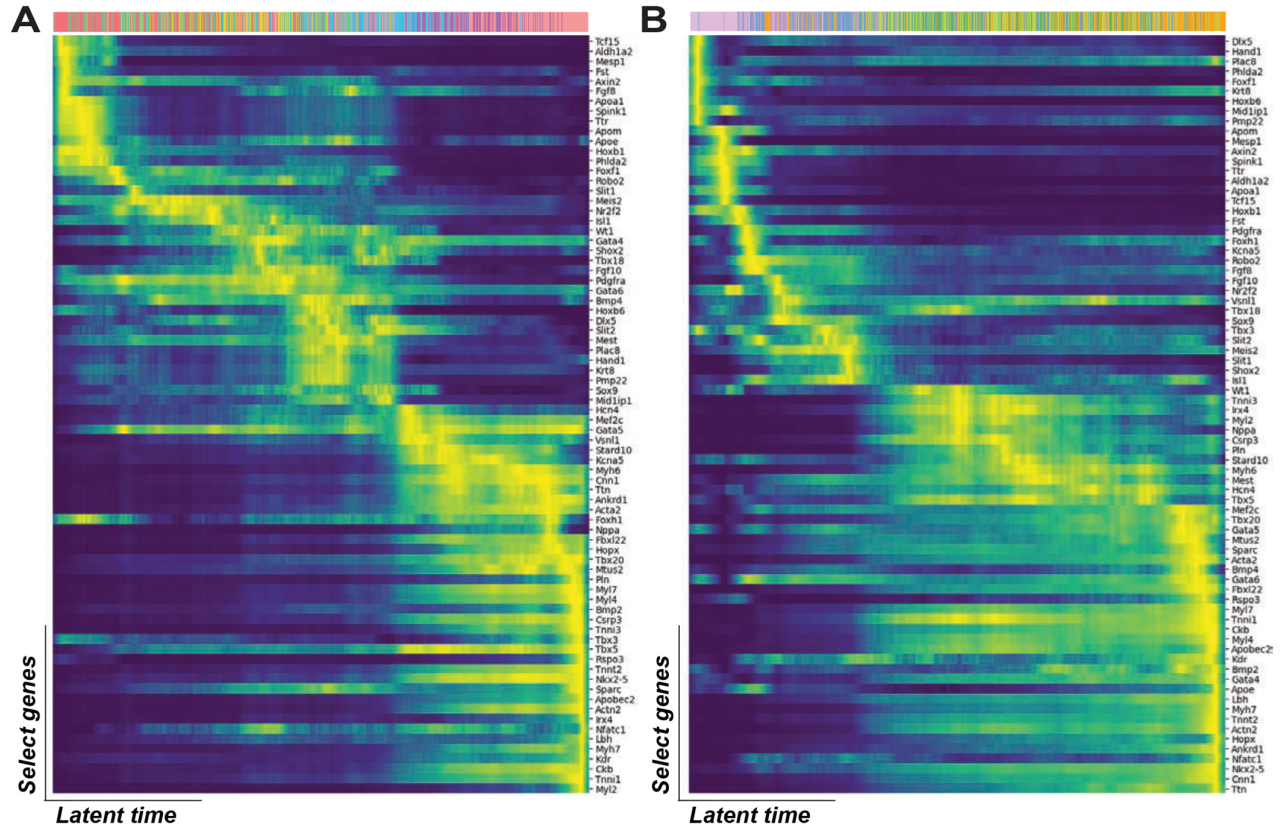
Fig. S6. Visualization of Dynamic Changes in Gene Expression Across LPM/ Heart Field Progenitor Differentiation Streams

A) Heatmap of selected genes of interest plotted against latent time for differentiation trajectory arising from pSHF/LPM progenitors (Figure 3F-H). **B)** Heatmap for selected genes of interest plotted against latent time for differentiation trajectory arising from aSHF progenitors (Figure 3I-K). **C)** Phase portraits (top) and expression dynamics (bottom) for selected genes found to be dynamically regulated across LPM/pSHF differentiation. S = spliced transcript; u = unspliced transcript; t = time. **D)** Phase portraits (top) and expression dynamics (bottom) for selected genes found to be dynamically regulated across aSHF differentiation.
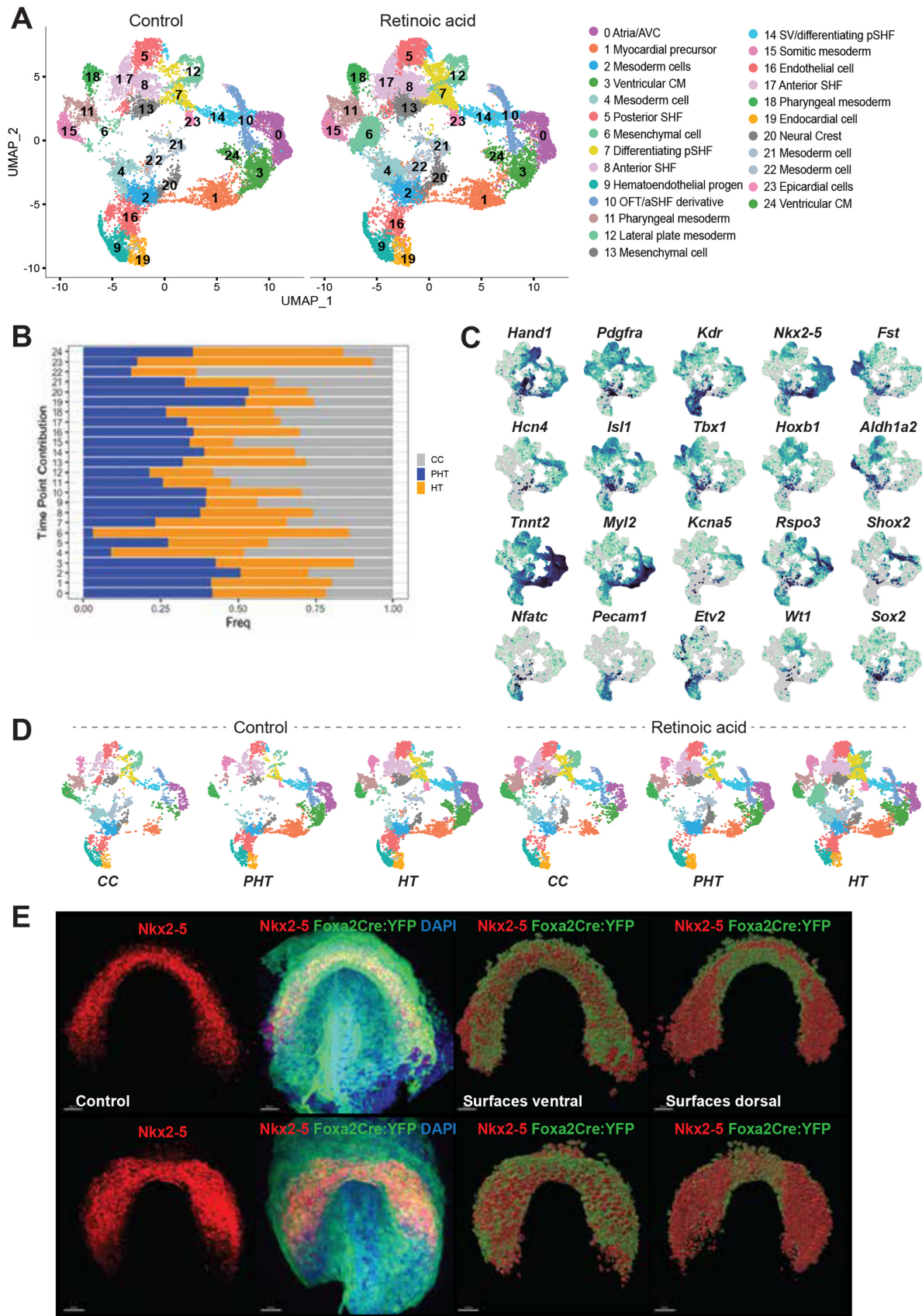
Fig. S7. Contribution of Control and RA-exposed Embryos to Cluster of Merged Data and Whole Mount Immunofluorescence of Cardiac Crescent Samples

A) UMAP projection of Normal/RA merged sample, split by condition (Normal and RA-exposed).

B) Quantification of overall contribution of cells from any one particular embryonic time point across Normal and RA treated samples. Frequency is calculated as the number of cells in a given cluster relative to the total number of cells in the sample. *Bottom***:** Sample contribution to UMAP projection. **C)** Feature plots for selected candidates of interest on merged Control and RA-treated samples. **D)** UMAP showing contribution from individual time points from normal and RA-exposed samples to combined cardiac UMAP projection. **E)** Representative images of whole mount immunofluorescence used for quantification of *Foxa2*-lineage traced cells within the cardiac crescent. *Left to right:* Single channel Nkx2-5 immunofluorescence; overlay of Nkx2-5 immunofluorescence and YFP (*Foxa2Cre;R26R:YFP* lineage-traced); Ventral view of surface volume construction of Nkx2-5 and *Foxa2* lineage-traced volumes; Dorsal view of surface volume construction.

Table S1. Ranked List of Most Dynamic Genes Per Cluster

Clusters are organized in columns from left to right, cluster ID is indicated in row 1.
Genes are organized in descending order, with most dynamical genes at the top of each
column.

Click here to download Table S1

**Table S2. List of Most Dynamic Genes for pSHF/LPM Derived
Differentiation Stream**

Genes are listed in second column, organized in descending order for the top 300 genes
across differentiation stream.

Click here to download Table S2

Table S3. List of Most Dynamic Genes for aSHF Derived Differentiation Stream

Genes are listed in second column, organized in descending order for the top 300 genes
across differentiation stream.

Click here to download Table S3

**Table S4. Differential Expression Results of Comparison Between
Differentiating aSHF (Cluster 8) and LPM/pSHF (Cluster 14).**

Results of differentiation expression testing performed in Seurat (negative binomial)
between Cluster 8 and Cluster 14. Results are organized in descending order by
adjusted p-value. pct.1 refers to the percentage of cells in test population #1 (Cluster 8)
expressing each gene compared to test population #2 (Cluster 14).

Click here to download Table S4

Table S5. Differential Expression Results of Comparison Between Ventricular (Clusters 2&17) and Atrial (Cluster 4) Cells.

Results of differentiation expression testing performed in Seurat (negative binomial) between atrial and ventricular cells. Results are organized in descending order by adjusted p-value. pct.1 refers to the percentage of cells in test population #1 (Cluster 2&17) expressing each gene compared to test population #2 (Cluster 4).

Click here to download Table S5

Table S6. Differential Expression Results of Comparison Between *EGFP+* and *EGFP-* cells in Cardiac Crescent.

Results of differentiation expression testing performed in Seurat (negative binomial) between EGFP+/- cells within CC sample. Results are organized in descending order by adjusted p-value. pct.1 refers to the percentage of cells in test population #1 (*EGFP* +) expressing each gene compared to test population #2 (*EGFP-*).

Click here to download Table S6