Resource

Developmental Cell

Single-Cell Resolution of Temporal Gene Expression during Heart Development

Highlights

- Single-cell RNA-seq data characterizing >1,200 murine cells from E9.5–P21 hearts
- Dynamic spatiotemporal gene expression defines distinct cardiomyocyte populations
- Human/mouse stem-cell-derived cardiomyocytes are developmentally immature
- *Nkx2.5*$^{+/-}$ mice have lineage-specific maturation defects in cardiac cells

Authors

Daniel M. DeLaughter, Alexander G. Bick, Hiroko Wakimoto, ..., Benoit G. Bruneau, J.G. Seidman, Christine E. Seidman

Correspondence

cseidman@genetics.med.harvard.edu

In Brief

Using spatiotemporal RNA-seq analyses of single cells isolated from E9.5–P21 mouse hearts, DeLaughter, Bick et al. reveal the dynamic transcriptional programs directing cardiomyocyte maturation during heart development. These data provide benchmarks for assessing lineage-specific maturation of differentiated stem cells and a basis for interrogating how human mutations cause congenital heart malformations.

DeLaughter et al., 2016, Developmental Cell *39*, 480–490
November 21, 2016
http://dx.doi.org/10.1016/j.devcel.2016.10.001

CellPress

CellPress

Developmental Cell
Resource

Single-Cell Resolution of Temporal Gene Expression during Heart Development

Daniel M. DeLaughter,[1,9] Alexander G. Bick,[1,9] Hiroko Wakimoto,[1] David McKean,[1] Joshua M. Gorham,[1] Irfan S. Kathiriya,[2,3] John T. Hinson,[4] Jason Homsy,[1] Jesse Gray,[1] William Pu,[5,6] Benoit G. Bruneau,[2,7] J.G. Seidman,[1] and Christine E. Seidman[1,8,10,*]

[1]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
[2]Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158, USA
[3]Department of Anesthesia and Perioperative Care, University of California, San Francisco, San Francisco, CA 92868, USA
[4]Division of Cardiovascular Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA
[5]Department of Cardiology, Harvard Medical School, Boston Children's Hospital, Boston, MA 02115, USA
[6]Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA
[7]Department of Pediatrics, Cardiovascular Research Institute, University of California, San Francisco, San Francisco, CA 94158, USA
[8]Cardiovascular Division, Howard Hughes Medical Institute, Brigham and Women's Hospital, Boston, MA 02115, USA
[9]Co-first author
[10]Lead Contact
*Correspondence: cseidman@genetics.med.harvard.edu
http://dx.doi.org/10.1016/j.devcel.2016.10.001

SUMMARY

Activation of complex molecular programs in specific cell lineages governs mammalian heart development, from a primordial linear tube to a four-chamber organ. To characterize lineage-specific, spatiotemporal developmental programs, we performed single-cell RNA sequencing of >1,200 murine cells isolated at seven time points spanning embryonic day 9.5 (primordial heart tube) to postnatal day 21 (mature heart). Using unbiased transcriptional data, we classified cardiomyocytes, endothelial cells, and fibroblast-enriched cells, thus identifying markers for temporal and chamber-specific developmental programs. By harnessing these datasets, we defined developmental ages of human and mouse pluripotent stem-cell-derived cardiomyocytes and characterized lineage-specific maturation defects in hearts of mice with heterozygous mutations in *Nkx2.5* that cause human heart malformations. This spatiotemporal transcriptome analysis of heart development reveals lineage-specific gene programs underlying normal cardiac development and congenital heart disease.

INTRODUCTION

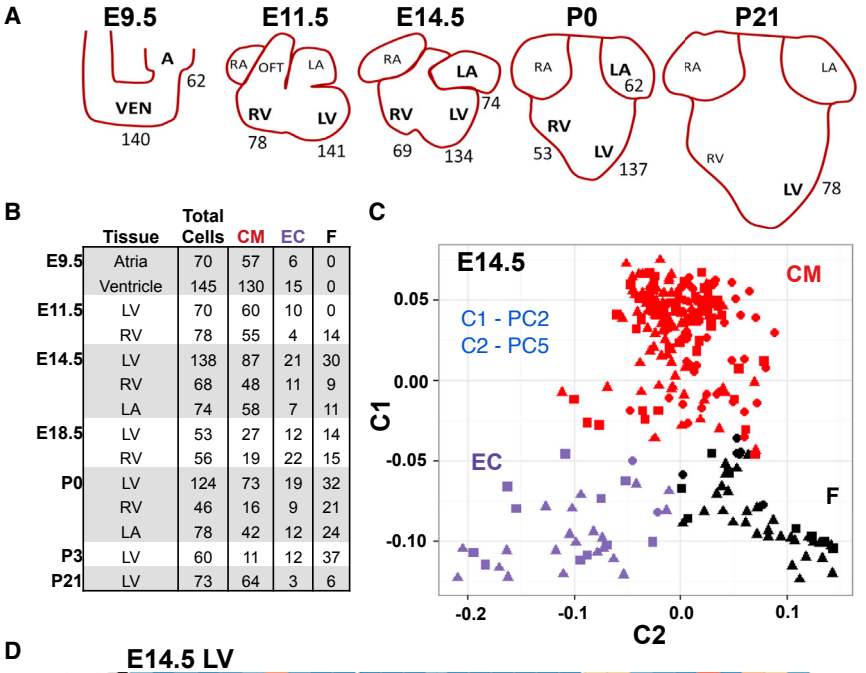
Formation and maturation of the mammalian heart is governed by signaling interactions among distinct cell lineages during development that transform a linear heart tube into a four-chamber heart. Transcriptional profiles of human, mouse, and chick hearts have identified dynamic changes in RNA expression throughout cardiogenesis Jensen et al. (2013); (Wagner and Siddiqui, 2007). Understanding how these relate to cellular differentiation and maturation remains incompletely understood, in part because of the inability of conventional transcriptome analysis experiments to deconvolute the multiple cell populations within the heart. Recent advances in single-cell transcription profiling approaches enable rapid capture and profiling of hundreds of cells (Buettner et al., 2015; Streets and Huang, 2014; Trapnell et al., 2014; Treutlein et al., 2014) that allow separation and characterization of lineages and cell lineage subtypes, independent of preexisting cell markers. Longitudinal assessments of transcriptional maps provide information that can be assembled into a transcriptional atlas across development or in response to disease.

To investigate cellular heterogeneity during cardiogenesis, we characterized the developmental temporal and spatial transcriptomes of heart development at single-cell resolution. Wild-type (WT) murine cells isolated from the left atrium (LA), primordial ventricle, and subsequently left (LV) and right (RV) ventricles at time points spanning embryonic to postnatal cardiac development were studied using single-cell RNA sequencing (RNA-seq). Cells were classified using unbiased transcriptome-wide analysis without a priori selection of cell-specific markers. We identified transcripts that delineated atrial and ventricular cardiomyocytes (CMs), and endothelial cell (EC) and fibroblast-enriched cell lineages. We also characterized CMs derived from stem cells and CMs and ECs derived from a congenital heart disease mouse model with haploinsufficiency of *Nkx2.5*, a critical transcription factor expressed early in cardiogenesis (Ashraf et al., 2014; Biben et al., 2000; Bruneau et al., 2001).

Single-cell transcriptional data of distinct cardiac lineages provided insights into lineage-specific markers and subpopulations that contribute to normal and aberrant cardiac development and provided a robust benchmark from which to assess the maturity of stem-cell-derived CMs.

RESULTS

We captured mouse (129SV) cardiac cells during important embryonic developmental milestones (Figure 1A): embryonic day



B

	Tissue	Total Cells	CM	EC	F
E9.5	Atria	70	57	6	0
	Ventricle	145	130	15	0
E11.5	LV	70	60	10	0
	RV	78	55	4	14
E14.5	LV	138	87	21	30
	RV	68	48	11	9
	LA	74	58	7	11
E18.5	LV	53	27	12	14
	RV	56	19	22	15
P0	LV	124	73	19	32
	RV	46	16	9	21
	LA	78	42	12	24
P3	LV	60	11	12	37
P21	LV	73	64	3	6

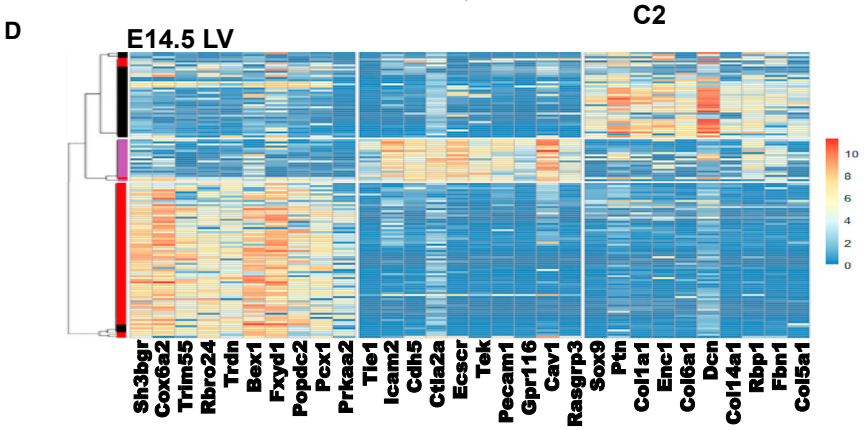


Figure 1. Single-Cell RNA-Seq of CMs, ECs, and Fibroblast-Enriched Cells in the Developing Heart

(A) Schematic representation of developing heart from which single cells were harvested (denoted by bold lettering). Numbers indicate numbers of captured wild-type cells. A, atrium; LA, left atrium; LV, left ventricle; OFT, outflow tract; RA, right atrium; RV, right ventricle; VEN, ventricle.

(B) The number single cells captured per tissue and time point and their lineage as determined by PCA. CM, cardiomyocytes; EC, endothelial cells; F, fibroblast-enriched cells.

(C) PCA components C1 and C2 separate E14.5 cells into three subgroups, each that robustly expressed genes associated with CMs (red), ECs (E, purple), or Fs (black). The percentage of variance explained by each principal component is provided in Table S2. pc2 - 3.3%; pc5 - 1.1%

(D) Unsupervised clustering (SC3) of E14.5 cells separates CMs, ECs, and fibroblast-enriched cells. The dendrogram was generated using consensus clustering with k = 3. The heatmap depicts the top ten most significant genes with enriched expression in each cluster. See also Figures S1 and S4; Tables S1, S2, and S3.

9.5 (E9.5) during heart looping, when the atria and ventricle become morphologically distinct; E11.5, when left and right cardiac chambers form; E14.5, when septation occurs with development of compact and trabecular layers of myocardium; and E18.5, when the heart undergoes metabolic remodeling. Cells were also captured at birth (postnatal day 0 [P0]) when the left atria and ventricle experience hemodynamic changes due to higher systemic pressures, and at P3, P7, and P21, when the heart transitions from proliferative growth (P3) to hypertrophic growth (P21).

The common atrium and ventricle (E9.5) were carefully dissected away from the adjacent atrioventricular canal or outflow tract, and at E14.5 to P21 the free wall was dissected from the LV and RV to exclude any valve or septal tissue. Tissues were digested into single-cell suspensions and isolated in a microfluidic device, after which single-cell RNA-seq analyses were performed. To assess technical variation and the dynamic range of transcript data, we compared and observed a good correlation of data derived from pooled single-cell RNA-seq and whole LV RNA-seq (Figure S1). Consistent with studies of single cells from other tissues, we found high single-transcript sensitivity and detected transcripts in a broad dynamic range (~10^5) between the most and least abundant transcripts.

Unbiased Transcriptome-wide Identification of Cell Populations

We studied more than 1,200 single cells isolated at different stages and chambers of murine heart development (Figures 1A and 1B; Table S1). To demonstrate that single-cell RNA-seq could distinguish distinct cell lineages in the heart, we employed principal component analysis (PCA) of all single cells from LV, RV, and LA tissues captured at E14.5 (261 cells) when the four-chambered heart is formed. Informative PCA components (Table S2) clustered cells into three groups (Figure 1C). Examining transcripts that were enriched in each of the three groups revealed that this unbiased clustering method faithfully identified previous markers known to be expressed in CMs (Figure 1C, red), ECs (Figure 1C, purple), or fibroblast-enriched cells (Figure 1C, black, e.g., non-CM, non-ECs). For example, among the top 30 genes that explain the largest proportion of the variation in PCA components (Table S2; see Experimental Procedures), CMs contained 11 sarcomere protein genes (e.g., *Myh6*, *Myh7*, *Ttn*), ECs contained 14 genes encoding cell surface markers (e.g., *Pecam1*, *Cd93*), and fibroblast-enriched cells contained nine genes encoding extracellular matrix (ECM) components (e.g., *Col1a2*, *Col1a1*).

To independently confirm cell classifications, we conducted unsupervised clustering using the Single-Cell Consensus Clustering (SC3) tool (Kiselev et al., 2016). SC3 uses a consensus clustering approach that integrates several methods for clustering of single-cell RNA-seq data to generate user-defined

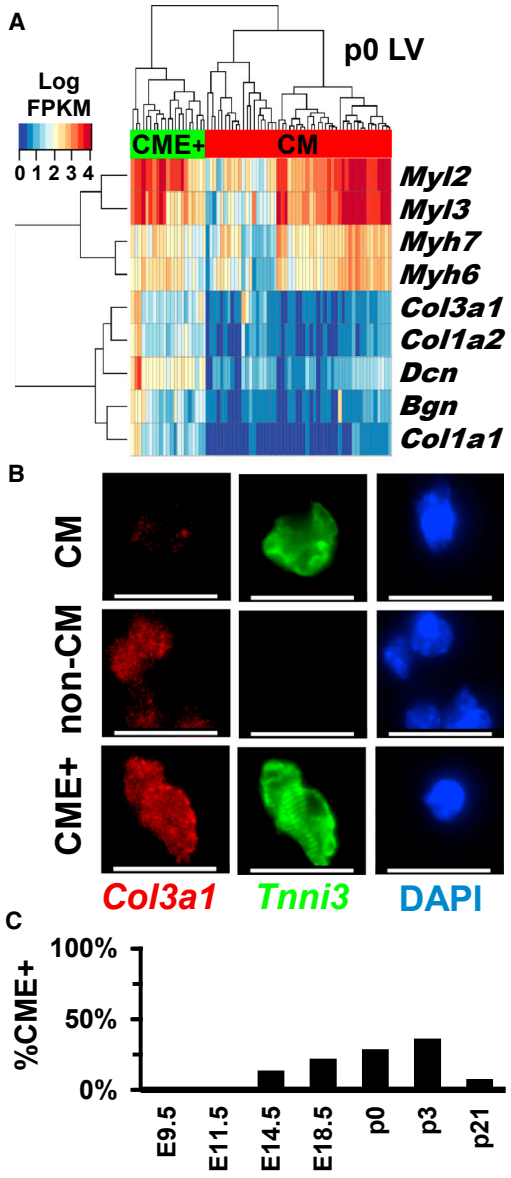


Figure 2. Subpopulations of CMs with Distinct Gene Expression Profiles

(A) Hierarchical clustering distinguishes CMs (red) from CMs expressing extracellular matrix genes (CME+, green) isolated from the LV at P0.

(B) Micrographs of P3 LV cells stained with *Tnni3* (green), *Col3* (red), or DAPI (blue). Scale bars, 20 μm.

(C) The proportion of LV CMs and CME+ cells varies during development.

See also Figures S1 and S2.

numbers of clusters (k). Assessments of variable numbers of clusters demonstrated that k = 3 and k = 6 maximized the silhouette width metric, which quantifies the ratio of each cell's transcriptional similarities with other co-clustered cells to the transcriptional differences with cells outside of its cluster (Rousseeuw, 1987). This independent approach yielded cell clusters that closely mirrored the PCA method (Figure 1B) with enriched expression of genes associated with CMs, ECs, or fibroblast-enriched cells (Figure 1D).

We undertook "perfect marker" gene analysis of PCA-separated cells (Treutlein et al., 2014) to identify markers for CMs, ECs, and fibroblast-enriched cells (Table S3). This approach identifies genes with high expression in specified cells and low expression in all other cells. "Perfect marker" gene analyses of PCA-separated CMs identified eight sarcomere protein genes (Table S3, boldface) and *Sh3bgr*, which encodes an SH3 domain binding glutamate-rich protein encoded within the trisomy chromosome 21 region that contributes to congenital heart defects. Morpholino-mediated knockdown of *Sh3bgr* in *Xenopus* results in disrupted segmentation of somites and heart malformations (Jang et al., 2015). Eight "perfect marker" genes for ECs encoded prototypic molecules (*Pecam1*, *Cdh5*) and two additional genes: *Klhl4* and *Gpr116* (Table S3, italics). *Klhl4* encodes an actin-binding kelch domain protein with unknown functions that has been implicated in developmental anomalies (Braybrook et al., 2001; Cheroki et al., 2008). Notably, an automated in situ hybridization screen (Gene Expression Database; Diez-Roux et al., 2011; Smith et al., 2014) of the E14.5 heart also identified *Klhl4* in the endocardium. *Gpr116* encodes a highly conserved G-protein-coupled adhesion receptor that is widely expressed in vascular beds, which when depleted from ECs produced vascular leakage (Niaudet et al., 2015; Yang et al., 2013). Six of ten "perfect marker" genes identified in fibroblast-enriched cells encoded ECM proteins (Table S3, underlined).

We then evaluated the distribution and maturation of CMs, ECs, and fibroblast-enriched cells throughout development. At E9.5, no fibroblast-enriched cells were identified in atrium, ventricle, or outflow tract, but subsequently the proportion of CMs in these regions decreased, while fibroblast-enriched cells increased (Figure 1B). At P0 approximately 30% of all LV cells and 45% of RV cells were fibroblast-enriched cells, a finding that is consistent with prior reports that cardiac fibroblasts appear by E12.5 and increase in number through P1 (Ieda et al., 2009). The proportion of ECs remained relatively constant from E9.5 (10%) to P0 (15%). At P21, the larger chip size used to capture CMs excluded many smaller non-CMs (Table S1), including hematopoietic-derived cells that comprise 5%–10% of non-CM cells in the adult mouse heart (Pinto et al., 2016). These cells were also not captured in embryonic hearts, likely due to their low abundance.

Among developing ventricular CMs, RNA-seq profiles were indicative of chamber myocardium and included two distinct subpopulations. One subpopulation (identified through consensus clustering using SC3 k = 6) expressed transcripts associated with cell division (Figure S2A). Using hierarchical clustering of the expression of cell-cycle genes (e.g., *Prc1*, *Ccna2*, *Cdca8*, *Cdca3*, *Top2a*, *Ccnb2*, *Mki67*, *Ccnb1*), we identified the proportion of CMs with proliferative capacity during development (Figure S2B). Approximately 60% of E9.5 and E14.5 CMs, but only 20% of P0 and P7 CMs, expressed these cell-cycle transcripts (Figure S2B). At P21, this subpopulation was absent.

A second CM subpopulation was identified by hierarchical clustering (Figure 2A) of marker genes for CMs and fibroblast-enriched cells (identified in Figure 1C and listed in Table S2). This CM subpopulation expressed transcripts encoding ECM proteins (denoted CME+) that are typically associated with cardiac fibroblasts (e.g., *Col1a1*, *Col3a1*, *Dcn*). RaceID, which identifies rare cell populations in complex data (Grun et al., 2015), also

defined a cluster of cells that expressed both sarcomere components and ECM (Figure S2C: cluster 7). To determine whether these cells were more similar to either CMs or fibroblast-enriched cells, we averaged gene expressions for all CMs, fibroblast-enriched cells, or CME+ at E14.5, E18.5, or P0 (Figure S1C). Comparisons of the mean gene expression values and Pearson correlations indicated that the average transcriptome of CME+ cells were more similar to CMs (E14.5, r = 0.77; E18.5, r = 0.96; P0, r = 0.88) than fibroblast-enriched cells (E14.5, r = 0.61; E18.5, r = 0.66; P0, r = 0.43). To independently confirm cell identities, we studied preparations of single-cell suspensions from P0 and P3 LV tissues using immunohistochemistry (Figures 2B and S2D) to detect the sarcomere protein troponin I (Tnni3) and collagen (Col3a1). Cell diameters (long axis) were compared (n = 145) from CMs and non-CMs expressing Tnni3 and/or Col3a1. In the P3 LV, 47% of CMs expressed Col3a1, similar to the proportion observed in RNA-seq data (36%). CMs and CME+ cells had a similar mean length (17.8 ± 3.7 versus 17.5 ± 3.3 μm) compared with all non-CMs (10.0 ± 3.2 μm) and non-CMs that expressed Col3a1 (8.2 ± 2.3 μm) (Figure S2E). Based on the comparable cell sizes, organized sarcomeres, and expression of Col3a1 and other ECM proteins, we deduced that CME+ cells are a subpopulation of CMs, and not fibroblasts that express sarcomere components, such as myofibroblasts (Mayer and Leinwand, 1997). Throughout LV development, CME+ cells comprised variable proportions of all CMs (Figure 2C): 9% at E14.5, 22% at E18.5, 29% at P0, 36% at P3, and 1% at P21.

Developmental Transcription Profiles of Non-CM Cells

Despite increased numbers of fibroblast-enriched cells throughout cardiac development, we observed minimal changes in the transcriptional profiles of non-CMs from E9.5 to P21 (Figure S3A), possibly due to limited numbers of captured cells (Figure 1B). By contrast, the transcriptional profiles of ventricular ECs (Figure S3B) showed a distinct temporal progression with three stages of evolution. During heart looping and chamber formation (E9.5, E11.5), unsupervised clustering using SC3 followed by gene-enrichment analysis (Figure S3C) identified high expression levels of *Hapln1* (hyaluronan and proteoglycan link protein 1) and *Hmga1* (high-mobility group AT-hook protein). *Hapln1 is* an important regulator of developmental ECM interactions (DeLaughter et al., 2013) while *Hmga1* regulates chromatin organization and promotes endothelial-mesenchymal transformation (Benecke and Eilebrecht, 2015). Expression of these genes during chamber formation is consistent with matrix remodeling that participates in early ventricular trabeculation. Later in development (E14.5, E18.5), ECs showed increased expression of *Fabp4* (fatty acid binding protein 4), a molecule that contributes to cardiac metabolic shifts from glycolysis to oxidative phosphorylation (Vander Heiden et al., 2010). From P0 through P21 the EC transcriptional program remained unchanged.

Developmental Transcriptional Profiles of CMs

E9.5 CMs isolated from the common atria or the primordial ventricle had differential expression of 563 genes (Table S4), and PCA analyses identified gene subsets that reliably discriminated atrial from ventricular CMs (Figure S4A and Table S2). For example, E9.5 ventricular CMs expressed high levels of *Hey2* and *Pln* while E9.5 atrial CMs expressed high levels of *COUP-TFII* (Figure S4B). As these profiles remained distinct over time, we characterized the individual temporal transcriptional programs governing the maturation of atrial or ventricular CMs separately.

Atrial CMs

Maturation of atrial CMs resulted in both decreased and increased expression of genes from E9.5 to E14.5 and from E14.5 to P0 (Table S4). PCA suggested progressive transcriptional states (Figure 3A and Table S2) that we characterized by hierarchical clustering of the genes associated with each state. Three distinct gene expression patterns were identified (Figures 3B and 3C). Pattern 1 consisted of genes with binary expression (i.e., on or off) in an increasingly larger proportion of cells. This pattern included genes that regulate many aspects of adult atrial CM physiology, including cell-cell adhesion (*Gja1*) (Desplantez et al., 2012; Luo et al., 2007), calcium buffering (*Csq2*, encoding *calsequestrin*), and electrophysiologic conductance (*Kcnk3*; Liang et al., 2014). Pattern 2 identified genes that were expressed early and globally in atrial development but were silent at E14.5 and thereafter. These genes included *Wnt2*, a regulator of early atrial development (Tian et al., 2010), genetic markers of immature CMs (*Acta2*), and a regulator of chromatin remodeling that may promote early CM differentiation (*Hmga2*; Monzen et al., 2008). Pattern 3 genes included globally expressed genes with increased expression as development progressed. This group includes genes with fundamental CM function including sarcomere proteins (*Tnni3*).

Right versus Left Ventricular CMs

Before considering the developmental programs in maturing ventricular CMs, we used PCA to evaluate the transcriptomes of E11.5 CMs within the developing RV and LV chambers (Figure S4A and Table S2). We identified 279 genes significantly differentially expressed ($p < 1 \times 10^{-7}$) between LV and RV CMs using a Mann-Whitney-Wilcoxon rank-sum test (Table S5). Included in this subset are genes previously described as regulators of ventricular chamber formation (Figure S4C), including *Hand1* and *Tbx5*, which had restricted expression in LV CMs (Liberatore et al., 2000; Thomas et al., 1998). By contrast, Tbx5-repressed genes had higher expression in RV CMs, findings that support prior studies indicating that *Tbx5* is critical for the early establishment of chamber identity (Bruneau et al., 2001). Genes enriched in RV CMs also included proteins associated with the ECM, microfibrillar-associated protein 2 (*Mfap2*), which binds active Tgfβ-1, Tgfβ-2, and Bmp2 (Combs et al., 2013), and a cell adhesion, melanoma cell adhesion molecule (*Mcam*) that marks neural crest-derived lineages (Medic and Ziman, 2010; Pujades et al., 2002).

Gene ontology and KEGG pathway analyses (Figure S4D) of 138 genes that were significantly enriched in RV CMs identified ECM organization (n = 11, $p = 2 \times 10^{-7}$), cell adhesion (n = 19, $p = 1.2 \times 10^{-6}$), and glycolysis (n = 7, $p = 1.9 \times 10^{-6}$) or enrichment in pathways involved in dilated cardiomyopathy (n = 7, $p = 7 \times 10^{-5}$) and arrhythmogenic RV cardiomyopathy (n = 7, $p = 1.8 \times 10^{-4}$). Although parallel analyses of 141 genes enriched in the LV CMs (Table S4) were less informative of specific pathways, we observed several genes that participate in processes implicated in congenital heart disease (e.g., retinoic acid signaling [Nash et al., 2015] and histone modification in

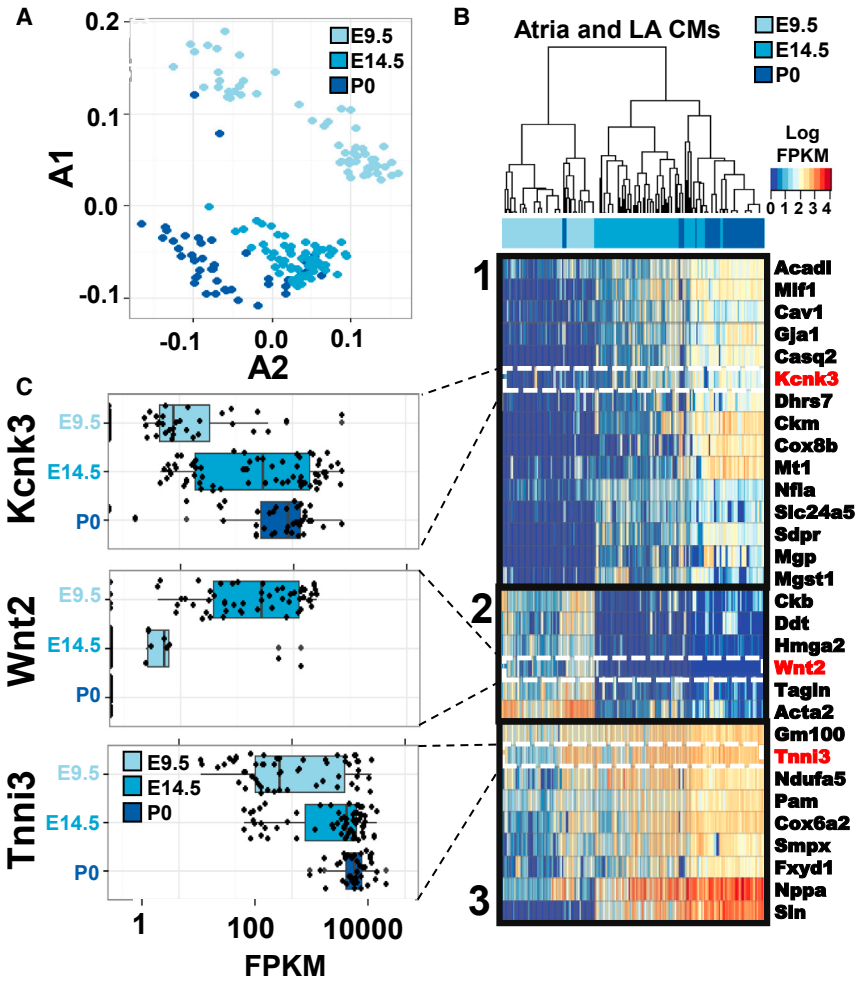


Figure 3. Distinct Temporal and Chamber Gene Expression Patterns in CMs

(A) PCA of gene expression in E9.5 common atria (light blue), E14.5 LA (blue), and P0 LA (dark blue) demonstrate that atrial CMs exhibited stepwise, temporal expression patterns. The percentage of variance explained by each principal component is provided in Table S2.

(B) Hierarchical clustering of E9.5 common atria (light blue), E14.5 LA (blue), and P0 LA (dark blue) CMs show three temporal gene expression programs.

(C) Representative genes depicting single-cell gene expression (black dots) from each of the three temporal atrial CM gene expression programs.

See also Figure S4; Tables S2 and S5.

response to Nodal signaling [Dahle et al., 2010; Homsy et al., 2015]). Notably, these right/left differences were not observed when comparing ECs or fibroblast-enriched cells. Thus, we suggest that developmental pathways that regulate left and right chamber ventricular formation are expressed primarily in CMs rather than in non-myocyte cells.

Between E14.5 and P0, LV and RV CMs became transcriptionally more similar to one another (Figures S4E–S4G and Table S2) and expressed genes that characterize CM maturation (e.g., *Myh6*, *Myl3*). Taking these data together, we propose that ventricular CMs exhibit distinct transcriptional programs underlying LV and RV chamber morphogenesis after which transcriptional convergence promotes CM maturation.

Developmental Transcriptional Programs in Ventricular CMs

Similarly to atrial myocytes, we observed a temporal evolution of ventricular CM gene expression (Figure 4A). PCA analyses of CMs isolated from four time points (E9.5, E14.5, P0, and P21) were undertaken using 997 genes expressed in at least 75% of CMs (Table S6). PCA component A2 (Figure 4A and Table S2) separated cells in overlapping, stepwise clusters ordered by age that were visualized using a density plot (Figure 4B). Among the genes that demarcated the transition from least to most mature ventricular CMs, ≥11 genes are involved in mitochondrial fatty acid β-oxidation (e.g., *Fabp3*, *Fapb4*, *Uqcr11*, *Cox6c*) and account for the developmental metabolic switch from embryonic glycolysis to mitochondrial fatty acid β-oxidation of adult CMs (Vander Heiden et al., 2010). Indeed, ventricular CMs expressed reciprocally increased and decreased transcript levels for *Fabp3* and *Fabp4*, *Myh7* and *Myh6*, and prenatal *Cox8a*, but postnatal *Cox8b* provided evidence of progressive maturation (Figure 4C). To examine whether CMs remained this distinct between all stages of cardiogenesis, we expanded the PCA to include three intermediate time points (E11.5, E18.5, and P3). From E9.5 to P21 (Figure 4B), ventricular CM transcriptomes exhibited a continuum of maturation, but also demarcation from the prior and subsequent stage. This resulted in a progressive wave of ventricular maturation, a developmental pattern that correlated with progressive cell-cycle exit (Figure S2A).

Within the developmental waves, density plots revealed three stages of CM maturation between E9.5 and P21 (Figure 4B). CMs on each extreme of this developmental axis are derived primarily from E9.5/E11.5 cells and from P21 cells. Between these extremes we observed a heterogeneous population of CMs comprising cells isolated from E14.5 to P3. This same clustering was also observed when unsupervised clustering using the SC3 tool was performed on ventricular myocytes (Figure S5A). To examine how gene expression changes between these stages of CM maturation, we performed hierarchical clustering on the 50 genes that explained most of the component A2 variance in Figure 4A. Cells ordered by their expression of component A2 (Figure 4D) also yielded three clusters with distinct patterns of gene expression. Parallel analyses of 400 genes confirmed these patterns (Figure S5B and Table S7). Cluster 1 consisted of genes whose expression began on or before E9.5 and increased steadily until P21, such as *Myl2*, *Myh6*, *Fabp3*, and *Cox6c* (Table S7). The second cluster was defined by genes whose expression

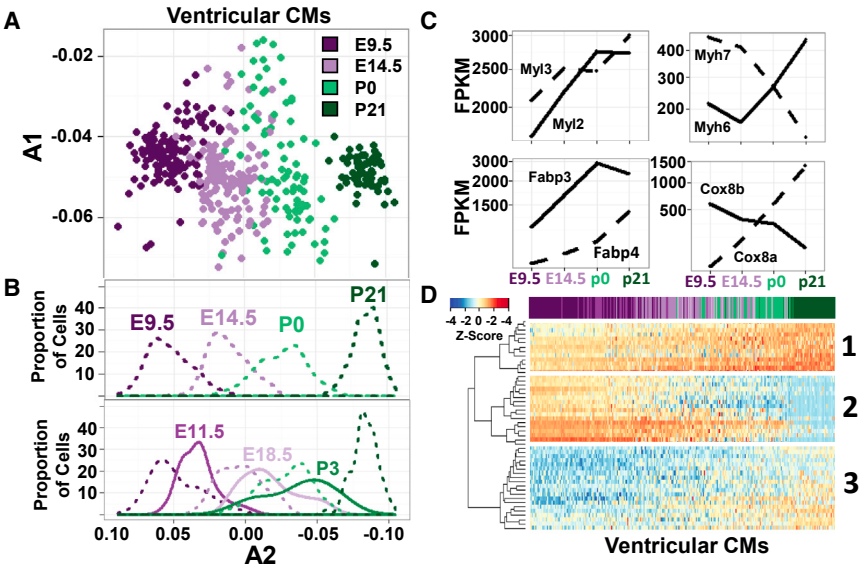


Figure 4. Temporal Gene Expression Profiles of Ventricular CMs

(A) PCA of ventricular CMs isolated at E9.5 (dark purple), E14.5 (light purple), P0 (light green), and P21 (dark green). The percentage of variance explained by each principal component is provided in Table S2.

(B) (Top) A density plot depicting the proportion of ventricular CMs isolated at four time points (E9.5, E14.5, P0, and P21) on component A2. These time points were evenly distributed along A2. (Bottom) Density plot depicting the proportion of cells with a given value of component A2 from PCA of CMs isolated at seven time points (E9.5, E11.5, E14.5, E18.5, P0, P3, and P21) show considerable overlap in temporal development from E14.5 to P3.

(C) Box plots depict mean fragments per kilobase of transcript per million mapped reads (FPKM) of genes with notable isoform switches in CMs isolated at E9.5, E14.5, P0, or P21.

(D) Hierarchical clustering of the top 50 genes in component A2 from (B). The genes clustered into three patterns (detailed in text). Columns represent individual single cells ordered by value of component A2 from (B) and colored by age at time of isolation. *Z* score reflects normalization by column. See also Figures S4 and S6; Tables S2, S6, and S7.

decreased as development proceeded, either at E14.5 (*Cdk1b* and *Pgam1* encoding phosphoglycerate mutase 1) or P21 (*Myl7* and *Hmgn2* encoding high-mobility group nucleosomal binding domain 2). Cluster 3 genes were lowly expressed early (E9.5–E11.5) but had increasing expression at E14.5 (*Oxct1* encoding succinyl-CoA:3-ketoacid-coenzyme A transferase 1, *Ryr2* encoding ryanodine receptor 2) or P21 (*S100a2* encoding S100 calcium binding protein A2). Thus, the genes observed in these three clusters implicate different biological processes critical in heart maturation: sarcomere structure and fatty acid oxidation in cluster 1, glycolysis and proliferation in cluster 2, and mitochondrial function and calcium handling in cluster 3.

The heterogeneity in maturity of individual CMs, isolated at the same time point, raised the question of whether whole-tissue samples at harvested at specified distinct embryonic times could be distinguished based on gene expression. To assess this, we processed whole LV free wall at E14.5 and E18.5 for RNA-seq, similar to single-cell methods. Analyses were restricted to 997 genes expressed at $\geq$1 FPKM (fragments per kilobase of transcript per million mapped reads) in 75% of single CMs harvested at E9.5, E11.5, E14.5, E18.5, P0, P3, and P21 (Table S6). A density plot was generated from the most informative temporal component (A2, Figure 5A). LV free wall tissue from E14.5 and E18.5 mice mapped to the middle of the distribution of single cells, confirming that the maturation profiles from many single cells accurately captured in vivo developmental transcriptional profiles and also confirmed that whole-tissue RNA-seq data averages the expression of CMs with different maturation stages.

Developmental Ages of Stem-Cell-Derived CMs

Applying the method described above (Figure 4B), we used published RNA-seq datasets that were processed through our pipeline to determine the developmental ages of mouse and human embryonic stem cells. Two replicate RNA-seq datasets from cardiac progenitors (mES-CP) and CMs (mES-CM) derived from murine embryonic stem cells (Wamstad et al., 2012) showed high concordance. Expression data from mES-CP corresponded to E9.5 single-cell CMs, while data from mES-CM cells corresponded to E14.5 cells (Figure 5A).

Similar analyses of human embryonic stem-cell-derived (hES) CMs were performed using 847 genes with human orthologs (Bick et al., 2012). RNA-seq datasets from hES-derived CMs cultured for 20 days (hES-D20) or approximately 1 (13.5 months) year (hES-Y1) (Kuppusamy et al., 2015) showed high concordance. The hES-D20 cells corresponded to E14.5 ventricular CMs while hES-Y1 cells corresponded to E18.5 ventricular CMs. To confirm that these conclusions did not reflect differences in mouse and human development, we evaluated RNA-seq data from fetal and neonatal human tissues, including two LV tissues from 8 to 12 weeks of gestation and five tissue samples from postnatal infant hearts. Human fetal heart samples corresponded to late E18.5, while neonatal samples corresponded to P0 to P3 (Figure 5B). Together, these findings support the conclusion by Kuppusamy et al. (2015) that prolonged culturing increases CM maturation, but also indicated that hES-Y1 CMs remain more immature than P21 cells (Figure 5B).

Nkx2.5 Haploinsufficiency Impairs Maturation of Distinct Cardiac Cell Lineages

Dominant mutations in *NKX2.5* cause human cardiac septation defects (Elliott et al., 2003; Schott et al., 1998; Yuan et al., 2015). *Nkx2.5*$^{+/-}$ mice recapitulate many of these cardiac phenotypes, albeit with incomplete penetrance (Ashraf et al., 2014; Biben et al., 2000; Bruneau et al., 2001). We hypothesized that haploinsufficiency of the transcription factor *Nkx2.5* causes heart malformations by altering developmental transcription in

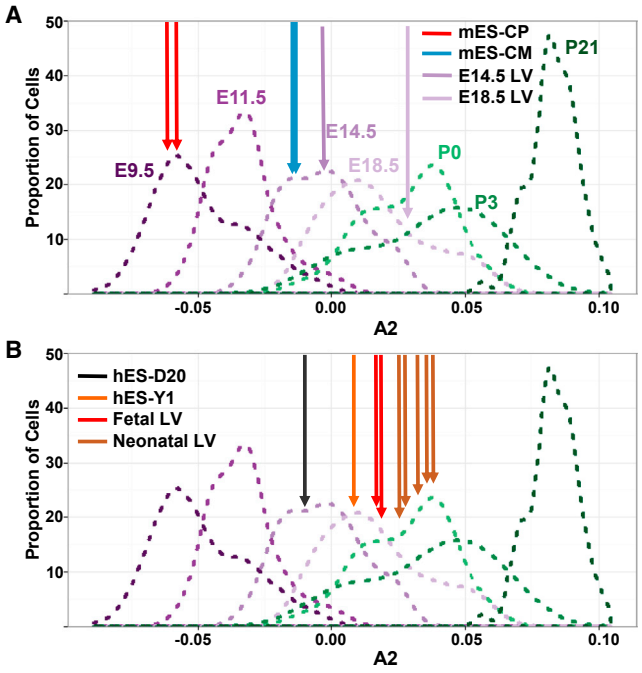


Figure 5. Developmental Ages of CMs Derived from Human and Mouse Stem Cell and LV Tissues, Depicted Using Density Plots as Described in Figure 4B

(A) Murine whole-tissue samples from E14.5 (purple arrow) and E18.5 LV (light purple arrow) have A2 values near the mean of the distribution of E14.5 and E18.5 LV single cells, respectively. Murine embryonic stem-cell-derived cardiac progenitors (mES-CP, red arrow) correspond to E9.5 CMs and murine embryonic stem-cell-derived CM lines (mES-CM, blue arrow) correspond to E14.5 CMs. mES transcriptome data are from two independent clonal lines (Wamstad et al., 2012).

(B) Similar analyses of transcriptional analyses from pooled human embryonic stem-cell-derived CMs (Kuppusamy et al., 2015) cultured for 20 days (hES-D20, black arrow) or 1 year (hES-Y1, orange arrow) indicate that these respectively correspond to mouse E14.5 and E18.5 CMs. The transcriptome of the free wall of human fetal LV (ages 8 and 12 weeks, red arrow, n = 2) and human neonatal LV or RV (brown arrow; n = 5) are shown for comparison. See also Table S6.

specific cardiac lineages. To test this model, we isolated cells from the LV at E14.5, P0, P7, and P21 in *Nkx2.5*$^{+/-}$ mice and compared PCA of single-cell RNA-seq data with that from WT mice (Table S4).

Nkx2.5$^{+/-}$ CMs expressed lower levels of *Myh7* (−5.0 fold, p = 2.3×10^{-11} compared with WT) implying a potential defect in CM maturation. We therefore compared the maturity of *Nkx2.5*$^{+/-}$ cardiac cell lineages using single-cell transcriptomes from mutant and WT CMs isolated at E14.5, P0, P7, and P21 (Figure 6A). To better visualize maturation differences, we compared density plots of cell distributions in mutant (solid waves) and WT (dashed waves) mice derived from PCA components A2 (Figure 6B). *Nkx2.5*$^{+/-}$ and WT CMs isolated at P0 or P21 had similar developmental distributions, but E14.5 *Nkx2.5*$^{+/-}$ CMs were skewed toward more immature CM populations. Comparing the mean FPKM (Student's t test), 420 genes were differentially expressed between *Nkx2.5*$^{+/-}$ and WT CMs, of which several were associated with myocyte maturation including *Ttn* and *Myh6* (Table S4).

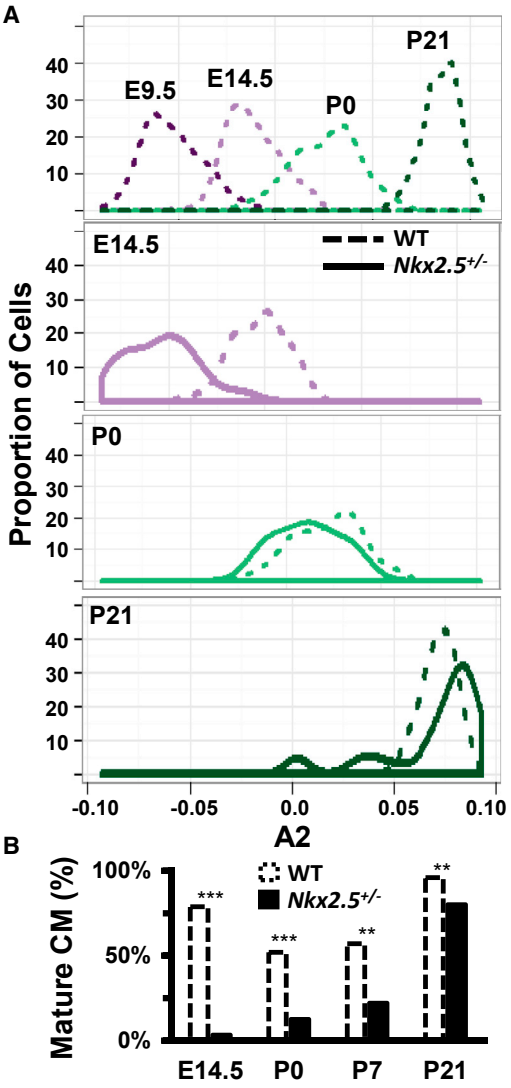


Figure 6. Maturation Defects in *Nkx2.5*$^{+/-}$ CMs

(A) Developmental profile of E9.5 VEN and E14.5, P0, and P21 LV CMs from wild-type (WT) (dashed line) and *Nkx2.5*$^{+/-}$ mice (solid line) depicted using density plots as described in Figure 4B. At E14.5 (second panel), *Nkx2.5*$^{+/-}$ CMs are less mature than E14.5 WT cells. This immaturity decreases gradually from P0 (third panel) to P21 (fourth panel).

(B) Using hierarchical clustering, the proportion of *Nkx2.5*$^{+/-}$ CMs (black boxes) that cluster with similarly aged WT CMs (dashed boxes) is significantly different at all time points (Fisher's exact p values: E14.5 ***p = 2.1×10^{-25}, P0 ***p = 0.0012, P7 **p = 0.010, P21 **p = 0.0071).

See also Figure S5 and Table S4.

To more precisely quantitate the maturity defect in *Nkx2.5*$^{+/-}$ CMs, we defined the top 20 genes contained in PCA component A2 that distinguished E14.5 mutant from WT CMs (Figure 6A and Table S2) and used these genes for hierarchical clustering of CMs. Two populations of WT CMs were identified: cells that were more mature (52%, 10/19) or less mature (48%, 9/19) based on gene expression (*Myl2*, *Myl3*, *Fabp3*; Figure 6B). At P7 (60%; 12/20) and P21 (97%; 62/64), increasing numbers of WT CMs were more mature.

In contrast, P0 *Nkx2.5*$^{+/-}$ CMs clustered predominantly as less mature (87%; 13/15) rather than more mature (2/15) cells (compared with WT, p = 0.012; Fisher's exact test). P7 gene expression profiles of *Nkx2.5*$^{+/-}$ CMs were more comparable with P0 WT CMs and only 20% of CMs exhibited mature gene expression patterns, in comparison with 60% of WT CMs (p = 0.01) (Figures 6B and S5C). At P21 most (78%, 21/27) of *Nkx2.5*$^{+/-}$ CMs were mature, albeit a significantly smaller proportion than in WT LV (97% 62/64, p = 0.0071, Fisher's exact test) (Figure 6B).

Prior studies demonstrate *Nkx2.5* expression principally in CMs (Kasahara et al., 1998), which we confirmed by single-cell analyses (Figure S5D). As recent evidence indicates developmental crosstalk between CM and EC lineages (Palencia-Desai et al., 2015), we also assessed the maturation of ECs from *Nkx2.5*$^{+/-}$ mice (Figure S3D). Transcriptional profiles of *Nkx2.5*$^{+/-}$ ECs were comparable with WT ECs at E14.5, but significantly delayed at P0 (*Nkx2.5*$^{+/-}$ 12/14 immature ECs versus WT 4/19; p = 0.038, Fisher's exact test).

Together, these data indicate that *Nkx2.5*$^{+/-}$ haploinsufficiency delays maturation in both CMs and ECs and provide further support that the myocardium is crucial for endocardial differentiation.

DISCUSSION

Using single-cell RNA-seq, we performed a high-resolution time-course analysis of cardiac development. Our efforts leveraged unbiased, single-cell analyses to separate and characterize more than 1,200 cardiac cells from three distinct lineages. In addition to characterizing longitudinal developmental programs, we identified distinct transcriptional profiles for CM subpopulations and chamber-specific profiles. Leveraging these analyses, we established developmental ages of CMs derived from stem cells, and characterized lineage-specific consequences of haploinsufficiency of the transcription factor *Nkx2-5*.

We identified considerable transcriptional heterogeneity among CMs. Among E14.5 ventricular CMs, a subpopulation (CME+) expressed ECM proteins that are emblematic of fibroblasts, even though their overall transcriptome typified CMs. The proportion of this CM subpopulation increased until P3, after which CME+ cells were rarely identified. We speculate that CME+ cells provide a developmental scaffold for the formation and maintenance of heart architecture and may be disease responsive, thereby accounting for increased cardiomyocyte expression of ECM proteins (Ambrosino et al., 2006) due to increased proportions of CME+ cells rather than global transcriptional changes in all CMs.

Another CM subpopulation was distinguished by the expression of cell-cycle genes, which we presumed implies active proliferation. The transcriptional profiles of these cells were not suggestive of stem cells or recently described cardiomyoblasts (Jain et al., 2015); earlier developmental time points are likely needed to capture these cells. Instead, cell-cycle gene expression was found in 60% of E9.5 CMs. With developmental progression, these transcripts rapidly declined. At P21 no CMs expressed cell-cycle genes.

Notably, the transcriptomes of proliferating and non-proliferating CMs were remarkably similar, indicating that these cells had comparable developmental maturity. Consistent with this, genes that demarcated distinct waves of development by CMs did not include cell-cycle genes. We suggest that throughout embryogenesis newly born CMs "kept pace" with differentiating cells, perhaps due to a rich milieu of intrinsic and secreted factors that promoted maturation. Identification of these molecules could be important for accelerating in vitro maturation of stem-cell-derived CMs.

Ventricular CMs exhibited overlapping, stepwise developmental patterns that we grouped into three transcriptional stages of maturation. At E9.5–E11.5, CMs expressed genes involved in proliferation (e.g., cell-cycle genes, Figure S2A) and morphogenesis (e.g., *Sall4*, a *Tbx5*-regulated transcription factor that promotes ventricular morphogenesis and chamber maturation; Koshiba-Takeuchi et al., 2006). CMs subsequently matured through sequential overlapping waves of gene expression (Figure 4B) between E14.5 and P3. These transitional CMs expressed genes involved in cellular growth (e.g., *Tcea1*, encoding transcription elongation factor), CM differentiation (*Cbx3*, an epigenetic silencer that promotes pluripotent cells differentiation; *Casq2*, calsequestrin 2; *Ckm*, creatinine kinase), and the metabolic switch from glycolysis (reduced *Pgam1* expression) to fatty acid oxidation (increased *Fabp3*, *Fabp4*, encoding fatty acid binding protein; Figure 4C) that characterizes mature CMs (Vander Heiden et al., 2010). Further maturation occurred by P21 when there was decreased expression of calmodulin-interacting proteins, *Bex1* and *Bex4*, which promote muscle regeneration (Koo et al., 2007).

Using these maturation stages we characterized stem-cell-derived CMs, which are advancing the understanding of human heart disease and propelling the development of new therapeutics, despite the limitation that these cells fail to recapitulate adult CM gene expression even after a year in culture (Kuppusamy et al., 2015). While a few markers have been previously used to gauge the maturation of stem-cell-derived CMs (Li et al., 2015), our spatiotemporal CM transcriptional profile of single cells and in vivo mouse and human tissues precisely defined the developmental "age" of stem cells and derived CMs (Figure 5). We observed remarkable homogeneity and consistent immaturity ($\leq$E18.5) in both mouse and human CMs derived from stem cells. Knowledge of the repertoire of genes expressed through development should advance chemical or genomic strategies to improve both the maturation and subpopulation diversity of stem-cell-derived CMs and cardiac lineages.

Single-cell-derived transcriptome maps also uncovered lineage-specific perturbations due to haploinsufficiency of the cardiac transcription factor *Nkx2.5*$^{+/-}$, a cause of human congenital heart disease. Our data extend earlier transcriptional analyses of *Nkx2.5*$^{+/-}$ mouse heart tissues (Bouveret et al., 2015) by delineating lineage-specific and temporal changes that persisted late into development. In comparison with WT cells, we demonstrate developmental immaturity of E14.5 *Nkx2.5*$^{+/-}$ LV CMs and P0 ECs (Figures 6 and S3D). As *Nkx2.5* is not expressed outside of CMs, we suggest that the dysregulated EC gene expression in *Nkx2.5*$^{+/-}$ hearts indicates non-autonomous signals that are evoked by haploinsufficiency of this transcription factor.

In conclusion, we demonstrate that single-cell RNA expression analyses improved delineation of lineage-specific and spatiotemporal developmental transcriptional programs. Expansion

of these analyses to include earlier developmental time points and additional lineages is expected to expand insights into molecular pathways that are critical for heart development.

EXPERIMENTAL PROCEDURES

Mice and Dissection

All animal experiments were approved by the Institutional Animal Care and Use Committee of Harvard Medical School. Pregnant female 129SV mice (Taconic) were used for tissue dissection. The construction of *Nkx2.5*-del (Tanaka et al., 1999) mice was previously described. *Nkx2.5* heterozygous males from a mixed background were crossed with 129SV females. Embryonic age was confirmed by ascertaining the Thieler stage of each pup. Tissue derived from one pup was used for each individual chip. Hearts were excised into PBS and incubated with agitation at 37°C for 20–60 min, depending on stage, in digestion buffer consisting of PBS, 0.70% (w/v) collagenase B, and 0.75% (w/v) collagenase D (Roche).

Cell Capture

Fluidigm integrated fluidic circuits (IFCs) were used to capture single cells. After capture, the IFC lysed individual cells, reverse-transcribed RNAs, and amplified complementary DNAs. The resulting cDNAs from individual cells were barcoded, indexed, and sequenced to an average depth of 1.5 million reads per library. IFCs were selected to capture all major cell populations from all cell size ranges observed using IFCs which capture cells of different sizes: 5–10 μm (embryonic), 10–17 μm (embryonic, neonatal), and 17–25 μm (3 weeks of age). No batch effects were observed between chips of the same size. Captured cells were imaged on chip to confirm the number of cells per site, and the viability of the cell was confirmed using a LIVE/DEAD cell assay (Life Technologies). Only single, viable cells were used for further analysis. Onboard cell lyses, in vitro transcription, and cDNA synthesis were performed using an IFC RNA-seq reagent kit (Fluidigm) and SMARTer Ultra Low RNA kit (Clontech Laboratories) reagents following the manufacturer's protocol.

Library Construction and DNA Sequencing

Libraries were constructed from cDNA yielded from IFCs in 96-well plates using an Nextera XT DNA sample preparation kit (Illumina). cDNA concentrations before and after library preparation was confirmed using High-Sensitivity DNA Tapes (Agilent) run on an Agilent Tape Station. For whole-tissue experiments, RNA was extracted from whole tissue using phenol-chloroform. Reverse transcription, initial cDNA synthesis, and cDNA library preparation were performed using the same reagents as for single cells. cDNA size distribution of single-cell, whole-tissue, and pooled libraries was performed using High-Sensitivity DNA Tapes (Agilent) on an Agilent Tape Station. Up to 96 single-cell libraries were pooled together for 50 bp, paired-end DNA sequencing on a single lane of an Illumina HiSeq2500 following the manufacturer's standard protocol. Single-cell libraries with >400,000 reads were used for further analysis. Reads were aligned to the mouse genome (mm10) using TopHat (Trapnell et al., 2009).

Statistical Analysis of Single-Cell Transcriptomes

PCA, Mann-Whitney-Wilcoxon rank-sum test, hierarchical clustering, and Student's t tests were performed using custom scripts in R. PCA was performed on $\log_{10}$ FPKM. Principal components with eigenvalues greater than 1 were evaluated to cluster cells. Genes identified by these components were used to perform hierarchical clustering. Euclidean distances were calculated from $\log_{10}$ FPKM and clustered by the complete linkage method. Consensus clustering was performed using the Single-Cell Consensus Clustering (SC3) package (Kiselev et al., 2016). Only cells expressing >3,000 genes were used for clustering. The k and distance metrics were chosen to maximize the silhouette width. Dendrograms were derived from consensus clustering and overlaid over heatmaps depicting marker gene expression also calculated with the SC3 tool. RaceID was performed as described by Grun et al. (2015). Student's t tests were two tailed and assuming unequal variance.

Gene ontology and KEGG pathway enrichment analysis were performed using DAVID software (Huang et al., 2007). "Perfect marker" gene analysis was performed as described by Treutlein et al. (2014). In brief, a "perfect marker" gene was defined for each cell lineage such that its expression was set at $\log_{10}$(FPKM) = 10 in all cells of the desired lineage (i.e., myocytes) and 0 in all other cells. A pairwise Pearson's correlation was calculated comparing the expression profile across all cells for a given gene to that of the perfect marker gene.

Stem-Cell-Derived CM Analysis

We derived a gene expression axis correlating with developmental time from the set of 997 genes expressed in 75% of ventricular myocytes at E9.5, E11.5, E14.5, E18.5, P0, P3, and P21. For comparisons with human cells, we limited this list to the 847 genes with human orthologs (Bick et al., 2012). Sequence from short-read files were downloaded from the NCBI GEO databank (GEO: GSE47948 and GSE62913, respectively) and rerun through our RNA-seq pipeline (Li et al., 2015; Wamstad et al., 2012). Human whole-tissue isolation, cDNA library construction, and mRNA-seq analysis were performed as previously described by McKean et al. (2016).

Immunohistochemistry

Cells isolated from P0 or P3 LVs were plated onto laminin-coated chamber slides and incubated at 37°C for 1 hr. Cells were fixed with 4% paraformaldehyde for 20 min and washed with PBS subsequent to permeabilization with 0.1% (v/v) Triton X-100 in PBS. Donkey serum (10% [v/v]))was used to block non-specific binding. Cells were incubated with two primary antibodies, rabbit anti-Tnni3 (Abcam ab56357), or goat anti-Col3a1 (Abcam ab7778) diluted 1:250 in blocking solution for 1 hr at room temperature. Cells were washed in PBS and fluorophore-conjugated secondary antibody (Molecular Probes) diluted 1:1,000 in PBS for 1 hr. Slides were washed and then mounted and counterstained using Vectashield antifade mounting medium with DAPI (Vector Laboratories).

ACCESSION NUMBERS

All single-cell sequencing data reported in this article have been deposited in the GNomEx database under accession numbers 272R, 274R, 275-292R, 439R, and 440R and can be found at https://b2b.hci.utah.edu/gnomex/ by using the guest login.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.devcel.2016.10.001.

AUTHOR CONTRIBUTIONS

Conceptualization, Methodology, and Writing – Original Draft, D.M.D., A.G.B., J.G.S., and C.E.S.; Investigation, D.M.D., A.G.B., H.W., and J.M.G.; Software, Formal Analysis, and Visualization, D.M.D. and A.G.B; Resources, D.M., I.S.K., J.T.H., and J.M.G.; Review & Editing, H.W., D.M., I.S.K., J.T.H., J.H., J.M.G., W.P., B.G.B., and C.E.S.; Supervision and Funding Acquisition, J.G., W.P., B.G.B., J.G.S., and C.E.S.

ACKNOWLEDGMENTS

This work was supported by grants from the NIGMS T32GM007753 (A.G.B.), John S. LaDue Memorial Fellowship at Harvard Medical School (J.H.), NHLBI Bench to Bassinet Program (U01HL098179/UM1HL098179 to B.G.B., 2UM1HL098147 to C.E.S., 2UM1HL098166 to W.P., J.G.S., and C.E.S.), the Howard Hughes Medical Institute (C.E.S.), NIMH (R01 MH101528-01 to J.G.), and NIH (HL125807 to J.T.H.). I.S.K. was supported by the Foundation for Anesthesia Education and Research. The authors thank David Conner and Steve DePalma for helpful discussions, advice, and technical assistance.

Received: December 27, 2015
Revised: March 30, 2016
Accepted: September 30, 2016
Published: November 10, 2016

REFERENCES

Ambrosino, C., Iwata, T., Scafoglio, C., Mallardo, M., Klein, R., and Nebreda, A.R. (2006). TEF-1 and C/EBPbeta are major p38alpha MAPK-regulated transcription factors in proliferating cardiomyocytes. Biochem. J. *396*, 163–172.

Ashraf, H., Pradhan, L., Chang, E.I., Terada, R., Ryan, N.J., Briggs, L.E., Chowdhury, R., Zarate, M.A., Sugi, Y., Nam, H.J., et al. (2014). A mouse model of human congenital heart disease: high incidence of diverse cardiac anomalies and ventricular noncompaction produced by heterozygous Nkx2-5 homeodomain missense mutation. Circ. Cardiovasc. Genet. *7*, 423–433.

Benecke, A.G., and Eilebrecht, S. (2015). RNA-mediated regulation of HMGA1 function. Biomolecules *5*, 943–957.

Biben, C., Weber, R., Kesteven, S., Stanley, E., McDonald, L., Elliott, D.A., Barnett, L., Koentgen, F., Robb, L., Feneley, M., et al. (2000). Cardiac septal and valvular dysmorphogenesis in mice heterozygous for mutations in the homeobox gene Nkx2-5. Circ. Res. *87*, 888–895.

Bick, A.G., Calvo, S.E., and Mootha, V.K. (2012). Evolutionary diversity of the mitochondrial calcium uniporter. Science *336*, 886.

Bouveret, R., Waardenberg, A.J., Schonrock, N., Ramialison, M., Doan, T., de Jong, D., Bondue, A., Kaur, G., Mohamed, S., Fonoudi, H., et al. (2015). NKX2-5 mutations causative for congenital heart disease retain functionality and are directed to hundreds of targets. Elife *4*, e06942.

Braybrook, C., Warry, G., Howell, G., Arnason, A., Bjornsson, A., Moore, G.E., Ross, M.T., and Stanier, P. (2001). Identification and characterization of KLHL4, a novel human homologue of the *Drosophila* Kelch gene that maps within the X-linked cleft palate and Ankyloglossia (CPX) critical region. Genomics *72*, 128–136.

Bruneau, B.G., Nemer, G., Schmitt, J.P., Charron, F., Robitaille, L., Caron, S., Conner, D.A., Gessler, M., Nemer, M., Seidman, C.E., et al. (2001). A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. Cell *106*, 709–721.

Buettner, F., Natarajan, K.N., Casale, F.P., Proserpio, V., Scialdone, A., Theis, F.J., Teichmann, S.A., Marioni, J.C., and Stegle, O. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat. Biotechnol. *33*, 155–160.

Cheroki, C., Krepischi-Santos, A.C., Szuhai, K., Brenner, V., Kim, C.A., Otto, P.A., and Rosenberg, C. (2008). Genomic imbalances associated with mullerian aplasia. J. Med. Genet. *45*, 228–232.

Combs, M.D., Knutsen, R.H., Broekelmann, T.J., Toennies, H.M., Brett, T.J., Miller, C.A., Kober, D.L., Craft, C.S., Atkinson, J.J., Shipley, J.M., et al. (2013). Microfibril-associated glycoprotein 2 (MAGP2) loss of function has pleiotropic effects in vivo. J. Biol. Chem. *288*, 28869–28880.

Dahle, O., Kumar, A., and Kuehn, M.R. (2010). Nodal signaling recruits the histone demethylase Jmjd3 to counteract polycomb-mediated repression at target genes. Sci. Signal. *3*, ra48.

DeLaughter, D.M., Christodoulou, D.C., Robinson, J.Y., Seidman, C.E., Baldwin, H.S., Seidman, J.G., and Barnett, J.V. (2013). Spatial transcriptional profile of the chick and mouse endocardial cushions identify novel regulators of endocardial EMT in vitro. J. Mol. Cell Cardiol. *59*, 196–204.

Desplantez, T., McCain, M.L., Beauchamp, P., Rigoli, G., Rothen-Rutishauser, B., Parker, K.K., and Kleber, A.G. (2012). Connexin43 ablation in foetal atrial myocytes decreases electrical coupling, partner connexins, and sodium current. Cardiovasc. Res. *94*, 58–65.

Diez-Roux, G., Banfi, S., Sultan, M., Geffers, L., Anand, S., Rozado, D., Magen, A., Canidio, E., Pagani, M., Peluso, I., et al. (2011). A high-resolution anatomical atlas of the transcriptome in the mouse embryo. PLoS Biol. *9*, e1000582.

Elliott, D.A., Kirk, E.P., Yeoh, T., Chandar, S., McKenzie, F., Taylor, P., Grossfeld, P., Fatkin, D., Jones, O., Hayes, P., et al. (2003). Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. J. Am. Coll. Cardiol. *41*, 2072–2076.

Grun, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature *525*, 251–255.

Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science *350*, 1262–1266.

Huang, D.W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C., et al. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. *35*, W169–W175.

Ieda, M., Tsuchihashi, T., Ivey, K.N., Ross, R.S., Hong, T.T., Shaw, R.M., and Srivastava, D. (2009). Cardiac fibroblasts regulate myocardial proliferation through beta1 integrin signaling. Dev. Cell *16*, 233–244.

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). HEART DEVELOPMENT. Integration of Bmp and Wnt signaling by Hopx specifies commitment of cardiomyoblasts. Science *348*, aaa6071.

Jang, D.G., Sim, H.J., Song, E.K., Medina-Ruiz, S., Seo, J.K., and Park, T.J. (2015). A thioredoxin fold protein Sh3bgr regulates Enah and is necessary for proper sarcomere formation. Dev. Biol. *405*, 1–9.

Jensen, B., Wang, T., Christoffels, V.M., and Moorman, A.F. (2013). Evolution and development of the building plan of the vertebrate heart. Biochim. Biophys. Acta *1833*, 783–794.

Kasahara, H., Bartunkova, S., Schinke, M., Tanaka, M., and Izumo, S. (1998). Cardiac and extracardiac expression of Csx/Nkx2.5 homeodomain protein. Circ. Res. *82*, 936–946.

Kiselev, V.Y., Kirschner, K., Schaub, M.T., Andrews, T., Chandra, T., Natarajan, K.N., Reik, W., Barahona, M., Green, A.R., and Hemberg, M. (2016). SC3-consensus clustering of single-cell RNA-Seq data. bioRxiv. http://dx.doi.org/10.1101/036558.

Koo, J.H., Smiley, M.A., Lovering, R.M., and Margolis, F.L. (2007). Bex1 knock out mice show altered skeletal muscle regeneration. Biochem. Biophys. Res. Commun. *363*, 405–410.

Koshiba-Takeuchi, K., Takeuchi, J.K., Arruda, E.P., Kathiriya, I.S., Mo, R., Hui, C.C., Srivastava, D., and Bruneau, B.G. (2006). Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart. Nat. Genet. *38*, 175–183.

Kuppusamy, K.T., Jones, D.C., Sperber, H., Madan, A., Fischer, K.A., Rodriguez, M.L., Pabon, L., Zhu, W.Z., Tulloch, N.L., Yang, X., et al. (2015). Let-7 family of microRNA is required for maturation and adult-like metabolism in stem cell-derived cardiomyocytes. Proc. Natl. Acad. Sci. USA *112*, E2785–E2794.

Li, G., Plonowska, K., Kuppusamy, R., Sturzu, A., and Wu, S.M. (2015). Identification of cardiovascular lineage descendants at single-cell resolution. Development *142*, 846–857.

Liang, B., Soka, M., Christensen, A.H., Olesen, M.S., Larsen, A.P., Knop, F.K., Wang, F., Nielsen, J.B., Andersen, M.N., Humphreys, D., et al. (2014). Genetic variation in the two-pore domain potassium channel, TASK-1, may contribute to an atrial substrate for arrhythmogenesis. J. Mol. Cell Cardiol. *67*, 69–76.

Liberatore, C.M., Searcy-Schrick, R.D., and Yutzey, K.E. (2000). Ventricular expression of tbx5 inhibits normal heart chamber development. Dev. Biol. *223*, 169–180.

Luo, M.H., Li, Y.S., and Yang, K.P. (2007). Fibrosis of collagen I and remodeling of connexin 43 in atrial myocardium of patients with atrial fibrillation. Cardiology *107*, 248–253.

Mayer, D.C., and Leinwand, L.A. (1997). Sarcomeric gene expression and contractility in myofibroblasts. J. Cell Biol. *139*, 1477–1484.

McKean, D.M., Homsy, J., Wakimoto, H., Patel, N., Gorham, J., DePalma, S.R., Ware, J.S., Zaidi, S., Ma, L., Patel, V., et al. (2016). Loss of RNA expression and allele-specific expression associated with congenital heart disease. Nat. Commun. *7*, 12824.

Medic, S., and Ziman, M. (2010). PAX3 expression in normal skin melanocytes and melanocytic lesions (naevi and melanomas). PLoS One *5*, e9977.

Monzen, K., Ito, Y., Naito, A.T., Kasai, H., Hiroi, Y., Hayashi, D., Shiojima, I., Yamazaki, T., Miyazono, K., Asashima, M., et al. (2008). A crucial role of a

high mobility group protein HMGA2 in cardiogenesis. Nat. Cell Biol. *10*, 567–574.

Nash, D., Arrington, C.B., Kennedy, B.J., Yandell, M., Wu, W., Zhang, W., Ware, S., Jorde, L.B., Gruber, P.J., Yost, H.J., et al. (2015). Shared segment analysis and next-generation sequencing implicates the retinoic acid signaling pathway in total anomalous pulmonary venous return (TAPVR). PLoS One *10*, e0131514.

Niaudet, C., Hofmann, J.J., Mae, M.A., Jung, B., Gaengel, K., Vanlandewijck, M., Ekvarn, E., Salvado, M.D., Mehlem, A., Al Sayegh, S., et al. (2015). Gpr116 receptor regulates distinctive functions in pneumocytes and vascular endothelium. PLoS One *10*, e0137949.

Palencia-Desai, S., Rost, M.S., Schumacher, J.A., Ton, Q.V., Craig, M.P., Baltrunaite, K., Koenig, A.L., Wang, J., Poss, K.D., Chi, N.C., et al. (2015). Myocardium and BMP signaling are required for endocardial differentiation. Development *142*, 2304–2315.

Pinto, A.R., Ilinykh, A., Ivey, M.J., Kuwabara, J.T., D'Antoni, M.L., Debuque, R., Chandran, A., Wang, L., Arora, K., Rosenthal, N.A., et al. (2016). Revisiting cardiac cellular composition. Circ. Res. *118*, 400–409.

Pujades, C., Guez-Guez, B., and Dunon, D. (2002). Melanoma cell adhesion molecule (MCAM) expression in the myogenic lineage during early chick embryonic development. Int. J. Dev. Biol. *46*, 263–266.

Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. *20*, 53–65.

Schott, J.J., Benson, D.W., Basson, C.T., Pease, W., Silberbach, G.M., Moak, J.P., Maron, B.J., Seidman, C.E., and Seidman, J.G. (1998). Congenital heart disease caused by mutations in the transcription factor NKX2-5. Science *281*, 108–111.

Smith, C.M., Finger, J.H., Hayamizu, T.F., McCright, I.J., Xu, J., Berghout, J., Campbell, J., Corbani, L.E., Forthofer, K.L., Frost, P.J., et al. (2014). The mouse gene expression database (GXD): 2014 update. Nucleic Acids Res. *42*, D818–D824.

Streets, A.M., and Huang, Y. (2014). How deep is enough in single-cell RNA-seq? Nat. Biotechnol. *32*, 1005–1006.

Tanaka, M., Chen, Z., Bartunkova, S., Yamasaki, N., and Izumo, S. (1999). The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. Development *126*, 1269–1280.

Thomas, T., Yamagishi, H., Overbeek, P.A., Olson, E.N., and Srivastava, D. (1998). The bHLH factors, dHAND and eHAND, specify pulmonary and systemic cardiac ventricles independent of left-right sidedness. Dev. Biol. *196*, 228–236.

Tian, Y., Yuan, L., Goss, A.M., Wang, T., Yang, J., Lepore, J.J., Zhou, D., Schwartz, R.J., Patel, V., Cohen, E.D., et al. (2010). Characterization and in vivo pharmacological rescue of a Wnt2-Gata6 pathway required for cardiac inflow tract development. Dev. Cell *18*, 275–287.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature *509*, 371–375.

Vander Heiden, M.G., Locasale, J.W., Swanson, K.D., Sharfi, H., Heffron, G.J., Amador-Noguez, D., Christofk, H.R., Wagner, G., Rabinowitz, J.D., Asara, J.M., et al. (2010). Evidence for an alternative glycolytic pathway in rapidly proliferating cells. Science *329*, 1492–1499.

Wagner, M., and Siddiqui, M.A. (2007). Signal transduction in early heart development (I): cardiogenic induction and heart tube formation. Exp. Biol. Med. *232*, 852–865.

Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. Cell *151*, 206–220.

Yang, M.Y., Hilton, M.B., Seaman, S., Haines, D.C., Nagashima, K., Burks, C.M., Tessarollo, L., Ivanova, P.T., Brown, H.A., Umstead, T.M., et al. (2013). Essential regulation of lung surfactant homeostasis by the orphan G protein-coupled receptor GPR116. Cell Rep. *3*, 1457–1464.

Yuan, F., Qiu, X.B., Li, R.G., Qu, X.K., Wang, J., Xu, Y.J., Liu, X., Fang, W.Y., Yang, Y.Q., and Liao, D.N. (2015). A novel NKX2-5 loss-of-function mutation predisposes to familial dilated cardiomyopathy and arrhythmias. Int. J. Mol. Med. *35*, 478–486.