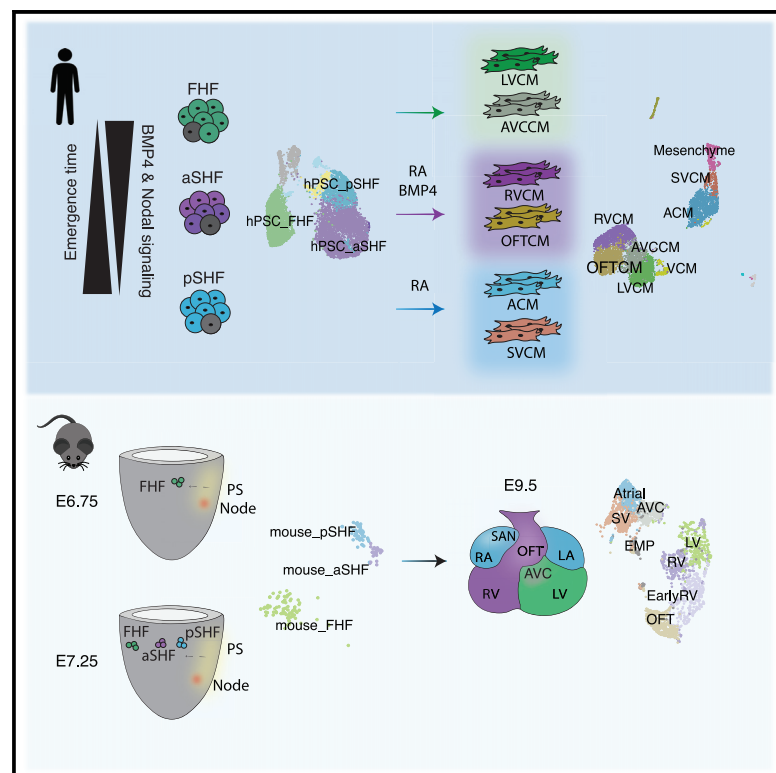# Modeling human multi-lineage heart field development with pluripotent stem cells

Graphical abstract

Authors

Donghe Yang, Juliana Gomez-Garcia, Shunsuke Funakoshi, ..., Gary D. Bader, Michael A. Laflamme, Gordon M. Keller

Correspondence

donghe.yang@mail.utoronto.ca (D.Y.), gordon.keller@uhnresearch.ca (G.M.K.)

In brief

We describe hPSC-derived models of FHF, aSHF, and pSHF development spanning the mesoderm, progenitor, and cardiomyocyte stages. Comparison between the hPSC-derived and the corresponding mouse populations revealed conserved developmental trajectories across the species. Access to cardiomyocyte subtypes derived from different heart fields enables modeling chamber-specific diseases and developing new therapies.

Highlights

- ScRNA-seq profiling of human heart field development spanning various stages

- Comparisons of hPSC-derived, human fetal-, and mouse fetal-cardiac populations

- Identification of pathways involved in the development of heart field lineages

- Generation of functionally and transcriptionally distinct cardiomyocyte (CM) subtypes

CellPress

Resource

Modeling human multi-lineage heart field development with pluripotent stem cells

Donghe Yang,[1,3,*] Juliana Gomez-Garcia,[1,6] Shunsuke Funakoshi,[1,9] Thinh Tran,[4,5,10] Ian Fernandes,[1,3] Gary D. Bader,[2,4,5] Michael A. Laflamme,[1,7] and Gordon M. Keller[1,2,3,8,*]

[1]McEwen Stem Cell Institute, University Health Network, Toronto, ON M5G 1L7, Canada
[2]Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada
[3]Department of Medical Biophysics, University of Toronto, Toronto, ON M5G 1L7, Canada
[4]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada
[5]The Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada
[6]Department of Biomedical Engineering, University of Toronto, Toronto, ON M5S 1A8, Canada
[7]Department of Laboratory Medicine & Pathobiology, University of Toronto, Toronto, ON M5S 1A8, Canada
[8]Lead contact
[9]Present address: Center for iPS Cell Research and Application, Kyoto University, Kyoto, 606-8507, Japan
[10]Present address: Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
*Correspondence: donghe.yang@mail.utoronto.ca (D.Y.), gordon.keller@uhnresearch.ca (G.M.K.)
https://doi.org/10.1016/j.stem.2022.08.007

SUMMARY

The cardiomyocyte (CM) subtypes in the mammalian heart derive from distinct lineages known as the first heart field (FHF), the anterior second heart field (aSHF), and the posterior second heart field (pSHF) lineages that are specified during gastrulation. We modeled human heart field development from human pluripotent stem cells (hPSCs) by using single-cell RNA-sequencing to delineate lineage specification and progression. Analyses of hPSC-derived and mouse mesoderm transcriptomes enabled the identification of distinct human FHF, aSHF, and pSHF mesoderm subpopulations. Through staged manipulation of signaling pathways identified from transcriptomics, we generated myocyte populations that display molecular characteristics of key CM subtypes. The developmental trajectory of the human cardiac lineages recapitulated that of the mouse, demonstrating conserved cardiovascular programs. These findings establish a comprehensive landscape of human embryonic cardiogenesis that provides access to a broad spectrum of cardiomyocytes for modeling congenital heart diseases and chamber-specific cardiomyopathies as well as for developing new therapies to treat them.

INTRODUCTION

The ability to differentiate human pluripotent stem cells (hPSCs) to the cardiovascular lineages has opened new avenues to study the earliest stages of human heart development, generate models of heart diseases, and create new therapies to treat some of some of the most devastating and debilitating forms of these diseases. Because different diseases target different regions of the heart, the therapeutic applications of this model are dependent on our ability to generate the appropriate cell types from hPSCs. The adult heart comprises different cardiomyocyte (CM) subtypes that include left and right ventricular and atrial CMs that form the working myocardium, the sinoatrial and atrioventricular nodal cells that represent the pacemakers, and the outflow and inflow tract cells that connect the heart to the vasculature (Litviňuková et al., 2020). Our current understanding of the origins of these heart lineages comes from

studies in model organisms that showed that the CM subtypes derive from distinct progenitors known as the first heart field (FHF), the anterior second heart field (aSHF), and the posterior second heart field (pSHF) progenitors (Galdos et al., 2017; Meilhac and Buckingham, 2018; Cai et al., 2003). These progenitors are specified by E7.5 in mouse and are distinguished at this stage by gene expression patterns and their position within the cardiac crescent region of the embryo (Cai et al., 2003; Kelly, 2012). The FHF progenitors, identified by the expression of Hcn4, Hand1, and Tbx5, give rise predominantly to left ventricular cardiomyocytes (LVCMs) and atrioventricular canal cardiomyocytes (AVCCMs) along with some atrial cardiomyocytes (ACMs) (Liang et al., 2013; Bruneau et al., 1999; Barnes et al., 2010; Später et al., 2013; de Soysa et al., 2019). The aSHF progenitors are characterized by the expression of Tbx1, Fgf8, and Fgf10 and contribute to right ventricular cardiomyocyte (RVCM) and outflow tract (OFT) lineages (Kelly et al., 2001; High et al., 2009; Nevis et al., 2013;

Francou et al., 2013), whereas the pSHF progenitors, defined by the expression of *Hoxb1* and *Nr2f2*, give rise to ACMs and sinus venosus (SV)-derived structures including the sinoatrial node pacemaker cells (Bertrand et al., 2011).

Although FHF and SHF progenitors are recognized as distinct populations in the cardiac crescent, lineage tracing and retrospective clonal studies have provided compelling evidence that the segregation of these fates occurs much earlier, at the time of mesoderm induction (Lescroart et al., 2014; Devine et al., 2014; Meilhac et al., 2004). Lescroart et al. showed that the FHF derivatives develop from *Mesp1*+ mesoderm that is induced earlier (E6.75) than the mesoderm that gives rise to the SHF lineages (E7.25) and that these temporally distinct populations display different transcriptomes (Lescroart et al., 2014). Single-cell RNA sequencing (scRNA-seq) analyses of these early stage populations provided the resolution that enabled the identification of FHF, aSHF, and pSHF mesoderm (Lescroart et al., 2018). More recent studies confirmed the spatiotemporal allocation of these fates and showed that the FHF mesoderm is induced earliest, followed by the aSHF mesoderm, which in turn is followed by the pSHF mesoderm (Ivanovitch et al., 2021).

Studies aimed at identifying the signaling pathways that control FHF and SHF development largely focused on the mesoderm and progenitor stages and provided evidence that the two populations are regulated differently. Bone morphogenetic protein (BMP) signaling plays a pivotal role in initiating the FHF lineage, which differentiates rapidly and forms the first contracting population within the heart tube (Klaus et al., 2007; Lescroart et al., 2018; Andersen et al., 2018). The SHF progenitors, by contrast, are exposed to an fibroblast growth factor (FGF)/Wnt signaling environment that promotes their proliferation, thereby delaying differentiation (Cohen et al., 2007; Ai et al., 2007; Rochais et al., 2009; Ilagan et al., 2006; Park et al., 2006). Specification of the atrial lineage from the pSHF progenitors is dependent on retinoic acid (RA) signaling (Stefanovic and Zaffran, 2017; Zhang et al., 2011; Gassanov et al., 2008), whereas the generation of CMs from the aSHF progenitors is controlled in part by BMP signaling (Thomas et al., 2014; Hutson et al., 2010). Regulation of FHF/SHF mesoderm induction is less well understood but likely involves pathways that control gastrulation, including BMP, nodal, Wnt, and FGF (Ciruna and Rossant, 2001; Mishina et al., 1995; Winnier et al., 1995; Liu et al., 1999).

The identification and characterization of comparable heart field lineages in the developing human heart has been hampered by the inaccessibility of early fetal tissue. Our insights to date are limited to recent molecular studies that identified distinct progenitor-like populations in the OFT and ventricular tissue of fetal hearts at 5 weeks of gestation (Sahara et al., 2019). These progenitors, known as con-ventricular progenitors (CVP) and free wall-ventricular progenitors (FVP), could represent different heart field lineages because they are molecularly and anatomically distinct. Studies using the hPSC model have provided evidence that FHF/SHF lineage diversification takes place during human cardiovascular development. We previously showed that the ventricular and atrial fates are established early during hPSC differentiation and that these CMs develop from distinct mesoderm populations induced by different concentrations of Activin A and BMP4 (Lee et al., 2017). In a more recent study, Zhang et al. used a dual reporter hPSC line to follow *TBX5* and

NKX2.5 expression in the differentiating populations and showed that *TBX5*+*NKX2.5*+ progenitors give rise to VCMs that likely represent FHF derivatives (Zhang et al., 2019). By using the same reporter strategy, Pezhouman et al. demonstrated that the *TBX5*+*NKX2.5*+ and *TBX5*−*NKX2.5*+ progenitors display expression patterns indicative of the FHF and SHF lineages, respectively (Pezhouman et al., 2021). The CM potential of the populations did not, however, reflect that expected from the mouse studies: both gave rise predominantly to atrial-like cells. To access heart-field-specific progenitors without genetic modification, Andersen et al. isolated populations based on expression of CXCR4 and showed that the CXCR4+ and CXCR4− cells express genes associated with the SHF and FHF, respectively (Andersen et al., 2018). Because the derivative CM populations were not characterized in this study, it is unknown whether the cardiogenic potential of these cells aligns with that of FHF and SHF lineages in the mouse. Taken together, these findings suggest that the FHF/SHF segregation does occur during human cardiovascular development. However, given the limited analyses to date, it is unknown whether the temporal diversification of cell fates in the human recapitulates that found in the mouse, nor is it known whether the developmental potential of different cardiac progenitors is conserved in the two species.
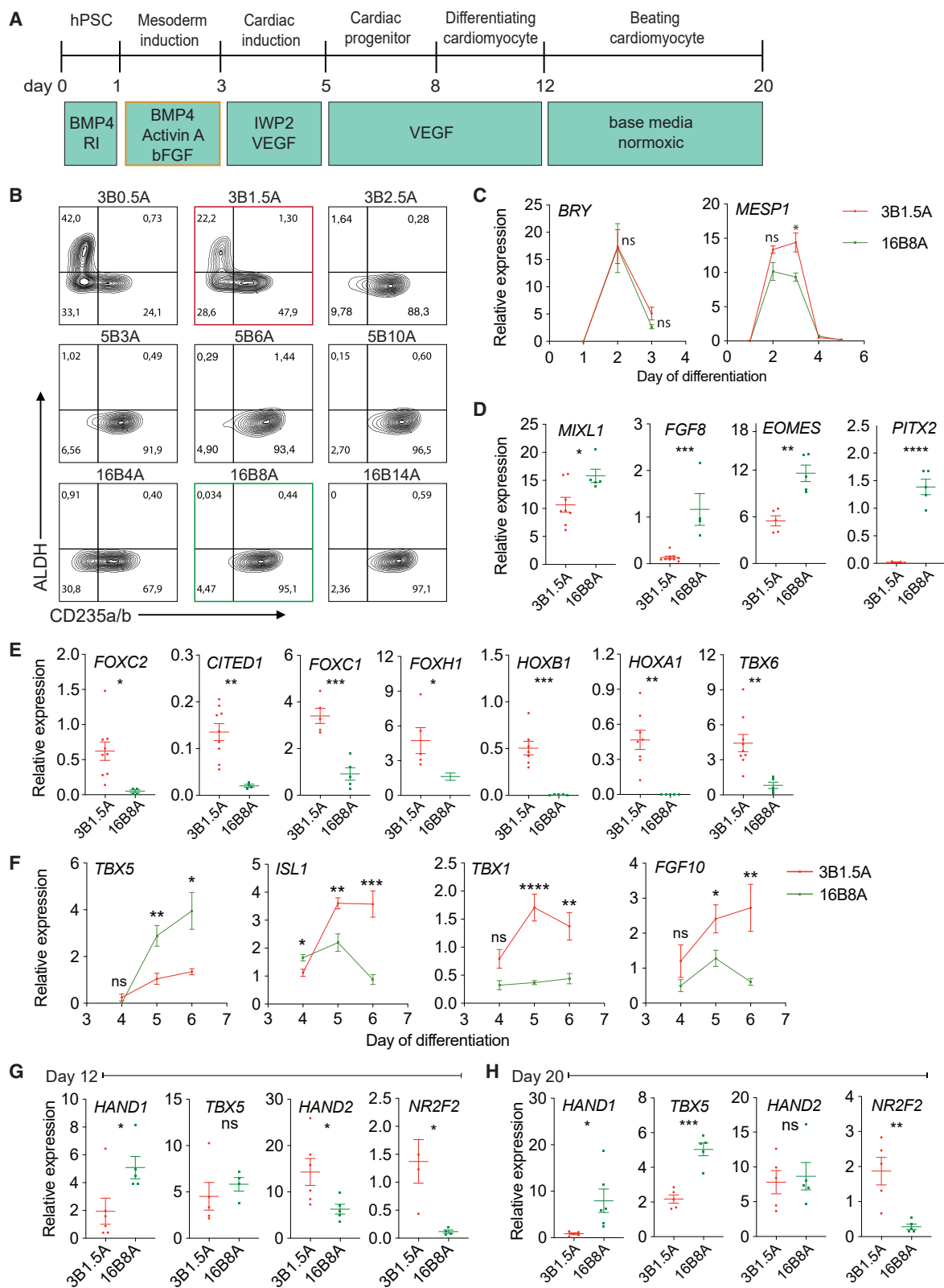
To further address the question of human heart fields, we carried out comprehensive scRNA-seq analyses of hPSC-derived mesoderm, progenitor, and CM populations induced under different conditions. With this approach, we identified distinct human FHF, aSHF, and pSHF lineages and showed that they display remarkable similarities to their counterparts in the mouse. Through the staged manipulation of signaling pathways identified from the scRNA-seq analysis, we were able to generate myocyte populations that display molecular characteristics of RVCMs, LVCMs, ACMs, AVCCMs, SVCMs, and OFTCMs.

RESULTS

Generation of FHF and SHF cardiac mesoderm populations

To generate the different heart field lineages, we varied Activin A and BMP4 concentrations between days 1 and 3 of differentiation to induce mesoderm subsets that display the defining features of populations with FHF and SHF potential (Figure 1A). At day 4, mesoderm induction was monitored by expression of PDGFRα together with either CD235a/b or ALDH1A2, markers that we have shown to track with the induction of ventricular and atrial mesoderm, respectively (Lee et al., 2017) (Figures 1B and S1A). For these analyses, we used the Aldefluor (ALDH) assay as a measure of ALDH1A2 activity (Jones et al., 1995).

As observed in our previous study, mesoderm populations generated with high concentrations of BMP4 and Activin A (over 5 ng/mL of BMP4 and 3 ng/mL of Activin A) expressed CD235a/b and lacked ALDH activity (Figure 1B), whereas populations specified by lower concentrations of factors (3 ng/mL BMP4 and 0.5–1.5 ng/mL Activin A) contained ALDH+ cells and relatively low numbers of CD235a/b+ cells. Notably, small changes in the concentration of Activin A (1.5 versus 2.5 ng/mL) dramatically impacted the proportion of ALDH+ cells induced (22% versus 1.6%) (Figure 1B). Comparison of

(legend on next page)

populations generated with 3B1.5A, 5B6A, 16B4A, and 16B8A showed that 16B8A-induced mesoderm contained the highest frequency of PDGFRα⁺CD235a/b⁺ cells, whereas 3B1.5A-induced mesoderm contained a high frequency of ALDH⁺ and PDGFRα⁺ cells (Figures 1B and S1A).

Molecular analyses of the populations induced with 3B1.5A and 16B8A showed similar expression kinetics of primitive streak (PS) (*BRY* or *TBXT*) and mesoderm (*MESP1*) genes, indicating comparable temporal development (Figure 1C). To determine whether the two hPSC-derived mesoderm subsets share similarities to the mouse FHF and SHF mesoderm identified by Lescroart et al., we analyzed these mouse populations for distinguishable gene expression patterns that could be used as markers of these lineages. The outcome of these analyses identified expression patterns (markers) that distinguished the early emerging (E6.75) and later emerging (E7.25) mesoderm (Figure S1B). Additionally, the findings showed that the early population expressed genes associated with the FHF mesoderm, whereas the later population contained subpopulations that displayed molecular profiles of aSHF and pSHF mesoderm (Figures S1C and S1D). Gene ontology (GO) analysis revealed an enrichment of nodal/activin signaling in the FHF mesoderm (Figure S1E), suggesting that this pathway plays a role in the induction of this population in mouse embryo. RT-qPCR analyses of the two hPSC-derived populations showed that the FHF genes were preferentially expressed in the 16B8A mesoderm (Figure 1D), whereas those associated with the SHF were more highly expressed in the 3B1.5A mesoderm (Figure 1E). Analyses of the populations generated with different concentrations of Activin A (3B0.5A, 3B1.5A, and 3B2.5A) indicated that the mesoderm induced with 3B1.5A expressed a combined aSHF/pSHF profile, suggesting that it is a heterogeneous population containing both aSHF and pSHF mesoderm (Figures S1F and S1G).

To evaluate the two hPSC-derived mesoderm populations with respect to FHF and SHF potential, we cultured the cells further and analyzed the resulting populations for expression of FHF and SHF genes. At days 5 and 6, the stage at which cardiovascular progenitors are specified, the cells generated from the 16B8A mesoderm expressed higher levels of the FHF progenitor marker *TBX5* and lower levels of the SHF progenitor markers *FGF10*, *TBX1*, and *ISL1* than did the derivatives of the 3B1.5A mesoderm (Figure 1F). Although both populations efficiently generated CMs (Figures S1H and S1I), they displayed differential gene expression patterns at days 12 and/or 20 of differentiation. CMs generated from the 16B8A induction expressed higher levels of the left ventricular (FHF) genes *HAND1* and *TBX5* (day 20) than did those derived from the 3B1.5A induction (Figures 1G and 1H). In contrast, expression of *HAND2* (day 12) associated with right ventricular (SHF) development and *NR2F2* indicative of atrial

(SHF) differentiation were expressed at higher levels in the 3B1.5A-derived CMs (Figures 1G and 1H). Taken together, these findings support the interpretation that the 16B8A- and 3B1.5A-induced populations represent FHF and SHF lineages, respectively.

Heterogeneity of hPSC-derived cardiac mesoderm populations

To further investigate the transcriptomes of the 3B1.5A- and 16B8A-induced mesoderm populations, we carried out scRNA-seq at day 3 of differentiation. Analyses of the day 3 data revealed the presence of *PDGFRA*⁺*MESP1*⁺ mesoderm clusters, a *POU5F1*⁺*NANOG*⁺*TBXT*ʰⁱᵍʰ PS cluster, and a *SOX17*⁺*FOXA2*⁺ endoderm cluster (Figures 2A, S2A, and S2B). Based on the expression of genes associated with mouse FHF, aSHF, and pSHF mesoderm development (Figure S1D), we annotated the 16B8A population as FHF mesoderm because it expressed the highest levels of the FHF markers *GSC*, *EOMES*, and *LHX1* (Figure 2B; Table S2). Within the 3B1.5A mesoderm, we identified a putative aSHF population that expressed *SIX1* and *FOXC2* and a putative pSHF cluster that expressed *HOXA1* and *ALDH1A2* (Figure 2B; Table S2). None of the clusters expressed genes indicative of the presence of paraxial mesoderm (*PAX3*, *MEOX1*, and *MYF5*) or neuroectoderm (*PAX6* and *SOX1*), consistent with the interpretation that they represent lateral plate mesoderm (Figure S2B).

We next sought to identify the gene expression patterns of each mesoderm subtype that are conserved between human and mouse by performing differential expression analyses on the human and mouse FHF, aSHF, and pSHF mesoderm clusters. Overall, 120 differentially expressed genes (DEGs) were detected that showed conserved cross-species patterns (Table S2; Figure 2C). We found that *PITX2* and *GSC*, conventionally recognized as PS laterality and polarity markers, were preferentially expressed in the FHF mesoderm, whereas *HOTAIRM1*, a long non-coding RNA located between the *HOXA1* and *HOXA2* loci (Zhang et al., 2009; Zhao et al., 2020), was expressed at higher levels in the pSHF mesoderm (Table S2; Figure 2D). RT-qPCR analyses confirmed some of the differential expression patterns and showed that the 16B8A-induced mesoderm expressed higher levels of the newly identified FHF genes (Figure 2E). By contrast, the pSHF and aSHF genes were detected at significantly higher levels in the 3B1.5A-induced mesoderm (Figures 2F and 2G). These findings further support the interpretation that 16B8A-induced population represents FHF mesoderm, whereas the one induced with 3B1.5A contains both aSHF and pSHF mesoderm.

To identify gene networks that regulate distinct mesoderm populations, we next performed GO and gene regulatory network (SCENIC) analyses (Aibar et al., 2017; Yu et al., 2012). The conserved FHF genes from this analysis included those

Figure 1. Generation of FHF and SHF mesoderm

(A) Protocol used to generate cardiomyocytes from hPSCs. RI, rock inhibitor.

(B) Flow cytometric analyses of ALDH activity and CD235a/b expression in day 4 populations. B, BMP4; A, activin A.

(C–H) RT-qPCR analyses of (C) *BRY* (*TBXT*) and *MESP1* expression in the 3B1.5A and 16B8A populations (n ≥ 3); (D and E) expression of the E6.75 mesoderm markers (D) and E7.25 mesoderm markers (E) in the day 3 populations (n ≥ 4); (F) expression of the indicated genes in the days 4, 5, and 6 cells; and (G and H) expression of the indicated genes in the day 12 and 20 cells (n ≥ 5). Unpaired t test; *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. Error bars represent SEM.
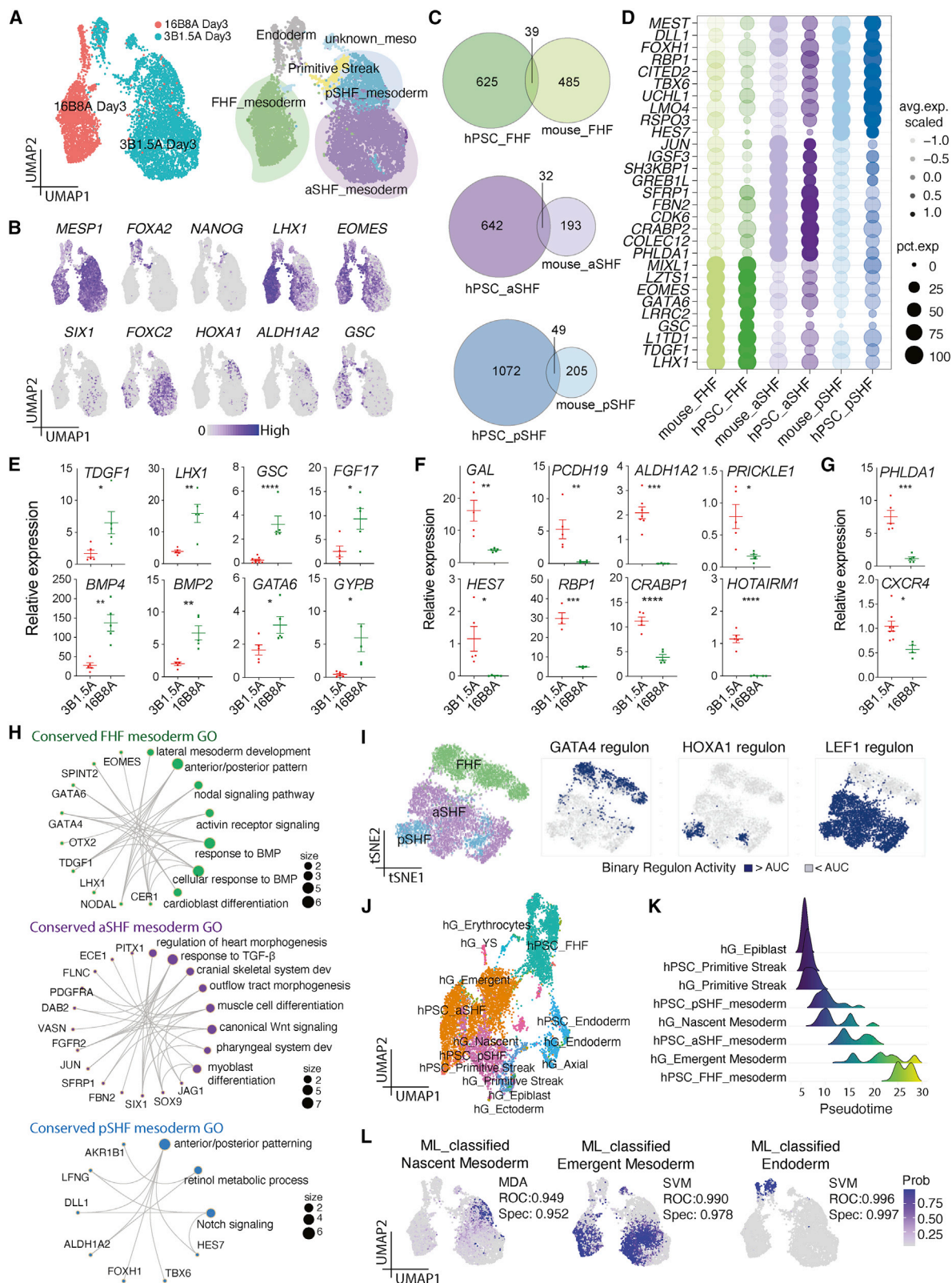
Figure 2. Transcriptomic analyses of day 3 mesoderm populations
(A) UMAP of day 3 3B1.5A and 16B8A populations.
(B) UMAP showing the expression of the indicated genes.

(legend continued on next page)

involved in nodal/activin and BMP signaling (Figure 2H)—pathways identified in the mouse as putative regulators of this population (Figure S1E) (Lescroart et al., 2018). Gene sets associated with OFT morphogenesis, pharyngeal development, and Wnt signaling were enriched in the aSHF mesoderm, a categorization consistent with findings in the mouse that this mesoderm is regulated in part by Wnt signaling and gives rise to OFT CMs and the pharyngeal system (Anderson et al., 2016; Kelly and Buckingham, 2002; Mjaatvedt et al., 2001; Yang et al., 2013). The conserved expression patterns of the pSHF included components of RA signaling, a pathway known to regulate specification of this mesoderm to derivative fates (Gassanov et al., 2008; Wiesinger et al., 2021; Zhang et al., 2011). In addition to the conserved patterns, we also identified genes specifically upregulated in the hPSC-derived mesoderm and found an enrichment of gene sets involved in mitochondrial metabolism in the FHF, those related to sterol synthesis and histone acetylation in the aSHF, and those associated with DNA replication in the pSHF (Figure S2C). To identify putative key regulators of the FHF, aSHF, and pSHF mesoderm populations, we carried out SCENIC analyses. These analyses identified the gene networks of *GATA4*, a master regulator of cardiac development (Yilbas et al., 2014; Zhou et al., 2012), in the FHF mesoderm; *HOXA1*, a downstream target of RA signaling (Stefanovic and Zaffran, 2017), in the pSHF mesoderm; and *LEF1*, a key component of Wnt signaling, in the aSHF and pSHF mesoderm (Figure 2I).

To determine whether the hPSC-derived FHF, aSHF, and pSHF represent temporally distinct populations, we integrated our data with scRNA-seq data from a human gastrulating embryo that contained temporally distinct nascent, emergent, and advanced mesoderm (Tyser et al., 2021). These integration analyses revealed a significant overlap between the hPSC- and human embryo-derived data (Figures 2J and S2D). Pseudotime analysis of the integrated data together with machine learning (ML)-based classification suggested that hPSC-derived FHF and aSHF populations represent embryonic emergent mesoderm, a population more advanced than the nascent mesoderm that shows similarity to the hPSC-pSHF mesoderm (Figures 2K, 2L, and S2E). Analyses of the stage-specific markers along pseudotime confirmed the transition from the *NANOG*⁺ PS to *MESP1*⁺ mesoderm (Figure S2F). Markers of FHF (*BMP2*), aSHF (*IRX1*), and pSHF (*HOXA1*) were expressed at distinct pseudotimes in a sequential order, supporting a temporal specification of these mesoderm subtypes in a pattern similar to that observed in the mouse (Figure S2F).

In addition to the day 3 populations, we also performed scRNA-seq on day 4 populations induced under either FHF or SHF conditions (Figure S2F). Analyses of these populations showed a reduction in the proportion of *MESP1*⁺ cells compared

to those seen in the day 3 populations and a high proportion of *GATA6*⁺ cells indicative of the specification of the cardiac lineage. *FOXA2*⁺ endoderm and *PECAM1*⁺ endothelial cell clusters were also detected (Figures S2G and S2H). Detailed analyses revealed the presence of a *HAND1*^high^*BMP4*^high^ FHF cluster, a *FOXC2*^high^*TBX1*^high^ aSHF cluster, and a *NR2F1*^high^*HOXB1*^high^ pSHF cluster (Figure S2I; Table S3). Based on these expression patterns, we consider the day 4 populations to represent late-stage mesoderm, undergoing the initial specification steps to cardiac progenitors.

Generation of distinct cardiac progenitors from purified mesoderm populations

To be able to separate the FHF, aSHF, and pSHF mesoderm populations for functional analyses, we next queried our datasets to identify fluorescence-activated cell sorting (FACS)-compatible markers that could be used to isolate these populations. We found that *CXCR4* and *ITGA3* were expressed at highest levels in the aSHF, that *CD1D* and *ALDH1A2* levels were highest in the pSHF, and that *GYPB* was preferentially expressed in the FHF (Figure S3A; Table S2). RT-qPCR analyses confirmed the differential expression of these markers in the 3B1.5A- and 16B8A-specified mesoderm (Figures 2E–2G, and S3B). Pearson correlation analysis showed that the expression of *GYPB*, *CXCR4*, and *ALDH1A2* positively correlated with the expression of the FHF, aSHF, and pSHF mesoderm genes, respectively (Figure 3A).

Flow cytometric analyses of PDGFRα⁺ mesoderm from the day 4 16B8A-induced FHF population showed that it expressed CD235a/b but little if any CXCR4, and it had no ALDH activity (Figure 3B). The day 4 3B1.5A population could be segregated into a CXCR4⁺ fraction that lacked ALDH activity indicative of aSHF mesoderm and a CXCR4⁻ fraction that contained ALDH⁺ cells reflective of pSHF mesoderm (Figure 3C). Although CD235a/b protein was also expressed in the aSHF mesoderm, its transcript did not correlate with the aSHF mesoderm markers. By using similar strategies, we validated *CD1D* and *ITGA3* (CD49c) as cell surface markers of pSHF and aSHF mesoderm, respectively (Figures S3C–S3E).

To determine the potential of these different day 4 subpopulations, they were isolated and cultured overnight as aggregates (Figure 3D). On day 5, they were harvested and analyzed for expression of cardiac progenitor markers identified from analyses of E7.75 mouse FHF, aSHF, and pSHF progenitors (Figure S3F) (de Soysa et al., 2019). As shown in Figure 3E, the FHF mesoderm (ALDH⁻CD235a/b⁺CXCR4⁻/low)-derived population displayed the FHF progenitor signature. The cells generated from the ALDH⁺CXCR4⁻ pSHF mesoderm preferentially expressed the pSHF markers (Figure 3F), whereas those

(C) Genes displaying conserved and non-conserved expression patterns between the hPSC-derived and mouse populations.

(D) Genes showing species-conserved expression patterns in the indicated populations.

(E–G) RT-qPCR analyses of expression levels of (E) FHF, (F) pSHF, and (G) aSHF mesoderm genes in the day 3 populations (n ≥ 5). Unpaired t test; *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. Error bars represent SEM
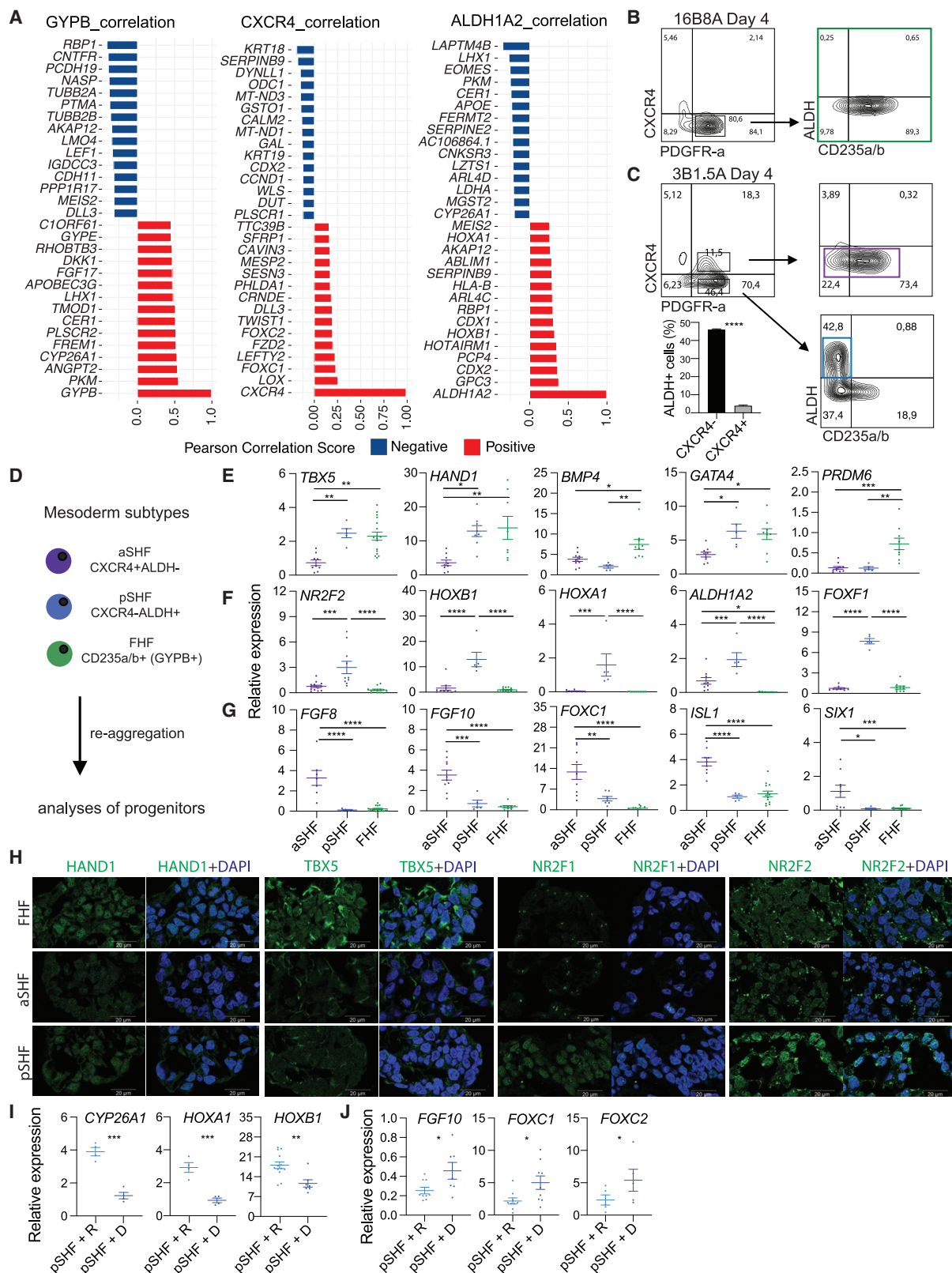
(H) Species-conserved genes and GO terms enriched in the mesoderm clusters.

(I) TSNE showing the indicated gene networks.

(J) Integrated hPSC-mesoderm and human gastrulating (hG) cells labeled by cell identity.

(K) Pseudotime ordering of hPSC-mesoderm and hG cells.

(L) UMAP showing the probability of classification.

(legend on next page)

derived from the ALDH⁻CXCR4⁺ aSHF mesoderm expressed higher levels of the aSHF markers (Figure 3G). Some of the FHF genes were expressed at comparable levels in the pSHF population, a finding consistent with studies in model organisms (Rankin et al., 2021; Ryan and Chin, 2003; Wiesinger et al., 2021). Immunocytochemistry analyses also showed differences among these populations, with the FHF progenitors displaying the highest levels of HAND1 and TBX5 and the pSHF progenitors showing elevated levels of NR2F1 and NR2F2 (Figures 3H and S3G).

To determine whether CD1d could be used in place of ALDH as a marker of pSHF mesoderm, we isolated the CD1d^{high}CXCR4⁻ and CD1d$^{low/−}$CXCR4⁺ fractions from the day 4 SHF population and analyzed the day 5 derivative cells for expression of pSHF and aSHF progenitor markers (Figure S3H). The CD1d^{high}CXCR4⁻ mesoderm-derived population expressed pSHF progenitor markers, whereas the cells generated from the CD1d$^{low/−}$CXCR4⁺ population expressed those associated with aSHF progenitors (Figure S3H). Taken together, these findings demonstrate that it is possible to isolate mesoderm populations with FHF, aSHF, and pSHF potential based on expression of CXCR4, GYPB, CD1d, and ALDH activity.

Expression of ALDH1A2 in the day 5 hPSC-derived pSHF population suggested that RA signaling could play a role in the establishment and/or maintenance of the pSHF progenitor fate (Figure 3F). To test this, we treated isolated day 4 ALDH⁺ pSHF cells with either dimethyl sulfoxide (DMSO) or retinol (ROH) and examined the expression of cardiac progenitor markers 24 h later. RT-qPCR analyses showed that treatment with ROH led to an upregulation of expression of the RA responsive gene CYP26A1 and the pSHF progenitor markers HOXA1 and HOXB1 (Figure 3I) and to a downregulation of expression of the aSHF progenitor markers FGF10, FOXC1, and FOXC2 (Figure 3J). These findings indicate that RA signaling at this stage enforces the pSHF molecular signature in the pSHF progenitors.

Single-cell transcriptomic analyses of day 6 progenitor populations

We next performed scRNA-seq on the day 6 populations generated from the day 4 mesoderm fractions and identified clusters that represent FHF, aSHF, and pSHF progenitors based on the expression of the lineage-specific markers (Figures 4A and 4B). Expression of the cardiac myosin gene MYH6 in the majority (85%) of the cells in the FHF population indicates that they are undergoing differentiation toward CMs. Significantly fewer MYH6⁺ cells were detected in the aSHF cluster (26%), and almost none were present the pSHF cluster (4.6%) (Figure S4A). These differences recapitulate the temporal pattern observed in the mouse cardiac crescent where the FHF lineage differentiates

to CMs prior to those of SHF lineages (van den Berg et al., 2009; Rochais et al., 2009). Analyses of these MYH6⁺ cells revealed that they retain lineage-specific markers (Figure S4B), highlighting molecular differences between the earliest-emerging CMs derived from these distinct progenitors.

We next analyzed the hPSC-derived and E7.75 mouse progenitor populations for conserved gene expression patterns and identified 116 for the FHF, 103 for the aSHF, and 71 for the pSHF (Figures 4C and 4D; Table S4). Among these, we found genes known to be differentially expressed in these populations as well as those have not been previously shown to differ, including S100A10 and CSRP2 expression in the FHF progenitors; RGS5, JAG1, and IRX1 expression in the aSHF progenitors; and MEIS3 and CPE expression in the pSHF progenitors (Figure 4D; Table S4).

GO analyses based on the species-conserved expression patterns showed that genes associated with cardiac ventricle development and cardiac right ventricle morphogenesis were expressed in the FHF and aSHF clusters, respectively (Figure 4E). These patterns are consistent with the development potential of the FHF and SHF lineages established in mouse. Genes implicated in OFT morphogenesis, pharyngeal development, and Wnt signaling were expressed in the aSHF progenitors, whereas those involved in RA signaling were expressed in the pSHF cluster, recapitulating the lineage-specific patterns observed in the mesoderm populations.

Although RA signaling is a well-established regulator of the pSHF-derived lineages, we found that the aSHF progenitors also expressed components of this pathway, including RDH10, ALDH1A2, and RARB (Figure S4C). Pseudotime ordering showed that the pSHF lineage expressed high levels of ALDH1A2 from the day 3 mesoderm stage onward, whereas expression in the aSHF lineage was not upregulated until day 4 (Figures 2F and 4G). These patterns correlate with ALDH activity, which was detected in more than 90% of the day 4 and 5 pSHF cells; by contrast, the day 4 aSHF cells were ALDH⁻, and less than 50% of the day 5 aSHF population showed this activity (Figures 3C, 4H, and 4I). Treatment of the day 5 aSHF cells with ROH for 24 h led to the upregulation of CYP26A1 and HOXB1 (Figure 4J), genes known to be expressed in the SV and OFT progenitors (Bertrand et al., 2011; Roux et al., 2015). DEG and GO analyses of the aSHF- and pSHF-ALDH1A2⁺ populations (Figure S4D) showed that they displayed molecular profiles indicative of aSHF and pSHF progenitors (Figures S4E and S4F). Additionally, these analyses revealed that genes associated with OFT and right ventricle morphogenesis, pharyngeal system development, and BMP signaling pathway were enriched in the ALDH1A2⁺ aSHF population (Figure S4F). Re-clustering of the aSHF cells demonstrated the presence of a large

Figure 3. Generation of cardiac progenitor populations from isolated mesoderm

(A) The top 15 genes whose expression patterns positively and negatively correlate with GYPB, CXCR4, and ALDH1A2; Pearson correlation; p < 0.05.

(B and C) Flow cytometric analyses of CXCR4, PDGFRα, and CD235a/b expression and ALDH activity in the day 4 populations induced with 16B8A (B) or 3B1.5A (C).

(D) Separation of mesoderm subtypes and analyses of their derivatives.

(E–G) RT-qPCR analyses of (E) FHF, (F) pSHF, and (G) aSHF progenitor marker expression in the day 5 populations (n ≥ 5). One-way ANOVA.

(H) Immunostaining analyses of the day 6 populations.

(I and J) RT-qPCR analyses of expression of (I) RA-responsive genes, and (J) aSHF genes in the day 5 pSHF cells treated with either DMSO (D) or ROH (2 uM) (R). Unpaired t test. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. Error bars represent SEM.
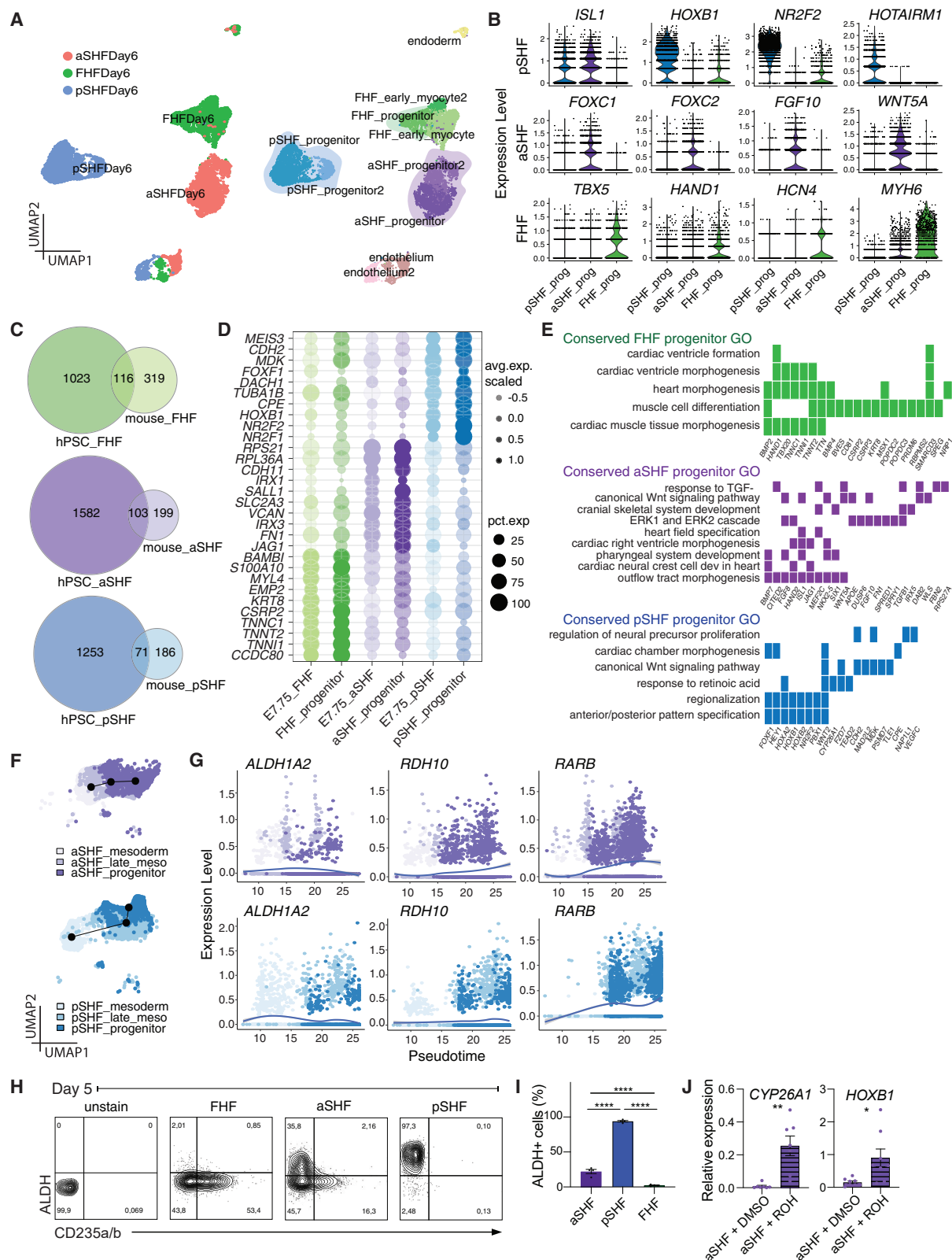
Figure 4. Transcriptomic analyses of day 6 progenitor populations
(A) UMAP of day 6 aSHF, pSHF, and FHF populations labeled by sample name (left) and cell identity (right).
(B) Violin plots showing expression of aSHF, pSHF, and FHF progenitor genes in the day 6 clusters.

(legend continued on next page)

ALDH1A2[+/low] cluster (cluster 0) that expressed the aSHF progenitor markers and a *ALDH1A2*[−] cluster (cluster 1) that expressed genes indicative of CM differentiation such as *TNNT2* and *NKX-2.5* (Figure S4G). Analyses of the E7.75 mouse aSHF progenitors identified a comparable *Aldh1a2*[+] cluster (E7.75_aSHF2) that expressed aSHF progenitor markers as well as an *Aldh1a2*[low] cluster (E7.75_aSHF1) that expressed CM genes (Figure S4H). Taken together, these findings strongly suggest that *ALDH1A2* expression in the aSHF lineage marks the progenitor stage of development.

The significance of BMP signaling in the aSHF progenitors is supported by gene set enrichment analysis (GSEA) that showed an enrichment of "response to Bmp" and "regulation of Bmp signaling pathway" in these progenitors and a corresponding downregulation of these pathways in the other progenitors (Figure 5A). Although most of the genes related to BMP signaling were upregulated in the aSHF progenitors (Figure 5B), the ligands (*BMP2*, *BMP4*, and *BMP5*) were more highly expressed in the FHF lineage (Figure 5B). These patterns fit with findings from a cell communication analysis showing that FHF progenitors are the predominant "sender" and aSHF cells are the "receiver" in the BMP signaling network (Figure 5C).

To determine whether BMP plays a role in the differentiation of the aSHF progenitors, we treated the day 5 progenitors with either a pathway agonist (10 ng/mL BMP4) or antagonist (0.1 μM LDN or 100 ng/mL Noggin) and then analyzed the population 24 h later. FHF and pSHF progenitors were included in some of the analyses, and ROH was added to pSHF (days 4–8) and aSHF (days 5–8) progenitors given that both populations display ALDH activity. As shown in Figures 5D and 5E, manipulation of the BMP pathway did not alter the expression of genes indicative of aSHF (*FGF10* and *ISL1*), pSHF (*HOXB1*), and FHF (*HAND1*) fates in the aSHF population. Inhibition of the pathway with either antagonist did, however, lead to an upregulation of expression of the pharyngeal markers *TBX1*, *FOXC2*, and *SIX1* as well as a downregulation of genes involved in CM development (*UNC45B* and *GATA4*) (Figure 5E) (Malek Mohammadi et al., 2017; Myhre et al., 2014; Rudeck et al., 2016). To further investigate the consequences of BMP manipulation, we extended the agonist/antagonist treatment for an additional 48 h and then cultured the treated cells for 12 days to promote CM development. FHF and pSHF progenitors were also cultured for 20 days without manipulation of BMP signaling. All three progenitor populations generated cTNT[+]/SIRPa[+] CMs by day 20 (Figures 5F and 5G). The aSHF-derived population also contained a PDGFRβ[+] mesenchymal cell subpopulation (Figure 5F). Treatment with BMP4 increased the proportion and number of CMs while reducing the number of PDGFRβ[+] cells; by contrast, inhibition of the pathway reduced the proportion of CMs (Figures 5F–5H and S4I). The CMs generated from the FHF

and aSHF progenitors expressed the VCM markers *IRX4* and *MYL2*, whereas those generated from the pSHF progenitors expressed the atrial marker *NR2F2* (Figure 5I). Consistent with the flow cytometry analyses (Figures 5F and 5G), treatment of the aSHF cells with LDN or Noggin resulted in a decrease in expression of the VCM markers and elevated levels of *PDGFRB* (Figure 5I). Taken together, these findings show that the FHF and aSHF progenitors generate VCMs, whereas the pSHF progenitors give rise to CMs that express atrial markers. Additionally, they demonstrate that efficient generation of VCMs from the aSHF progenitors is dependent on BMP signaling.

Transcriptional profiles of FHF-, aSHF-, and pSHF-derived CMs

ScRNA-seq analyses of the day 20 CMs generated from the FHF, aSHF, and pSHF progenitors revealed developmental trajectories similar to those described in the mouse. Analyses of the *TNNT2*[+] clusters showed that the FHF progenitors gave rise to *MYL2*[high]*HAND1*[high]*IRX1*[low] LV-like CMs (LVLCMs) and *MYL2*[+]*BMP2*[high]*TBX2*[high]*MSX2*[high] AVC-like CMs (AVCLCMs). The population derived from the aSHF progenitors contained clusters representing *MYL2*[high]*HAND1*[−]*IRX1*[high] RV-like CMs (RVLCMs) and *MYL2*[+]*HAND2*[+]*SEMA3C*[+] OFT-like CMs (OFTLCMs), whereas the pSHF progenitors gave rise to atrial-like CMs (ALCMs) (*NR2F2*[+]*TBX18*[−]*CAV1*[+]*NKX2-5*[+]) and SV-like CMs (SVLCMs) (*NR2F2*[+]*TBX18*[+]*CAV1*[−]*NKX2-5*[−]) (Figures 6A and 6B; Table S5).

We next compared the data from the hPSC-derived CMs to published mouse E9.25 heart data (Figure S5A) (de Soysa et al., 2019) to identify species-conserved expression patterns for each of the CM subtypes. For these analyses, we paired CM subtypes derived from the same heart field lineage including ACM versus SV (pSHF), RV versus OFT (aSHF), and AVC versus LV (FHF). Additionally, we compared the LVCM and RVCM populations. Our analysis of the aSHF derivatives identified 94 RVCM markers (*NPPB*, *PLN*, *IRX1*, and others) and 112 OFTCM markers (*CFC1*, *FHL1*, *SEMA3C*, and others) that are shared by humans and mice (Figure 6C; Table S5). GO analysis indicated that OFTLCMs expressed elevated levels of genes related to OFT morphogenesis and cardiac neural crest cell development as well as those associated with transforming growth factor β (TGF-β) signaling (Figure 6D). These findings are relevant because TGF-β pathway plays a pivotal role in endocardial cushion formation in the OFT region (Azhar et al., 2009; Brown et al., 1999; Combs and Yutzey, 2009), and neural crest cells participate in the maturation and remodeling of OFT CMs and valvular cells (Jain et al., 2011; Waldo et al., 1999; Yelbuz et al., 2003). As expected, RVLCMs were enriched for genes associated with VCM development and function including myofibril assembly and muscle contraction (Figure 6D). Analysis of the

---

(C) FHF, aSHF, and pSHF progenitor genes displaying conserved and non-conserved expression patterns.

(D) Progenitor genes that show species-conserved expression patterns.

(E) Species-conserved genes and GO terms enriched in the indicated progenitors.

(F) Slingshot pseudotime inference of day 3, day 4, and day 6 populations.

(G) Expression of the indicated genes in the aSHF (purple) and pSHF (blue) lineages along pseudotime.

(H) Flow cytometric analyses of ALDH activity and CD235a/b expression in the day 5 populations.

(I) Percentage of ALDH[+] cells in the day 5 populations (n = 4). One-way ANOVA. (J) Expression of *CYP26A1* and *HOXB1* in the day 6 aSHF cells treated with ROH or DMSO (n ≥ 5). Unpaired t test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$. Error bars represent SEM.
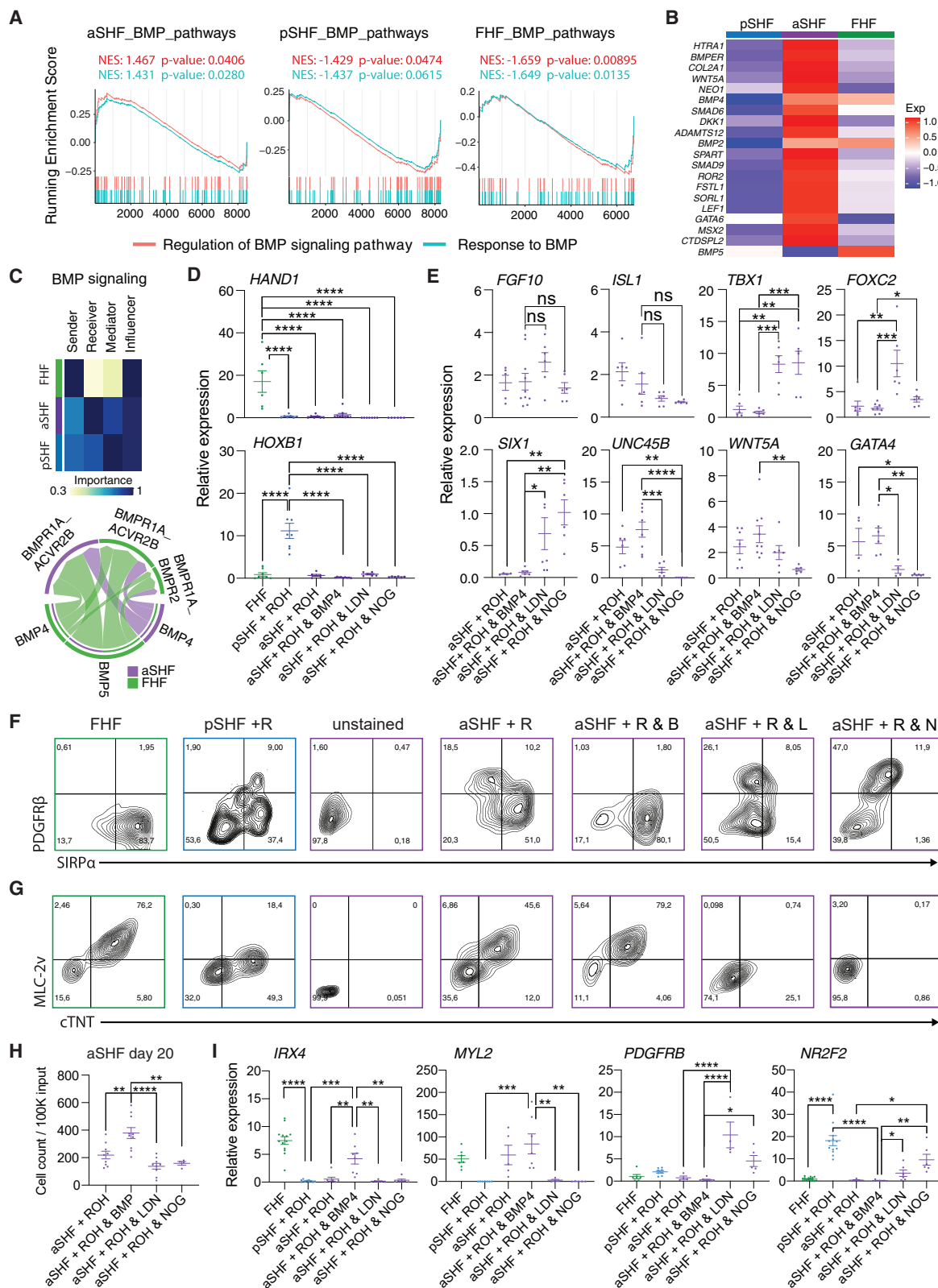
Figure 5. Bmp signaling is required for ventricular CM differentiation of aSHF progenitors
(A) GSEA of the indicated pathways in the day 6 progenitors.
(B) Expression of genes related to Bmp signaling in the progenitor clusters.

(legend continued on next page)

pSHF derivatives revealed 142 conserved ACM and 190 SVCM markers (Figure S5B; Table S5). GO analysis indicated that ALCMs expressed genes involved in cardiac muscle contraction and chamber development, whereas SVLCMs were enriched for neural crest cell migration and synapsis related genes (Figure S5B). Analysis of the FHF derivatives identified 122 conserved AVCCM and 147 LVCM markers (Figure S5C;Table S5). GO analysis showed that LVLCMs expressed genes involved in translation and transcription, whereas AVCLCMs expressed genes associated with endocardial cushion development, atrium development, and cardiac conduction (Figure S5C), defining characteristics of AVCCMs (Aanhaanen et al., 2009; Bakker et al., 2010; Cai et al., 2011; de Lange et al., 2004; de Vlaming et al., 2012). Finally, comparison between LVCMs and RVCMs confirmed differential expression of known chamber-specific markers including *HAND1* and *TBX5* in LVCMs and *IRX1*, *PLN*, and *NPPB* in RVCMs (Figure 6E; Table S5) (de Soysa et al., 2019). GO analysis based on the 145 LV and 64 RV markers indicated that RNA-processing-related genes were highly expressed in LVLCMs, whereas those associated with calcium and calcineurin signaling were preferentially expressed in RVLCMs (Figure 6F). Lastly, GSEA analyses demonstrated that the RVLCMs were enriched for genes associated with arrhythmogenic right ventricular cardiomyopathy (ARVC) (Figure S5D), a disease that predominantly affects the right ventricle (Azaouagh et al., 2011; Lombardi and Marian, 2010).

RT-qPCR analyses confirmed differential expression of the identified LV and RV genes in the FHF- and aSHF-derived CMs (Figures 6H, 6I, and S5E). Genes associated with ACMs and SVCMs were highly expressed in the pSHF-CMs (Figure 6J). Immunostaining analyses supported this lineage assignment and showed that the FHF-CMs expressed higher levels of HAND1, TBX5, and GJA1 than did the aSHF-CMs; in contrast, the aSHF derivatives expressed higher levels of IRX1 (Figures 6G, S5F, and S5G). HEY2, a compact VCM marker, was expressed in the FHF and aSHF but not in the pSHF derivatives (Figures S5F and S5G). Analyses of human gestation week 15 fetal heart samples confirmed the differential expression of IRX1 and HAND1 in the LV and RV cells, respectively (Figures S5H and S5I).

Integration of the data from the day 20 CM populations with a published human fetal heart dataset (Cui et al., 2019) revealed a significant overlap between hPSC-derived VCMs and fetal VCMs, hPSC-derived ACMs and fetal ACMs, and hPSC-derived mesenchyme and fetal valvular cells (Figure 6K). The overlapping VCMs (*TNNT2*⁺*MYL2*⁺) formed clusters 5 and 6, whereas atrial CMs (*NR2F2*⁺*MYL2*⁻) formed cluster 2 (Figures 6K, S6A, and S6B). A ML-based classification using the human fetal heart data as reference suggested that the transcriptomes of

the hPSC-derived LVLCMs and RVLCMs align with those of LV and RV CMs from gestation week 7 fetal heart tissue (Figures 6L and S6C).

Access to transcriptomes of hPSC-derived FHF, aSHF, and pSHF mesoderm; progenitors; and CMs identified in this study provides an unprecedented opportunity to compare our populations to those isolated from primary tissue or to those generated from hPSCs using different cell lines or protocols. With these comparisons, we found that *LGR5*, a defining marker of human fetal CVPs (Sahara et al., 2019) was preferentially expressed in the aSHF progenitors (Figure S6D). GSEA analyses further showed an enrichment of the CVP gene set in the aSHF progenitors and that of the FVP gene set in the FHF progenitors (Figure S6E), demonstrating that the hPSC-derived progenitors share molecular features with anatomically distinct populations in the human fetal heart.

To determine whether the method of induction can influence cardiovascular development, we next analyzed hPSC-derived populations generated with the small-molecule CHIR protocol (Friedman et al., 2018). ML models predicted that most of the day 2 CHIR-induced *MESP1*⁺*SOX17*⁻ mesoderm cells represent the aSHF lineage (Figure S6F). The day 5 population was more heterogeneous as it was predicted to contain *ISL1*⁺ aSHF progenitors and some *HAND1*⁺ FHF progenitors (Figure S6G). In line with the early mesoderm profiles, most of the day 30 *IRX4*⁺⁻*PDGFRB*⁻ VCMs were classified as RVLCMs (Figure S6H). This RV assignment is supported by the integration of the BMP/ACT- and CHIR-induced *MYL2*⁺ VCMs (Figure S6I), which showed that the overlapping VCM clusters (clusters 1 and 2) were primarily composed of CHIR-induced VCMs and the BMP/ACT-RVLCMs (Figures S6J and S6K). Very few CHIR-induced VCMs were detected in cluster 4, which contains most of the BMP/ACT- LVLCMs (Figure S6K). A subpopulation of VCMs generated with the CHIR protocol did not cluster with our day 20 CM data, suggesting they could represent cell types not generated with the conditions used in our study.

Electrophysiological characterization of hPSC-derived cardiomyocytes

To further characterize the CMs generated from different heart fields, we used optical imaging to analyze their electrophysiological properties. Analyses of day 20 spontaneously beating CMs showed that those generated from the pSHF had the fastest beating rate, those from the FHF showed an intermediate rate, and those from the aSHF had the slowest rate (Figures S7A and S7B). Comparison of action potential durations measured at 50% and 90% repolarization (APD50 and APD90) from spontaneously beating cells revealed that the aSHF-derived CMs had higher values than the CMs generated from the other two lineages (Figure S7B). When the day 20 aSHF- and FHF-CMs were

---

(C) Network centrality measures of the BMP signaling network in each progenitor group (upper); the inferred BMP signaling network in the FHF and aSHF progenitors (lower).

(D and E) RT-qPCR analyses of expression of the indicated genes in the day 6 progenitors treated as shown (n ≥ 4).

(F and G) Flow cytometric analyses of expression of (F) PDGFRβ and SIRPα or (G) MLC-2v and cTNT in the day 20 aSHF, pSHF, and FHF populations. R, ROH; B, BMP4; L, LDN; N, noggin.

(H) Cell numbers in the day 20 aSHF populations.

(I) RT-qPCR analyses of expression of the indicated genes in the day 20 populations (n ≥ 5). One-way ANOVA. Error bars represent SEM *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
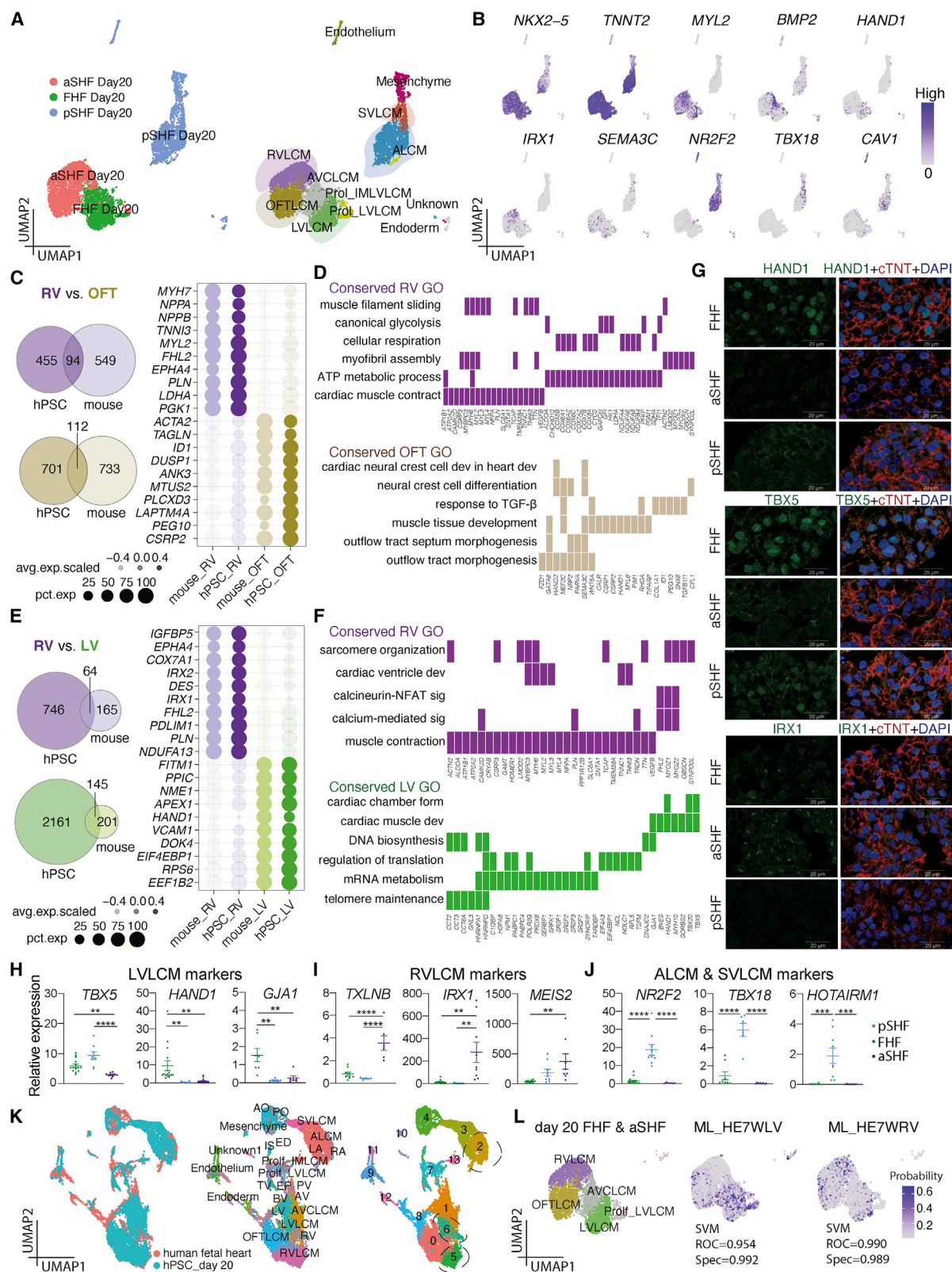
Figure 6. Transcriptomic analyses of the FHF-, aSHF-, and pSHF-derived CMs

(A) UMAP of day 20 FHF, aSHF, and pSHF populations labeled by sample name (left) and cell type (right).

(B) UMAP showing the expression of genes indicative of different CM subtypes.

(legend continued on next page)

paced at 0.5Hz (Figure 7A), the former still displayed longer APD50 and APD90 than the latter (Figure 7B). PSHF-CMs were excluded from the pacing analysis given their fast spontaneous beating rate (approximately 3 Hz).

One factor that can contribute to these differences is the state of maturation of the cells. Comparison of the VCM populations revealed that the RVLCMs expressed gene sets associated with maturation such as sarcomere organization, fatty acid (FA) metabolic maturation, and calcium signaling (Figure 7C) (Feyen et al., 2020; Funakoshi et al., 2021). These cells also expressed elevated levels of *CACNA1D*, an L-type Ca^{2+} channel that generates a depolarizing current (Molina et al., 2016; Pinggera and Striessnig, 2016) (Figures 7A, S7A, and S7D). The FHF-CMs, by contrast, expressed low levels of these genes and contained a subpopulation of proliferating cells indicative of an immature stage (Figures 7C and S7C). These differences are in line with the observation that LV cells from fetal heart tissue at 7 weeks of gestation expressed higher levels of genes involved in proliferation than the RV cells (Figure S7E). To promote maturation, the FHF- and aSHF-CMs were cultured as aggregates in the base media from days 20–35. RT-qPCR analyses showed that the expression patterns of the LV (*HAND1* and *TBX5*) and RV (*IRX1*) genes did not change appreciably over this time, indicating the chamber specificity of the CMs remained stable. In addition to these chamber-specific genes, we also observed higher expression levels of *KCND3*, a K^+ channel involved in transient outward current (I_{to}) (Patel and Campbell, 2005), in the day 35 aSHF-CMs than those observed in the FHF-CMs (Figure 7D). This differential expression in LV and RV CMs has also been observed in the mouse (Martin et al., 2012). Optical imaging showed that the day 35 FHF- and aSHF-CMs displayed VCM-like action potentials and that the APD50 and APD90 increased in the FHF-CMs to levels comparable to those in the aSHF-CMs (Figures 7E and 7F). RT-qPCR analyses revealed comparable expression levels of *CACNA1D* and *CACNA1C* in these VCMs (Figure S7F), suggesting a maturation process following the proliferation phase in the FHF population.

Development of the FHF, aSHF, and pSHF lineages from HES3 and H1 hPSCs

To determine whether the approach for modeling FHF, aSHF, and pSHF development is applicable to other hPSC lines, we next investigated the generation of these lineages from HES3-NKX2-5eGFP/w hESCs and H1 hESCs. Titration of BMP4 and Activin A concentrations between days 1 and 3 of differentiation revealed that 15B8A and 15B6A were optimal to induce FHF mesoderm from the HES3 and H1 lines, respectively (Figure S7G), whereas 3B1A (HES3) and 1B0A (H1) induced mesoderm that expressed pSHF and aSHF genes (Figure S7H).

Flow cytometric analyses of the day 4 populations showed patterns comparable to those observed with the HES2 cells. For both lines, the majority of the FHF PDGFRα$^+$ population was CD235a/b$^+$CXCR4$^{-/low}$ALDH$^-$, whereas the SHF PDGFRα$^+$ population contained CXCR4$^+$ALDH$^-$ and CXCR4$^-$ALDH$^+$ fractions (Figures S7I and S7J). When isolated and cultured for 24 h, the FHF mesoderm gave rise to FHF progenitors, whereas the CXCR4$^+$ALDH$^-$ and CXCR4$^-$ALDH$^+$ fractions generated aSHF and pSHF progenitors, respectively (Figure S7K). Further culture of the FHF, aSHF, and pSHF progenitors resulted in the development of CMs from each population. Analyses of the day 20 CMs showed that the FHF-CMs expressed LVCM genes, the aSHF-CMs expressed RVCM genes, and the pSHF-CMs expressed ACM and SVCM genes (Figure S7L). Collectively, these findings indicate that the strategy to model the development of FHF, aSHF, and pSHF lineages can be translated to other hPSC lines.

DISCUSSION

Studies in the mouse have provided compelling evidence that the CM subtypes that make up the adult heart develop from three distinct lineages known as the FHF, the aSHF, and the pSHF (Brade et al., 2013; Buckingham et al., 2005). In this study, we have recapitulated FHF, aSHF, and pSHF development from hPSCs and showed that the human lineages develop from distinct mesoderm populations that transition through a defined progenitor stage to give rise to different CMs (Figure 7G). Comparisons of the molecular profiles of the mouse and human populations uncovered conserved transcriptional patterns that we used to define the hPSC-derived FHF, aSHF, and pSHF mesoderm; progenitor; and CM stages of development. Analyses of the derivative CM populations revealed that the human lineages largely recapitulate the developmental trajectories of those of the mouse, enabling the generation of populations that display molecular characteristics of LVCMs, AVCCMs, RVCMs, ACMs, OFTCMs, and SVCMs.

Our findings that human heart field lineages develop from different mesoderm populations are consistent with those in the mouse, which showed that the FHF and SHF lineages are induced and segregate during gastrulation (Lescroart et al., 2014; Devine et al., 2014). Although the pathways regulating these early developmental steps *in vivo* have not been well characterized, our findings provide strong evidence that different levels of activin/nodal and BMP signaling play a pivotal role in the induction of the FHF and SHF mesoderm. Studies in the mouse embryo show that FHF mesoderm expresses high levels of BMP signaling-related genes (Lescroart et al., 2018), and our observation that this population expresses elevated levels of

(C) RVCM and OFTCM genes displaying conserved and non-conserved expression patterns (left); top 10 species-conserved RVCM and OFTCM markers (right).
(D) Species-conserved genes and GO terms enriched in the RVCMs and OFTCMs.
(E) RVCM and LVCM genes displaying conserved and non-conserved expression patterns (left); top 10 species-conserved RVCM and LVCM markers (right).
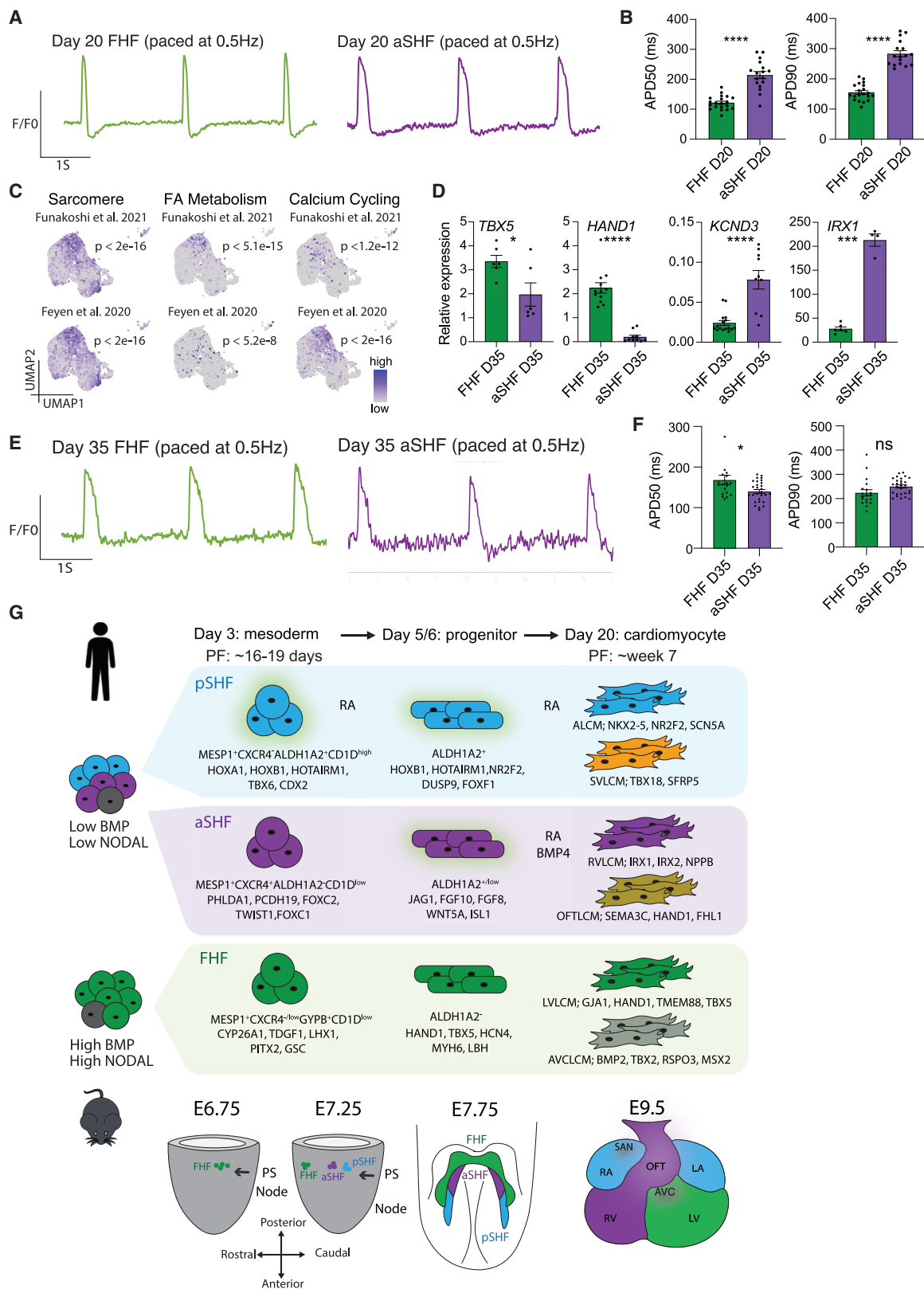(F) Species-conserved genes and GO terms enriched in the RVCMs and LVCMs.
(G) Immunostaining analyses of HAND1, IRX1, and TBX5 in the day 20 CMs.
(H–J) RT-qPCR analyses of expression of (H) LVLCM genes, (I) RVLCM genes, and (J) ALCM and SVLCM genes in the day 20 populations (n ≥ 5). One-way ANOVA. Error bars represent SEM *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
(K) Integrated human fetal and hPSC-derived cardiac cells labeled by data source (left), cell identity (middle), and cluster number (right).
(L) UMAP showing the probability of classification.

components of nodal/activin signaling compared to those seen in the aSHF and pSHF mesoderm (Figures S1D and S1E) suggest that these pathways play a role in the generation of this subpopulation of mesoderm *in vivo*. In addition to these pathways, our analyses reveal an enrichment of components of the Wnt pathway in the SHF lineages (Figures 2H and 2I). Although we did not manipulate the Wnt pathway in this study, a role in human SHF development is supported by comparative analyses of our profiles to those of CHIR-induced populations, which suggest that early stage Wnt signaling preferentially induces aSHF mesoderm and derivative RVLCMs (Figures S6F and S6H). Although primarily assigned as RVLCMs, the CHIR-CMs were found to express *HAND1* (not shown). This observation could reflect low levels of expression of *HAND1* in RVLCMs generated from this hPSC line, as we have observed that HES3- and H1-derived RVLCMs also express low levels of this gene (Figure S7L). Further analyses of populations produced with different protocols and hPSC lines using the stage- and lineage-specific molecular signatures identified in this study will provide a clearer picture of the lineages generated with different methods.

Although lineage tracing and retrospective studies established the lineage relationship between distinct mesoderm and CM subtypes, the transition from mesoderm to cardiovascular progenitors remains largely uncharacterized. Through our ability to isolate the hPSC-derived mesoderm subpopulations, we formally established lineage-specific mesoderm-progenitor relationships and showed that CD235a/b$^+$ mesoderm gives rise to FHF progenitors, CXCR4$^+$ALDH$^-$ mesoderm generates aSHF progenitors, and CXCR4$^-$ALDH$^+$ (or CXCR4$^-$CD1D^{high}) mesoderm differentiates to pSHF progenitors. Our analyses of the progenitor populations revealed that the aSHF and pSHF progenitors express *ALDH1A2*, indicating that RA signaling plays a role in both lineages (Figures 4G–4J). The observation that the *ALDH1A2*$^+$ aSHF progenitors express genes involved in pharyngeal and OFT development is consistent with findings from lineage tracing experiments in mouse demonstrating that OFT progenitors express *Raldh2* and that RA-activated cells contribute to OFT development (Dollé et al., 2010). Our analyses also revealed that the aSHF progenitors express components of the BMP pathway. Based on this observation, we were able to demonstrate that signaling through this pathway is required for the generation of VCMs from these progenitors (Figure 5). These findings are supported by mouse studies showing that BMP signaling, possibly initiated by Bmp2 or Bmp4 secreted from the adjacent endoderm, is essential for the myocardial differentiation of aSHF progenitors (Hutson et al., 2010, 2010, 2010; Tirosh-Finkel et al., 2010).

Access to isolated populations of mesoderm has enabled us to track the origin of the different human CM subtypes and to establish a developmental map of the human heart field line-

ages. Our findings show that the human lineages display developmental potential similar to those defined in the mouse (Figure 7G). Although previous studies have described the development of ventricular, atrial, and sinoatrial pacemaker cells from hPSCs (Cyganek et al., 2018; Devalla et al., 2015; Lee et al., 2017; Protze et al., 2017; Scavone et al., 2013), methods for generating RVCMs, LVCMs, OFTCMs, and AVCCMs have not yet been reported. Distinguishing CM subtypes, in particular, those from left versus right chambers such as LVCMs and RVCMs in the absence of heart structure, can be challenging. Our designation of LVCMs versus RVCMs is based on the demonstration that these CMs develop from different subpopulations of mesoderm, that they show species-conserved LV and RV gene expression patterns, and that their molecular profiles align with those of the LV and RV CMs from fetal hearts at 7 weeks of gestation (Figure 6L). The ability to generate distinct populations of LVCMs and RVCMs from hPSCs is important for both cell therapy and disease modeling applications. For example, LVCMs are likely the best cell type for transplantation for remuscularization of an infarcted region of the left ventricle, whereas RVCMs would be the appropriate population for modeling ARVC.

In conclusion, by using precise stage-specific induction strategies and extensive transcriptomic analyses, we have identified and characterized the human equivalent of the FHF, aSHF, and pSHF cardiac lineages that together establish a comprehensive map of human cardiovascular development. These findings will establish the basis for optimizing protocols for the generation of a broad spectrum of human CM subtypes, providing unprecedented opportunities for detailed genetic and epigenetic studies on human cardiac development, for modeling cardiovascular diseases that target specific regions of the heart, and for developing cell-based therapies with appropriate chamber specific populations.

Limitations of study

In this study, we address the transcriptomic similarities between the FHF, aSHF, and pSHF populations generated from hPSCs and the corresponding populations in the mouse embryo as a first step in establishing *in vitro* models of human heart field development. However, we do acknowledge the importance of investigating human- and mouse-specific features which could shed light on differences in the heart fields between species. In addition, findings shown here were collected from cells generated from hESCs; further studies using iPSCs could provide more compelling evidence that the methods we established are broadly applicable. Lastly, future work including investigation of additional signaling pathways important for pSHF and aSHF development is needed to improve the efficiency of CM differentiation from these lineages.

Figure 7. Characterization of electrophysiological properties of CM subtypes
(A) Imaging of optical action potentials (oAPs) of day 20 FHF- and aSHF-CMs (0.5 Hz).
(B) APD50 and APD90 of day 20 FHF- and aSHF-CMs (0.5 Hz). Unpaired t test.
(C) Average expression levels of the indicated gene sets. Wilcoxon test.
(D) RT-qPCR analyses of expression of the indicated genes in the day 35 aSHF- and FHF-CMs (n ≥ 5). Unpaired t test.
(E) Imaging of oAPs of the day 35 FHF- and aSHF-CMs (0.5 Hz).
(F) APD50 and APD90 of the day 35 FHF- and aSHF-CMs (0.5 Hz). Unpaired t test. Error bars represent SEM *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
(G) Schematic illustration of human and mouse FHF, aSHF, and pSHF development from the mesoderm to CM stages.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- ● KEY RESOURCES TABLE
- ● RESOURCE AVAILABILITY
  - ○ Lead contact
  - ○ Materials availability
  - ○ Data and code availability
- ● EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - ○ Human embryonic stem cell lines
  - ○ Human fetal heart samples
- ● METHOD DETAILS
  - ○ Directed differentiation of human pluripotent stem cells
  - ○ Flow cytometry
  - ○ Aldefluor assay
  - ○ Quantitative reverse transcription PCR
  - ○ Immunocytochemistry and imaging
  - ○ Imaging of Optical action potentials
  - ○ Sample preparation, single-cell library generation and raw data processing
  - ○ Cell filtering and cell-type clustering analysis
  - ○ Batch correction and integration of multiple datasets
  - ○ Gene regulatory network analysis
  - ○ Gene ontology, gene set enrichment analyses, and human to mouse gene transformation
  - ○ Pseudotime trajectory analysis
  - ○ Machine learning-based classification
  - ○ Intercellular communication analysis
- ● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stem.2022.08.007.

AUTHOR CONTRIBUTIONS

D.Y. designed the project, performed experiments, analyzed all the scRNA-seq data, and wrote the manuscript. G.M.K. designed the project and wrote the manuscript. J.G.G. and M.A.L. performed and analyzed optical imaging experiments. S.F. and I.F. performed immunostaining experiments. T.T. and G.B. analyzed the mouse mesoderm scRNA-seq data.

DECLARATION OF INTERESTS

G.M.K. is a scientific founder and paid consultant for BlueRock Therapeutics, LP, a paid consultant for VistaGen Therapeutics, and a board member of Anagenesis Biotechnologies. M.A.L. is a paid consultant for BlueRock Therapeutics, LP. G.M.K. and D.Y. declare a patent titled "Cardiomyocyte subtypes and methods of making and using" (no. PCT/IB2022/055,425) related to this work.

REFERENCES

Aanhaanen, W.T., Brons, J.F., Domínguez, J.N., Rana, M.S., Norden, J., Airik, R., Wakker, V., de Gier-de Vries, C., Brown, N.A., Kispert, A., et al. (2009). The Tbx2⁺ Primary Myocardium of the Atrioventricular Canal Forms the Atrioventricular Node and the Base of the Left Ventricle. Circ. Res. *104*, 1267–1274. https://doi.org/10.1161/CIRCRESAHA.108.192450.

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. *104*, 9319–9324. https://doi.org/10.1073/pnas.0701212104.

Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat. Methods *14*, 1083–1086. https://doi.org/10.1038/nmeth.4463.

Alquicira-Hernandez, J., Sathe, A., Ji, H.P., Nguyen, Q., and Powell, J.E. (2019). scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. Genome Biol. *20*, 264. https://doi.org/10.1186/s13059-019-1862-5.

Andersen, P., Tampakakis, E., Jimenez, D.V., Kannan, S., Miyamoto, M., Shin, H.K., Saberi, A., Murphy, S., Sulistio, E., Chelko, S.P., and Kwon, C. (2018). Precardiac organoids form two heart fields via Bmp/Wnt signaling. Nat. Commun. *9*, 3140. https://doi.org/10.1038/s41467-018-05604-8.

Anderson, R.H., Mori, S., Spicer, D.E., Brown, N.A., and Mohun, T.J. (2016). Development and Morphology of the Ventricular Outflow Tracts. World J. Pediatr. Congenit. Heart Surg. *7*, 561–577. https://doi.org/10.1177/2150135116651114.

Azaouagh, A., Churzidse, S., Konorza, T., and Erbel, R. (2011). Arrhythmogenic right ventricular cardiomyopathy/dysplasia: a review and update. Clin. Res. Cardiol. Off. J. Ger. Card. Soc. *100*, 383–394. https://doi.org/10.1007/s00392-011-0295-2.

Azhar, M., Runyan, R.B., Gard, C., Sanford, L.P., Miller, M.L., Andringa, A., Pawlowski, S., Rajan, S., and Doetschman, T. (2009). Ligand-specific function of transforming growth factor beta in epithelial-mesenchymal transition in heart development. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *238*, 431–442. https://doi.org/10.1002/dvdy.21854.

Bakker, M.L., Moorman, A.F., and Christoffels, V.M. (2010). The Atrioventricular Node: Origin, Development, and Genetic Program. Trends Cardiovasc. Med. *20*, 164–171. https://doi.org/10.1016/j.tcm.2011.02.001.

Barnes, R.M., Firulli, B.A., Conway, S.J., Vincentz, J.W., and Firulli, A.B. (2010). Analysis of the Hand1 Cell Lineage Reveals Novel Contributions to Cardiovascular, Neural Crest, Extra-Embryonic, and Lateral Mesoderm Derivatives. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *239*, 3086–3097. https://doi.org/10.1002/dvdy.22428.

van den Berg, G., Abu-Issa, R., de Boer, B.A., Hutson, M.R., de Boer, P.A., Soufan, A.T., Ruijter, J.M., Kirby, M.L., van den Hoff, M.J., and Moorman, A.F. (2009). A Caudal Proliferating Growth Center Contributes to Both Poles of the Forming Heart Tube. Circ. Res. *104*, 179–188. https://doi.org/10.1161/CIRCRESAHA.108.185843.

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. https://doi.org/10.1016/j.ydbio.2011.02.029.

Brade, T., Pane, L.S., Moretti, A., Chien, K.R., and Laugwitz, K.-L. (2013). Embryonic Heart Progenitors and Cardiogenesis. Cold Spring Harb. Perspect. Med. *3*, a013847. https://doi.org/10.1101/cshperspect.a013847.

Brown, C.B., Boyer, A.S., Runyan, R.B., and Barnett, J.V. (1999). Requirement of type III TGF-beta receptor for endocardial cell transformation in the heart. Science *283*, 2080–2082. https://doi.org/10.1126/science.283.5410.2080.

Bruneau, B.G., Logan, M., Davis, N., Levi, T., Tabin, C.J., Seidman, J.G., and Seidman, C.E. (1999). Chamber-Specific Cardiac Expression of Tbx5 and Heart Defects in Holt–Oram Syndrome. Dev. Biol. *211*, 100–108. https://doi.org/10.1006/dbio.1999.9298.

Buckingham, M., Meilhac, S., and Zaffran, S. (2005). Building the mammalian heart from two sources of myocardial cells. Nat. Rev. Genet. *6*, 826–835. https://doi.org/10.1038/nrg1710.

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. https://doi.org/10.1016/s1534-5807(03)00363-0.

Cai, X., Nomura-Kitabayashi, A., Cai, W., Yan, J., Christoffels, V.M., and Cai, C.-L. (2011). Myocardial Tbx20 regulates early atrioventricular canal formation and endocardial epithelial-mesenchymal transition via Bmp2. Dev. Biol. *360*, 381–390. https://doi.org/10.1016/j.ydbio.2011.09.023.

Carlson, M.R.J., Pagès, H., Arora, S., Obenchain, V., and Morgan, M. (2016). Genomic Annotation Resources in R/Bioconductor. In Statistical Genomics: Methods and Protocols, E. Mathé and S. Davis, eds. (New York, NY: Springer), pp. 67–90.

Ciruna, B., and Rossant, J. (2001). FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. Dev. Cell *1*, 37–49. https://doi.org/10.1016/S1534-5807(01)00017-X.

Cohen, E.D., Wang, Z., Lepore, J.J., Lu, M.M., Taketo, M.M., Epstein, D.J., and Morrisey, E.E. (2007). Wnt/β-catenin signaling promotes expansion of Isl-1–positive cardiac progenitor cells through regulation of FGF signaling. J. Clin. Invest. *117*, 1794–1804. https://doi.org/10.1172/JCI31731.

Combs, M.D., and Yutzey, K.E. (2009). Heart valve development: regulatory networks in development and disease. Circ. Res. *105*, 408–421. https://doi.org/10.1161/CIRCRESAHA.109.201566.

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. https://doi.org/10.1016/j.celrep.2019.01.079.

Cyganek, L., Tiburcy, M., Sekeres, K., Gerstenberg, K., Bohnenberger, H., Lenz, C., Henze, S., Stauske, M., Salinas, G., Zimmermann, W.-H., Hasenfuss, G., and Guan, K. (2018). Deep phenotyping of human induced pluripotent stem cell–derived atrial and ventricular cardiomyocytes. JCI Insight *3*, 99941. https://doi.org/10.1172/jci.insight.99941.

de Lange, F.J., Moorman, A.F., Anderson, R.H., Männer, J., Soufan, A.T., Vries, C d G d, Schneider, M.D., Webb, S., van den Hoff, M.J., and Christoffels, V.M. (2004). Lineage and Morphogenetic Analysis of the Cardiac Valves. Circ. Res. *95*, 645–654. https://doi.org/10.1161/01.RES.0000141429.13560.cb.

Devalla, H.D., Schwach, V., Ford, J.W., Milnes, J.T., El-Haou, S., Jackson, C., Gkatzis, K., Elliott, D.A., Chuva de Sousa Lopes, S.M., Mummery, C.L., et al. (2015). Atrial-like cardiomyocytes from human pluripotent stem cells are a robust preclinical model for assessing atrial-selective pharmacology. EMBO Mol. Med. *7*, 394–410. https://doi.org/10.15252/emmm.201404757.

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. https://doi.org/10.7554/eLife.03848.

Dollé, P., Fraulob, V., Gallego-Llamas, J., Vermot, J., and Niederreither, K. (2010). Fate of retinoic acid–activated embryonic cell lineages. Dev. Dyn. *239*, 3260–3274. https://doi.org/10.1002/dvdy.22479.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biolog-

ical databases and microarray data analysis. Bioinforma. Oxf. Engl. *21*, 3439–3440. https://doi.org/10.1093/bioinformatics/bti525.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. *4*, 1184–1191. https://doi.org/10.1038/nprot.2009.97.

Elliott, D.A., Braam, S.R., Koutsis, K., Ng, E.S., Jenny, R., Lagerqvist, E.L., Biben, C., Hatzistavrou, T., Hirst, C.E., Yu, Q.C., et al. (2011). NKX2-5(eGFP/w) hESCs for isolation of human cardiac progenitors and cardiomyocytes. Nat. Methods *8*, 1037–1040. https://doi.org/10.1038/nmeth.1740.

Feyen, D.A., McKeithan, W.L., Bruyneel, A.A., Spiering, S., Hörmann, L., Ulmer, B., Zhang, H., Briganti, F., Schweizer, M., Hegyi, B., et al. (2020). Metabolic Maturation Media Improve Physiological Function of Human iPSC-Derived Cardiomyocytes. Cell Rep. *32*, 107925. https://doi.org/10.1016/j.celrep.2020.107925.

Francou, A., Saint-Michel, E., Mesbah, K., Théveniau-Ruissy, M., Rana, M.S., Christoffels, V.M., and Kelly, R.G. (2013). Second heart field cardiac progenitor cells in the early mouse embryo. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1833*, 795–798. https://doi.org/10.1016/j.bbamcr.2012.10.003.

Friedman, C.E., Nguyen, Q., Lukowski, S.W., Helfer, A., Chiu, H.S., Miklas, J., Levy, S., Suo, S., Han, J.-D.J., Osteil, P., et al. (2018). Single-Cell Transcriptomic Analysis of Cardiac Differentiation from Human PSCs Reveals HOPX-Dependent Cardiomyocyte Maturation. Cell Stem Cell *23*, 586–598.e8. https://doi.org/10.1016/j.stem.2018.09.009.

Funakoshi, S., Fernandes, I., Mastikhina, O., Wilkinson, D., Tran, T., Dhahri, W., Mazine, A., Yang, D., Burnett, B., Lee, J., et al. (2021). Generation of mature compact ventricular cardiomyocytes from human pluripotent stem cells. Nat. Commun. *12*, 3155. https://doi.org/10.1038/s41467-021-23329-z.

Galdos, F.X., Guo, Y., Paige, S.L., VanDusen, N.J., Wu, S.M., and Pu, W.T. (2017). Cardiac Regeneration: Lessons From Development. Circ. Res. *120*, 941–959. https://doi.org/10.1161/CIRCRESAHA.116.309040.

Gassanov, N., Er, F., Zagidullin, N., Jankowski, M., Gutkowska, J., and Hoppe, U.C. (2008). Retinoid acid-induced effects on atrial and pacemaker cell differentiation and expression of cardiac ion channels. Differ. Res. Biol. Divers. *76*, 971–980. https://doi.org/10.1111/j.1432-0436.2008.00283.x.

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

Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat. Biotechnol. *36*, 421–427. https://doi.org/10.1038/nbt.4091.

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. https://doi.org/10.1172/JCI38922.

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. Dev. Camb. Engl. *137*, 3001–3011. https://doi.org/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. Dev. Camb. Engl. *133*, 2435–2445. https://doi.org/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. https://doi.org/10.1371/journal.pbio.3001200.

Jain, R., Engleka, K.A., Rentschler, S.L., Manderfield, L.J., Li, L., Yuan, L., and Epstein, J.A. (2011). Cardiac neural crest orchestrates remodeling and functional maturation of mouse semilunar valves. J. Clin. Invest. *121*, 422–430. https://doi.org/10.1172/JCI44244.

Jin, S., Guerrero-Juarez, C.F., Zhang, L., Chang, I., Ramos, R., Kuan, C.-H., Myung, P., Plikus, M.V., and Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. Nat. Commun. 12, 1088. https://doi.org/10.1038/s41467-021-21246-9.

Jones, R.J., Barber, J.P., Vala, M.S., Collector, M.I., Kaufmann, S.H., Ludeman, S.M., Colvin, O.M., and Hilton, J. (1995). Assessment of Aldehyde Dehydrogenase in Viable Cells. Blood 85, 2742–2746. https://doi.org/10.1182/blood.V85.10.2742.bloodjournal85102742.

Kelly, R.G. (2012). Chapter two - The Second Heart Field. In Current Topics in Developmental Biology, B.G. Bruneau, ed. (Academic Press), pp. 33–65.

Kelly, R.G., and Buckingham, M.E. (2002). The anterior heart-forming field: voyage to the arterial pole of the heart. Trends Genet. 18, 210–216. https://doi.org/10.1016/S0168-9525(02)02642-2.

Kelly, R.G., Brown, N.A., and Buckingham, M.E. (2001). The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm. Dev. Cell 1, 435–440. https://doi.org/10.1016/s1534-5807(01)00040-5.

Kennedy, M., D'Souza, S.L., Lynch-Kattman, M., Schwantz, S., and Keller, G. (2007). Development of the hemangioblast defines the onset of hematopoiesis in human ES cell differentiation cultures. Blood 109, 2679–2687. https://doi.org/10.1182/blood-2006-09-047704.

Klaus, A., Saga, Y., Taketo, M.M., Tzahor, E., and Birchmeier, W. (2007). Distinct roles of Wnt/β-catenin and Bmp signaling during early cardiogenesis. Proc. Natl. Acad. Sci. USA 104, 18531–18536. https://doi.org/10.1073/pnas.0703113104.

Lee, J.H., Protze, S.I., Laksman, Z., Backx, P.H., and Keller, G.M. (2017). Human Pluripotent Stem Cell-Derived Atrial and Ventricular Cardiomyocytes Develop from Distinct Mesoderm Populations. Cell Stem Cell 21, 179–194.e4. https://doi.org/10.1016/j.stem.2017.07.003.

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. https://doi.org/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. https://doi.org/10.1126/science.aao4174.

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. https://doi.org/10.1161/CIRCRESAHA.113.301588.

Litviňuková, M., Talavera-López, C., Maatz, H., Reichart, D., Worth, C.L., Lindberg, E.L., Kanda, M., Polanski, K., Heinig, M., Lee, M., et al. (2020). Cells of the adult human heart. Nature 588, 466–472. https://doi.org/10.1038/s41586-020-2797-4.

Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. Nat. Genet. 22, 361–365. https://doi.org/10.1038/11932.

Lombardi, R., and Marian, A.J. (2010). Arrhythmogenic right ventricular cardiomyopathy is a disease of cardiac stem cells. Curr. Opin. Cardiol. 25, 222–228. https://doi.org/10.1097/HCO.0b013e3283376daf.

Macosko, E., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A., Kamitaki, N., Martersteck, E., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002.

Malek Mohammadi, M., Kattih, B., Grund, A., Froese, N., Korf-Klingebiel, M., Gigina, A., Schramck, U., Rudat, C., Liang, Q., Kispert, A., et al. (2017). The transcription factor GATA4 promotes myocardial regeneration in neonatal mice. EMBO Mol. Med. 9, 265–279. https://doi.org/10.15252/emmm.201606602.

Martin, C.A., Siedlecka, U., Kemmerich, K., Lawrence, J., Cartledge, J., Guzadhur, L., Brice, N., Grace, A.A., Schwiening, C., Terracciano, C.M., and

Huang, C.L.H. (2012). Reduced Na(+) and higher K(+) channel expression and function contribute to right ventricular origin of arrhythmias in Scn5a+/- mice. Open Biol 2, 120072. https://doi.org/10.1098/rsob.120072.

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

Meilhac, S.M., Esner, M., Kelly, R.G., Nicolas, J.-F., and Buckingham, M.E. (2004). The clonal origin of myocardial cells in different regions of the embryonic mouse heart. Dev. Cell 6, 685–698. https://doi.org/10.1016/s1534-5807(04)00133-9.

Mishina, Y., Suzuki, A., Ueno, N., and Behringer, R.R. (1995). Bmpr encodes a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. Genes Dev. 9, 3027–3037. https://doi.org/10.1101/gad.9.24.3027.

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. https://doi.org/10.1006/dbio.2001.0409.

Molina, C.E., Heijman, J., and Dobrev, D. (2016). Differences in Left Versus Right Ventricular Electrophysiological Properties in Cardiac Dysfunction and Arrhythmogenesis. Arrhythmia Electrophysiol. Rev. 5, 14. https://doi.org/10.15420/aer.2016.8.2.

Myhre, J.L., Hills, J.A., Jean, F., and Pilgrim, D.B. (2014). Unc45b is essential for early myofibrillogenesis and costamere formation in zebrafish. Dev. Biol. 390, 26–40. https://doi.org/10.1016/j.ydbio.2014.02.022.

Nevis, K., Obregon, P., Walsh, C., Guner-Ataman, B., Burns, C.G., and Burns, C.E. (2013). Tbx1 is required for second heart field proliferation in zebrafish. Dev. Dyn. Off. Publ. Am. Assoc. Anat. 242, 550–559. https://doi.org/10.1002/dvdy.23928.

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. Dev. Camb. Engl. 133, 2419–2433. https://doi.org/10.1242/dev.02367.

Patel, S.P., and Campbell, D.L. (2005). Transient outward potassium current, 'Ito', phenotypes in the mammalian left ventricle: underlying molecular, cellular and biophysical mechanisms. J. Physiol. 569, 7–39. https://doi.org/10.1113/jphysiol.2005.086223.

Pezhouman, A., Engel, J.L., Nguyen, N.B., Skelton, R.J.P., Gilmore, W.B., Qiao, R., Sahoo, D., Zhao, P., Elliott, D.A., and Ardehali, R. (2021). Isolation and characterization of human embryonic stem cell-derived heart field-specific cardiomyocytes unravels new insights into their transcriptional and electrophysiological profiles. Cardiovasc. Res. 118, 828–843. https://doi.org/10.1093/cvr/cvab102.

Pinggera, A., and Striessnig, J. (2016). Cav 1.3 (CACNA1D) L-type Ca2+ channel dysfunction in CNS disorders. J. Physiol. 594, 5839–5849. https://doi.org/10.1113/JP270672.

Protze, S.I., Liu, J., Nussinovitch, U., Ohana, L., Backx, P.H., Gepstein, L., and Keller, G.M. (2017). Sinoatrial node cardiomyocytes derived from human pluripotent cells function as a biological pacemaker. Nat. Biotechnol. 35, 56–68. https://doi.org/10.1038/nbt.3745.

Rankin, S.A., Steimle, J.D., Yang, X.H., Rydeen, A.B., Agarwal, K., Chaturvedi, P., Ikegami, K., Herriges, M.J., Moskowitz, I.P., and Zorn, A.M. (2021). TBX5 Drives Aldh1a2 Expression to Regulate a RA-Hedgehog-Wnt Gene Regulatory Network Coordinating Cardiopulmonary Development.

Rochais, F., Mesbah, K., and Kelly, R.G. (2009). Signaling pathways controlling second heart field development. Circ. Res. 104, 933–942. https://doi.org/10.1161/CIRCRESAHA.109.194464.

Roux, M., Laforest, B., Capecchi, M., Bertrand, N., and Zaffran, S. (2015). Hoxb1 regulates proliferation and differentiation of second heart field progenitors in pharyngeal mesoderm and genetically interacts with Hoxa1 during cardiac outflow tract development. Dev. Biol. 406, 247–258. https://doi.org/10.1016/j.ydbio.2015.08.015.

Rudeck, S., Etard, C., Khan, M.M., Rottbauer, W., Rudolf, R., Strähle, U., and Just, S. (2016). A compact unc45b-promoter drives muscle-specific

expression in zebrafish and mouse. Genes. N. Y. N *54*, 431–438. https://doi.org/10.1002/dvg.22953.

Ryan, K., and Chin, A.J. (2003). T-box genes and cardiac development. Birth Defects Res. Part C Embryo Today - Rev. *69*, 25–37. https://doi.org/10.1002/bdrc.10001.

Sahara, M., Santoro, F., Sohlmér, J., Zhou, C., Witman, N., Leung, C.Y., Mononen, M., Bylund, K., Gruber, P., and Chien, K.R. (2019). Population and Single-Cell Analysis of Human Cardiogenesis Reveals Unique LGR5 Ventricular Progenitors in Embryonic Outflow Tract. Dev. Cell *48*, 475–490.e7. https://doi.org/10.1016/j.devcel.2019.01.005.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502. https://doi.org/10.1038/nbt.3192.

Scavone, A., Capilupo, D., Mazzocchi, N., Crespi, A., Zoia, S., Campostrini, G., Bucchi, A., Milanesi, R., Baruscotti, M., Benedetti, S., et al. (2013). Embryonic stem cell-derived CD166+ precursors develop into fully functional sinoatrial-like cells. Circ. Res. *113*, 389–398. https://doi.org/10.1161/CIRCRESAHA.113.301283.

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. https://doi.org/10.1038/s41586-019-1414-x.

Später, D., Abramczuk, M.K., Buac, K., Zangi, L., Stachel, M.W., Clarke, J., Sahara, M., Ludwig, A., and Chien, K.R. (2013). A HCN4+ cardiomyogenic progenitor derived from the first heart field and human pluripotent stem cells. Nat. Cell Biol. *15*, 1098–1106. https://doi.org/10.1038/ncb2824.

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

Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genom. *19*, 477. https://doi.org/10.1186/s12864-018-4772-0.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., et al. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888–1902, In press. https://doi.org/10.1016/j.cell.2019.05.031.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

Thomas, P.S., Rajderkar, S., Lane, J., Mishina, Y., and Kaartinen, V. (2014). AcvR1-mediated BMP signaling in second heart field is required for arterial pole development: Implications for myocardial differentiation and regional identity. Dev. Biol. *390*, 191–207. https://doi.org/10.1016/j.ydbio.2014.03.008.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. Science *282*, 1145–1147. https://doi.org/10.1126/science.282.5391.1145.

Tirosh-Finkel, L., Zeisel, A., Brodt-Ivenshitz, M., Shamai, A., Yao, Z., Seger, R., Domany, E., and Tzahor, E. (2010). BMP-mediated inhibition of FGF signaling promotes cardiomyocyte differentiation of anterior heart field progenitors. Development *137*, 2989–3000. https://doi.org/10.1242/dev.051649.

Tyser, R.C.V., Mahammadov, E., Nakanoh, S., Vallier, L., Scialdone, A., and Srinivas, S. (2021). Single-cell transcriptomic characterization of a gastrulating

human embryo. Nature *600*, 285–289. https://doi.org/10.1038/s41586-021-04158-y.

de Vlaming, A., Sauls, K., Hajdu, Z., Visconti, R.P., Mehesz, A.N., Levine, R.A., Slaugenhaupt, S.A., Hagège, A., Chester, A.H., Markwald, R.R., and Norris, R.A. (2012). Atrioventricular valve development: new perspectives on an old theme. Differ. Res. Biol. Divers. *84*, 103–116. https://doi.org/10.1016/j.diff.2012.04.001.

Waldo, K., Zdanowicz, M., Burch, J., Kumiski, D.H., Stadt, H.A., Godt, R.E., Creazzo, T.L., and Kirby, M.L. (1999). A novel role for cardiac neural crest in heart development. J. Clin. Invest. *103*, 1499–1507. https://doi.org/10.1172/JCI6501.

Wiesinger, A., Boink, G.J., Christoffels, V.M., and Devalla, H.D. (2021). Retinoic acid signaling in heart development: Application in the differentiation of cardiovascular lineages from human pluripotent stem cells. Stem Cell Rep. *16*, 2589–2606. https://doi.org/10.1016/j.stemcr.2021.09.010.

Winnier, G., Blessing, M., Labosky, P.A., and Hogan, B.L. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev. *9*, 2105–2116. https://doi.org/10.1101/gad.9.17.2105.

Yang, Y.-P., Li, H.-R., Cao, X.-M., Wang, Q.-X., Qiao, C.-J., and Ya, J. (2013). Second heart field and the development of the outflow tract in human embryonic heart. Dev. Growth Differ. *55*, 359–367. https://doi.org/10.1111/dgd.12050.

Yelbuz, T.M., Waldo, K.L., Zhang, X., Zdanowicz, M., Parker, J., Creazzo, T.L., Johnson, G.A., and Kirby, M.L. (2003). Myocardial volume and organization are changed by failure of addition of secondary heart field myocardium to the cardiac outflow tract. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *228*, 152–160. https://doi.org/10.1002/dvdy.10364.

Yilbas, A.E., Hamilton, A., Wang, Y., Mach, H., Lacroix, N., Davis, D.R., Chen, J., and LI, Q. (2014). Activation of GATA4 gene expression at the early stage of cardiac specification. Front. Chem. *2*, 12. https://doi.org/10.3389/fchem.2014.00012.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS A J. Integr. Biol. *16*, 284–287. https://doi.org/10.1089/omi.2011.0118.

Zhang, J.Z., Termglinchan, V., Shao, N.-Y., Itzhaki, I., Liu, C., Ma, N., Tian, L., Wang, V.Y., Chang, A.C., Guo, H., et al. (2019). A Human iPSC Double-Reporter System Enables Purification of Cardiac Lineage Subpopulations with Distinct Function and Drug Response Profiles. Cell Stem Cell *24*, 802–811.e5. https://doi.org/10.1016/j.stem.2019.02.015.

Zhang, Q., Jiang, J., Han, P., Yuan, Q., Zhang, J., Zhang, X., Xu, Y., Cao, H., Meng, Q., Chen, L., et al. (2011). Direct differentiation of atrial and ventricular myocytes from human embryonic stem cells by alternating retinoid signals. Cell Res. *21*, 579–587. https://doi.org/10.1038/cr.2010.163.

Zhang, X., Lian, Z., Padden, C., Gerstein, M.B., Rozowsky, J., Snyder, M., Gingeras, T.R., Kapranov, P., Weissman, S.M., and Newburger, P.E. (2009). A myelopoiesis-associated regulatory intergenic noncoding RNA transcript within the human HOXA cluster. Blood *113*, 2526–2534. https://doi.org/10.1182/blood-2008-06-162164.

Zhao, Y., Wang, W., Guan, C., Hu, Z., Liu, L., Li, W., and Jiang, X. (2020). Long noncoding RNA HOTAIRM1 in human cancers. Clin. Chim. Acta *511*, 255–259. https://doi.org/10.1016/j.cca.2020.10.011.

Zhou, P., He, A., and Pu, W.T. (2012). Regulation of GATA4 Transcriptional Activity in Cardiovascular Development and Disease. In Current Topics in Developmental Biology (Elsevier), pp. 143–169.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal to SIRPα (clone SE5A5), PeCy7 conjugated	Biolegend	Cat.# 323,808; RRID: AB_1236443
Mouse monoclonal to CD140b (PDGFRβ) (clone 28D4), BV421 conjugated	BD Horizon	Cat.# 564,124; RRID: AB_2738609
Mouse monoclonal to PDGFRα, PE conjugated	BD Pharmingen	Cat.# 556,002; RRID: AB_396286
Mouse monoclonal to CD235a/b (GYPB) (clone HIR2), APC conjugated	BD PharMingen	Cat.# 551,336; RRID: AB_398499
Mouse monoclonal to human CD49C (ITGA3), APC conjugated	ThermoFisher	Cat.# 17-0494-42; RRID: AB_10670221
Mouse monoclonal to human CD1D, PE conjugated	BioLegend	Cat.# 350,306; RRID: AB_10641845
Mouse monoclonal to CTNT (clone 13-11), non-conjugated	ThermoFisher	Cat.# MS-295-P; RRID: AB_11000742
Rabbit polyclonal to MLC2V (clone 13-11), non-conjugated	Abcam	Cat.# ab79935; RRID: AB_1952220
Monoclonal mouse ant-Human CD184 (CXCR4) Antibody	BioLegend	Cat.# 306,518; RRID: AB_11146018
Polyclonal Rabbit anti-Human HAND1 Antibody	LSBio	Cat. # LS-C118745; RRID: AB_10796650
Polyclonal Rabbit anti- TBX5 Antibody	ThermoFisher	Cat. # 42-6500; RRID: AB_2533533
Monoclonal Mouse anti-Human NR2F2 Antibody	R&D systems Bio-Techne	Cat. # PP-H7147-00; RRID: AB_2155627
Monoclonal Mouse anti-Human NR2F1 antibody	R&D systems Bio-Techne	Cat. # PP-H8132-00; RRID: AB_2155494
Polyclonal Rabbit anti-Human IRX1 Antibody	ThermoFisher	Cat. # PA5-60261; RRID: AB_2642783
Polyclonal Rabbit anti-Human HEY2 Antibody	Proteintech	Cat. # 10597-1-AP; RRID: AB_2118415
Polyclonal Rabbit anti-Human CX43 (GJA1) Antibody	Abcam	Cat. #AB11370; RRID: AB_297976
Goat anti-mouse IgG (H + L), APC conjugated	BD Pharmingen	Cat.# 550,826; RRID: AB_398465
Donkey anti-rabbit IgG (H + L), PE conjugated	Jackson ImmunoResearch	Cat.# 711-116-152 RRID: AB_2340599
Donkey anti-rabbit IgG (H + L), Alexa Fluor 488	ThermoFisher	Cat.# A21206; RRID: AB_2535792
Donkey anti-mouse IgG (H + L), Alexa Fluor 488	ThermoFisher	Cat.# A21202; RRID: AB_141607
Donkey anti-rabbit IgG (H + L), Alexa Fluor 555	ThermoFisher	Cat.# A31572; RRID: AB_162543
Donkey anti-mouse IgG (H + L), Alexa Fluor 555	ThermoFisher	Cat.# A31570; RRID: AB_2536180
Chemicals, peptides, and recombinant proteins		
Penicillin/streptomycin	ThermoFisher	Cat.# 15,070,063
Potassium Chloride	Sigma	Cat.#P9541-1KG

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
D-(+)-Glucose	Sigma	Cat.#G8270-1KG
Calcium Chloride	Sigma	Cat.#C1016-500G
Sodium Chloride	Sigma	Cat.#S7653-1KG
Magnesium Chloride	Sigma	Cat.# 208,337-1KG
Sodium Phosphate Monobasic monohydrate	Sigma	Cat.#S9G38-250G
HEPES	Sigma	Cat.#H3375-250G
Sodium Hydroxide	Sigma	Cat.#S5881-500G
Blebbistatin	Sigma	Cat.B0560-5MG
L-glutamine	ThermoFisher	Cat.# 25,030,081
Non-essential amino acids	ThermoFisher	Cat.# 11,140-050
Transferrin	ROCHE	Cat.# 10,652,202,001
Ascorbic acid	Sigma	Cat.# A4544-100G
Monothioglycerol	Sigma	Cat.# M-6145-25mL
β -Mercaptoethanol	ThermoFisher	Cat.# 21,985-023
ROCK inhibitor Y-27632	Tocris	Cat.# 1254
Recombinant human BMP4	R&D	Cat.# 314-BP
Recombinant human ActivinA	R&D	Cat.# 338-AC
Recombinant human bFGF	R&D	Cat.# 233-FB
Recombinant human VEGF	R&D	Cat.# 293-VE
Retinol	Sigma	Cat.#R7632
Retinoic acid	Sigma	Cat.#R2625
IWP2 (Wnt inhibitor)	Tocris	Cat.# 3533
LDN 193189 (BMP inhibitor)	Tocris	Cat.# 6053
Collagenase type2	Worthington	Cat.# LS004-176
DAPI	Biotium	Cat.# 40,043
Fetal calf serum (FCS)	Wisent	Cat.# 080-150
BSA	Sigma	Cat.# A1470
Noggin	R&D	Cat.# 3344-NG
Critical commercial assays		
RNAqueous-micro kit with RNase-free Dnase treatment	Ambion	Cat.# AM1931
Fluovolt Membrane Potential Kit	ThermoFisher	Cat.#F10488
Superscript III Reverse Transcriptase kit	ThermoFisher	Cat.# 18,080,044
QuantiFast SYBR Green PCR kit	QIAGEN	Cat.# 204,057
Aldefluor assay kit	STEMCELL Technologies	Cat.# 1700
10x Genomics Single Cell 3′ v3.1 Reagent Kit	10x Genomics	Single Cell 3′ v3.1
Deposited data		
mouse mesoderm scRNA-seq	Lescroart et al., 2018	GSE100471
hPSC scRNA-seq	Friedman et al., 2018	E-MTAB-6268
human gastrulation scRNA-seq	Tyser et al., 2021	E-MTAB-9388
mouse E7.75 – E9.25 cardiac cell scRNA-seq	de Soysa et al., 2019	GSE126128
human fetal heart scRNA-seq	Cui et al., 2019	GSE106118
hPSC-derived cardiac mesoderm (day 3), late mesoderm (day 4), progenitor (day 6) and cardiomyocyte (day 20) scRNA-seq	This study	GSE173486
Experimental models: Cell lines		
Human ESC: HES3-NKX2-5eGFP/w line	Gift from Drs. E. Stanley and A. Elefanty, Monash University, AU (Elliott et al., 2011)	NA

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human ESC: HES2 line	WiCell	Cat.# ES02
Human ESC: H1 line	Jones lab (Thomson et al., 1998)	
Oligonucleotides		
PCR primer sequences	This study	Table S1
Software and algorithms		
FlowJo	Tree Star	https://www.flowjo.com
MATLAB vR2020b	MathWorks	https://www.mathworks.com/products/new_products/release2020b.html
Metamorph v7.8.2.0	Molecular Devices LLC	
FV10-ASW	Olympus	https://www.olympus-lifescience.com
GraphPad Prism	GraphPad Software	http://www.graphpad.com/scientific-software/prism
10x Chromium Single Cell Software Suite v3.1		https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
R v3.6.1		https://cran.r-project.org/
Bioconductor v3.10		https://bioconductor.org/news/bioc_3_10_release/
Batchelor	Haghverdi et al., 2018	
Seurat v3.2.2	Stuart et al., 2019	https://github.com/satijalab/seurat
Seurat-wrappers		https://github.com/satijalab/seurat-wrappers
CellChat v0.0.2	Jin et al., 2021	https://github.com/sqjin/CellChat
ScPred v1.9.0	Alquicira-Hernandez et al., 2019	https://github.com/powellgenomicslab/scPred
ClusterProfiler v3.18	Yu et al., 2012	https://github.com/YuLab-SMU/clusterProfiler
MsigDB v7.2.1		http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp
Enrichplot v1.10.2		https://yulab-smu.top/biomedical-knowledge-mining-book/
Slingshot v1.4.0	Street et al., 2018	https://github.com/kstreet13/slingshot
BiomaRt v2.42.1	Durinck et al., 2009	https://m.ensembl.org/info/data/biomart/biomart_r_package.html
Org.Mm.eg.db		https://bioconductor.org/packages/release/data/annotation/html/org.Mm.eg.db.html
Org.Hs.eg.db		https://bioconductor.org/packages/release/data/annotation/html/org.Hs.eg.db.html
SingleCellExperiment v1.8.0		https://github.com/drisso/SingleCellExperiment
Ggplot2		https://ggplot2.tidyverse.org/
dplyr		https://dplyr.tidyverse.org/
SCENIC v1.1.2.1	Aibar et al., 2017	https://scenic.aertslab.org/
RcisTarget v1.11.10	Aibar et al., 2017	https://scenic.aertslab.org/
AUCell v1.13.3	Aibar et al., 2017	https://scenic.aertslab.org/
Other		
Human fetal heart	Lunenfeld-Tanenbaum Research Institute	UHN#2018-0370
StemPro-34 media (kit)	ThermoFisher	Cat.# 10,639,011
DMEM/F12	Cellgro	Cat.# 10-092-CV

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Hanks Buffered Salt solution	GIBCO	Cat.# 14,175-079
KnockOut serum replacement	ThermoFisher	Cat.# 10,828,028
TrypLE	ThermoFisher	Cat.# 12,605,010
96-well clear flat bottom TC-treated culture microplate	Falcon	Cat.# 353,072
24-well clear flat bottom TC-treated culture microplate	Falcon	Cat.# 353,047
96-well clear flat bottom ultra-low attachment microplate	Corning	Cat.# 3474
Petri dishes 60 × 15 mm	VWR	Cat.# 25,384-090
Micro cover glasses 12mm	VWR	Cat.# 89,015-725
ProLong diamond antifade mountant	ThermoFisher	Cat.#P36965

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Dr. Gordon M. Keller (Gordon.Keller@uhnresearch.ca).

Materials availability
HPSC lines used in this study can be shared upon request following submission of a Material Transfer Agreement.

Data and code availability
ScRNA-seq datasets (raw files and processed count matrices) collected in this study have been deposited at GEO under the accession number provided in the key resources table.

This paper does not report original code. This study uses referenced sources of code that can be found in the vignettes of the cited packages. Details of the packages are provided in the key resources table and the parameters are provided in the method details section. Any additional information required to reanalyze the data reported in this paper is available from the Lead Contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human embryonic stem cell lines
The HES3-NKX2-5$^{eGFP/w}$ cell line (karyotype: 46, XX) was generated by targeting eGFP-encoding sequences to the NKX2-5 locus of HES3 cells using previously described protocol (Elliott et al., 2011). The HES2 cell line (karyotype: 46, XX) was purchased from WiCell. The H1 cell line (karyotype: 46, XX) was generated using previously described protocol (Thomson et al., 1998).The hPSC lines were maintained on irradiated mouse embryonic fibroblasts in hPSC culture media consisting of DMEM/F12 (Cellgro), penicillin/strepto-mycin (1%, ThermoFisher), L-glutamine (2 mM, ThermoFisher), non-essential amino acids (1x, ThermoFisher), b-Mercaptoethanol (55 mM, ThermoFisher) and Knock-Out serum replacement (20%, ThermoFisher) as described previously (Kennedy et al., 2007).

Human fetal heart samples
The fetal heart tissue samples (sample number 31499; gestation week 15; sex unknown; sample size = 1) were obtained from the Research Center for Women's and Infants Health (RCWIH) BioBank program at the Lunenfeld-Tanenbaum Research Institute (UHN #2018-0370). Use of the human tissue samples for this study was approved by the Research Ethics Board of the University Health Network, Toronto, ON, Canada. Informed consent was obtained from all tissue donors by Novogenix Laboratories or RCWIH.

METHOD DETAILS

Directed differentiation of human pluripotent stem cells
HPSCs were differentiated using our previously published embryoid body (EB)-based protocol (Lee et al., 2017). Briefly, at day 0, hPSCs at 80% confluency were dissociated into single cells (TrypLE, ThermoFisher) and aggregated into embryonic bodies (EBs) by 20 h of culture on an orbital shaker (70rpm). For this EB formation step, the cells were cultured in 6 cm Petri-grade dishes, each well containing 4 mL of media (StemPro-34 media (ThermoFisher) supplemented with 1% penicillin/streptomycin (ThermoFisher), 2mM L-glutamine (ThermoFisher), 150 μg/mL transferrin (ROCHE), 50 μg/mL ascorbic acid (Sigma) and 50 μg/mL

monothioglycerol (Sigma)) (base media). 10 μM ROCK inhibitor Y-27632 (TOCRIS) and 1 ng/ml rhBMP4 (R&D) were added to the media for EB formation. Cultures were incubated in a low oxygen environment (5% CO2, 5%O2, 90% N2). At day 1, the EBs were transferred to mesoderm induction media consisting of base media, 5 ng/mL rhbFGF (R&D), and various concentration of rhBMP4 (R&D) and rhActivinA (R&D) as described in the results section. Ebs were cultured (static) in 6 well plates in a low oxygen environment. At day 3, the Ebs were transferred to the media consisting of base media, 2 μM Wnt inhibitor IWP2 (TOCRIS) and 10 ng/mL rhVEGF (R&D) and from day 5 to day 12, they were cultured in base media with 5 ng/mL rhVEGF (low oxygen). From day 12 to day 20, the Ebs were cultured in base media and in a normoxic environment (5% CO2, 20% O2). For pSHF lineage differentiation, the sorted day 4 ALDH$^+$ mesoderm cells were cultured in ultra-low cluster 96 well plates in base media containing 1 μM IWP2 (TOCRIS) and 5 ng/mL rhVEGF (R&D). ROH (2 μM; Sigma) was added from day 4 to day 8. PSHF-CMs used for optical imaging analysis were generated by treating the unsorted ALDH$^+$ mesoderm (3B0.25A) with 500 nM RA, 1 μM IWP2 (TOCRIS), and 5 ng/mL rhVEGF (R&D) from days 3–5 of differentiation as described by Lee et al. (2017). The sorted aSHF mesoderm was cultured in ultra-low cluster 96 well plates in the media containing 1 μM IWP2 (TOCRIS) and 5 ng/mL rhVEGF (R&D); in addition, ROH (2 μM) and rhBMP4 (10 ng/mL rhBMP4) were added from day 5 to day 8 of aSHF differentiation.

Flow cytometry

Early stage Ebs (day 3 to day 6) were dissociated to single cells by treatment with TrypLE for 3–5 min at room temperature. The resulting cell suspension was filtered and transferred to IMDM media for staining. Day 20 Ebs were dissociated by treatment with 0.5 mg/ml collagenase type 2 (Worthington) in HANKs buffer for 1.5 h at 37 °C. The cells were filtered and then transferred to FACS buffer consisting of PBS with 5% fetal calf serum (Wisent) and 0.02% sodium azide for staining. The following antibodies were used for staining cells obtained from various stages of differentiation: anti-PDGFRa-PE (R&D Systems, 1:20), anti-CD235a/b-APC (BD PharMingen, 1:200), anti-SIRPa-PeCy7 (Biolegend, 1:2000), anti-CXCR4 (Biolegend, 1:100), anti-CD1d-PE (Biolegend, 1:100), anti-CD49c-PE (ThermoFisher, 1:100), anti-cardiac isoform of CTNT (ThermoFisher Scientific, 1:2000), or anti-myosin light chain 2 (Abcam, 1:1000). For unconjugated primary antibodies, the following secondary antibodies were used for detection: goat anti-mouse IgG-APC (BD Pharmigen, 1:250), or donkey anti-rabbit IgG-PE (Jackson ImmunoResearch, 1:250). Detailed antibody information is described in the Key Resources Table. To stain live cells with antibodies against cell-surface proteins, the dissociated single cells were stained for 15 min at room temperature in FACS buffer and washed twice before they were subject to further analyses. For intracellular staining, cells were fixed for 15 min at 4 °C in PBS containing 4% PFA followed by permeabilization using 90% methanol for 15 min at 4 °C. The permeabilized cells were washed twice with PBS containing 0.5% BSA (Sigma) and stained with unconjugated primary antibodies in FACS buffer for 18 h at 4 °C. The cells were then washed with PBS with 0.5% BSA and stained with secondary antibodies in FACS buffer for 30 min at 4 °C. Following the washing steps, the stained cells were analyzed on a Fortessa (BD) analyzer. The data were analyzed using FlowJo software (Tree Star). For cell sorting, stained cells were kept in StemPro-34 media and sorted using Influx (BD), FACSAriaII (BD), MoFlo-XDP (BD) or FACSAria Fusion (BD) sorters. Data were analyzed using FlowJo software (Tree Star).

Aldefluor assay

The aldefluorTM assay (STEMCELL Technologies) was used to detect aldehyde dehydrogenase (ALDH) activity within the cells. For the analyses, cells were dissociated using the method described in 'Flow cytometry' and then stained in the aldefluor assay buffer containing 0.1% BSA and BAAA substrate (0.12 mg/mL) for 40 min at 37 °C. This step was done in an environment free of light. The aldehyde dehydrogenase inhibitor DEAB (0.75nM) was added to one group of cells as the negative control. To stop the reaction, cells were washed with cold wash media consisting of IMDM and 10% aldefluor assay buffer. During analyses, the cells were kept in cold wash media. For sorting ALDH$^+$ SHF mesoderm, stained cells were maintained in cold StemPro-34 containing 10% aldefluor assay buffer throughout the sorting process. The sorted cells were collected and re-aggregated in base media containing 1 mM IWP2 and 5 ng/mL rhVEGF.

Quantitative reverse transcription PCR

RNA extraction was performed using the RNAqueous-micro-Kit (invitrogen). Purified RNA was treated with Rnase-free Dnase (invitrogen), and then reverse transcribed into cDNA using oligo (dT) primers and random hexamers and iscript Reverse Transcriptase (ThermoFisher). RT-qPCR was performed on an EP Real- Plex MasterCycler (Eppendorf) using the QuantiFast SYBR Green PCR kit (QIAGEN). The copy number of each gene was determined based on a standard curve generated using human genomic DNA. The relative expression levels of genes were obtained by normalizing their copy number to that of *TBP*. The detailed primer sequences are provided in the Table S1.

Immunocytochemistry and imaging

EBs were fixed in 4% paraformaldehyde, embedded, and sectioned. After deparaffinization and rehydration, heat-induced epitope retrieval was performed followed by antibody staining. The following antibodies were used for staining: mouse anti-cardiac isoform of cTNT (ThermoFisher Scientific, 1:200), rabbit anti-human cTNT (abcam, 1:200), rabbit anti-CX43 (abcam, 1:800), rabbit anti-TBX5 (ThermoFisher, 1:200), rabbit anti-HAND1 (LSBio, 1:200), rabbit anti-HEY2 (Proteintech, 1:200), mouse anti-NR2F2 (Bio-Techne, 1:200), mouse anti-NR2F1 (Bio-Techne, 1:200), rabbit anti-IRX1 (ThermoFisher Scientific, 1:200). For detecting unconjugated primary antibodies, the following secondary antibodies were used: donkey anti-mouse IgG-Alexa488 (ThermoFisher, 1:500), donkey

anti-rabbit IgG-Alexa488 (ThermoFisher, 1:500), donkey anti-mouse IgG-Alexa555 (ThermoFisher, 1:500), or donkey anti-rabbit IgG-Alexa555 (ThermoFisher, 1:500). IRX1 and CX43 (GJA1) expression was measured by counting the number of CX43$^+$ cells in one field of view (×40 magnification) of cTNT$^+$ cardiomyocytes randomly selected from 1 area in each EB. Data were collected from 3 independent experiments. All images were captured using a Zeiss LSM700 confocal microscope.

Imaging of Optical action potentials

CM aggregates were loaded with the voltage-sensitive dye Fluovolt (Molecular Probes, Thermofisher Scientific) as per kit instructions. Fluovolt-loaded aggregates were then resuspended with the following buffer (in mM): 1 MgCl$_2$, 1.8 CaCl$_2$, 5.4 KCl, 140 NaCl, 0.33 NaH$_2$PO$_4$, 10 HEPES, 5 glucose, pH adjusted to 7.4, and warmed to 37 °C. To avoid motion artifacts, blebbistatin (15 μM; Sigma) was added to the buffer. Aggregates were imaged using an Olympus IX73 inverted microscope with a Plan N 4x objective NA = 0.1 (Olympus) and an X-Cite 120 LED light source (Lumen Dynamics). OAPs were acquired at 500 fps using a high-speed, high-sensitivity EM-CCD camera (Evolve 128, Photometrics) operated via Metamorph software (version 7.8.2.0, Molecular Devices LLC). OAPs were first acquired under spontaneous beating to determine spontaneous rate, APD50, and APD90. Aggregates were then imaged during pacing at 0.5 Hz using custom-made parallel wire electrodes and a PowerLab 8/35 system (AD Instruments). OAP recordings were analyzed using custom MATLAB (R2020b; MathWorks) scripts. Briefly, values for the average fluorescence intensity were normalized to the background fluorescence and filtered using a Gaussian-weighted moving average over a 20-element sliding window followed by a linear detrend filter for noise reduction. Fluorescence transients from aggregates that did not show reliable 1:1 capture at 0.5 Hz pacing were disregarded. All oAP results reflect at least 15 recordings from at least 3 independent differentiation experiments.

Sample preparation, single-cell library generation and raw data processing

Mesoderm (day 3), late mesoderm (day 4), progenitor (day 6) and myocyte (day 20) populations were generated from the HES2 hPSC line. Cells were dissociated to single cells as described above and stained with DAPI. Live cells were then sorted using FACSAria Fusion (BD) at the Sickkids/UHN flow cytometry facility. Samples of the same stage (for example, day 3 FHF and day 3 SHF samples) were prepared, handled, and sequenced at the same time to avoid batch effects. Single-cell libraries from these cell suspensions were generated in the 10X Genomics Chromium controller using the Chromium Single Cell 3′ Reagent Kit v3. FHF, aSHF and pSHF cells of the same developmental stage/differentiation phase were sequenced together and all libraries were sequenced simultaneously. Chromium Single Cell Software Suite v3 was used for processing the single cell RNA-seq raw data produced in the 10x Chromium Plat-form, which includes sample demultiplexing, read alignment, barcode processing, and UMI counting. "Cellranger mkfastq" was used to generate FASTQ files from BCL files. Next, "cellranger count" was used to generate single-cell gene counts for a single library. Reads in the FASTQ files were mapped to the human reference genome (GRCH38) with STAR software. Reads were confidently mapped to the exonic loci with MAPQ 255. Chromium cellular barcodes were used to generate gene-barcode matrices. Only reads that were confidently (uniquely) mapped to the transcriptome were used for the UMI count. Filtered gene-barcode matrices containing only cellular barcodes were used for downstream analyses. The data was preprocessed by Princess Margaret Genomics Center.

Cell filtering and cell-type clustering analysis

We captured and sequenced the transcriptomes of 2672 days 3 FHF cells, 5228 days 3 SHF cells, 4108 days 4 FHF cells, 8606 days 4 SHF cells, 3146 days 6 FHF cells, 3241 days 6 aSHF cells, 3231 days 6 pSHF cells (ROH treatment day 4 - day 6), 2352 days 20 FHF cells, 2909 days 20 aSHF cells (ROH and BMP4 treatment day 5 - day 8), and 3535 days 20 pSHF cells (ROH treatment day 4 - day 8). Prior to the downstream analyses, raw data (or filtered count matrix) from cells of the same stage and various lineages (FHF, aSHF and pSHF) were merged as a single object. Further filtering of low-quality cells (based on RNA content and mitochondrial genes) was performed as described in the tutorials (http://satijalab.org/seurat/) (Macosko et al., 2015; Satija et al., 2015). For the analyses of hPSC datasets (day 3, day 4, day 6, and day 20 samples), filtered cells were analyzed using the 'SCTransform' pipeline before principal component analysis (Hafemeister and Satija, 2019). Specifically, 'SCTransform' function was employed and mitochondrial (percent.mt) and cell cycle factors (S.Score and G2M.Score) were regressed out. Next, significant principal components were calculated (RunPCA function) and the top 25 PCAs were used for downstream graph-based, semi-unsupervised clustering into distinct populations (FindClusters function) and uniform manifold approximation and projection (UMAP) dimensionality reduction was used to project these cells in two dimensions (RunUMAP function). For the analyses of mouse progenitor and CM data, the filtered data were normalized for genes expressed per cell and total expression, then multiplied by a scale factor of 10,000 and log-transformed (NormalizeData and ScaleData functions) using the standard Seurat normalization pipeline (the standard normalization method was employed by the authors who published these mouse data). Next, the top 25 PCAs were calculated followed by UMAP dimensionality reduction. For clustering, the resolution parameter was approximated based on the number of cells and distinct marker expressed in discernible clusters. Specifically, resolutions from 0.4 to 0.8 were used for aggregated datasets in the present study (0.5 for the day 3 data, 0.4 for the days 4 and 6 data, and 0.8 for the day 20 data). To identify marker genes or upregulated genes, the clusters of interest were subset and compared for differential gene expression using the Wilcoxon rank-sum test (FindAllMarkers function; only.pos = TRUE, min.pct = 0.1, logfc.threshold = 0.1).

Batch correction and integration of multiple datasets

Batch effects in hPSC-derived cells from multiple stages (Figures 4F and 4G) were corrected by matching mutual nearest neighbors (mnn) with the package batchelor originally described by Haghverdi L et al. (Haghverdi et al., 2018) and adopted by Seurat wrappers. The merged object was normalized and processed with the RunFastMNN function. Note that the object was split by 'timepoint' (for example, day 3 vs. day 4). The integration of human fetal-derived and hPSC-derived data (Figures 2J and S2D, and 6K) and that of CHIR-induced and BMP/ACT-induced cells (Figure S6I-S6J) were achieved by canonical correlation analysis provided in the Seurat package, as described in the manuscript (Stuart et al., 2019). Specifically, PrepSCTIntegration function was used prior to identifying anchors. Next, FindIntegrationAnchors and IntegrateData functions were applied with the normalization.method parameter being set to 'SCT'.

Gene regulatory network analysis

SCENIC was employed to identify gene regulatory networks (or regulons) in mesoderm subtypes (Aibar et al., 2017). The R implemented SCENIC was run on the filtered raw count matrix combining all mesoderm subtypes (geneFiltering function) using the GENIE3 method for gene network (or co-expression modules) reconstruction. The cisTarget motif datasets (hg19-500bp-upstream-7species.mc9nr.feather, hg19-tss-centered-10kb-7species.mc9nr.feather) were used to construct regulons for each transcription factor. Potential direct-binding targets were then selected based on DNA-motif analysis with the RcisTarget package. Cellular enrichment of the network activity in each cell was subsequently assessed by AUCell. For visualization, the network activity was converted into binary activity matrix ('ON'/'OFF'), and only the 'ON' cells that passed the AUC thresholds were shown in the t-Distributed Stochastic Neighbor Embedding (tSNE) plots.

Gene ontology, gene set enrichment analyses, and human to mouse gene transformation

GO and gene set enrichment analyses were performed with the ClusterProfiler, MSigDB, org.Hs.eg.db and org.Mm.eg.db packages (Carlson et al., 2016; Subramanian et al., 2005; Yu et al., 2012). Upregulated genes in the clusters of interest were obtained using the FindAllMarkers function provided in the Seurat package (only.pos = TRUE, min.pct = 0.1, logfc.threshold = 0.1). The gene symbols were transformed to Entrez Gene IDs using the annotations provided in the org.Hs.eg.db package. To perform one to one transformation between human and mouse ortholog (analysis of human and mouse conserved genes), useMart and getLDS functions provided in the biomaRt package were used (Durinck et al., 2005, 2009). The genes enriched in these clusters were subsequently analyzed for the enrichment of biological processes (BP) using the enrichGO function (ontology = BP, pvalueCutoff = 0.05, OrgDb = org.Hs.eg.db) with the package ClusterProfiler. Heatplot and cnetplot functions provided in the Enrichplot package together with ggplot2 were used to visualize the results of GO analyses. For GSEA analysis, all the genes expressed in each cluster were retrieved and then ranked by their expression level (LogFC). With the ranked data frame, enrichment analyses using custom gene set, KEGG annotations, or BP annotations were performed with the package ClusterProfiler. Finally, the selected terms/gene sets were visualized with the package Enrichplot (gseaplot2 function).

Pseudotime trajectory analysis

Pseudotime and cell trajectory analyses shown in the Figures 2K and 4F were performed with the slingshot package as described in the paper and tutorials (Street et al., 2018). The Seurat object was transformed to a single cell experiment object which was then analyzed with the slingshot package (slingshot and getLineages functions). Note that the start cluster was pre-determined for the slingshot analysis (epiblast for the integrated mesoderm data). The resulting pseudotime values of all the cell types were added to the metadata of their corresponding Seurat objects for downstream visualization with ggplot2.

Machine learning-based classification

Machine learning-based classification of hPSC-derived cells was performed with the scPred package as described by Jose Alquicira-Hernandez et al. (2019). The models (mixture discriminant analysis (MDA) and support vector machine (SVM) in this study) were trained using the selected datasets as references (getFeatureSpace and trainModel functions). The model type, ROC and specificity were indicated in the figures. Subsequently, the query data were analyzed for classification of cell types based on the models (scPredict function; recompute_alignment = FALSE, default threshold = 0.55). The outcome of scPred analysis was transferred to the metadata of the corresponding Seurat object for visualization.

Intercellular communication analysis

Intercellular communication network analysis was performed with the package CellChat which predicts major signaling inputs and outputs for cells and how those cells and signals coordinate for functions using network analysis and pattern recognition approaches (Jin et al., 2021). RNA data matrices were retrieved and transformed to CellChat objects. Over-expressed ligands or receptors were estimated using the identifyOverExpressedGenes and identifyOverExpressedInteractions functions provided in the package, and the gene expression data was projected onto a protein-protein interaction (PPI) network. Subsequently, communication probability was inferred using the functions computeCommunProb and computeCommunProbPathway. Finally, the results were visualized using the netVisual_signalingRole and netVisual_aggregate functions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Standard statistical analyses were performed using GraphPad Prism 8. The number of replicates, type of statistical test and test results are described in the figure legends. All data are represented as mean ± SE of mean (SEM). Statistical significance of two group comparisons was determined by unpaired Student's t test and that of three or more groups was determined by one-way ANOVA analysis with Bonferroni post-hoc test in GraphPad Prism 8 software. Results are significant at $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****). Sample size of all the experiments was not pre-determined, and no randomization or investigator blinding approaches were implemented during the experiments and data analyses given the nature of the study.